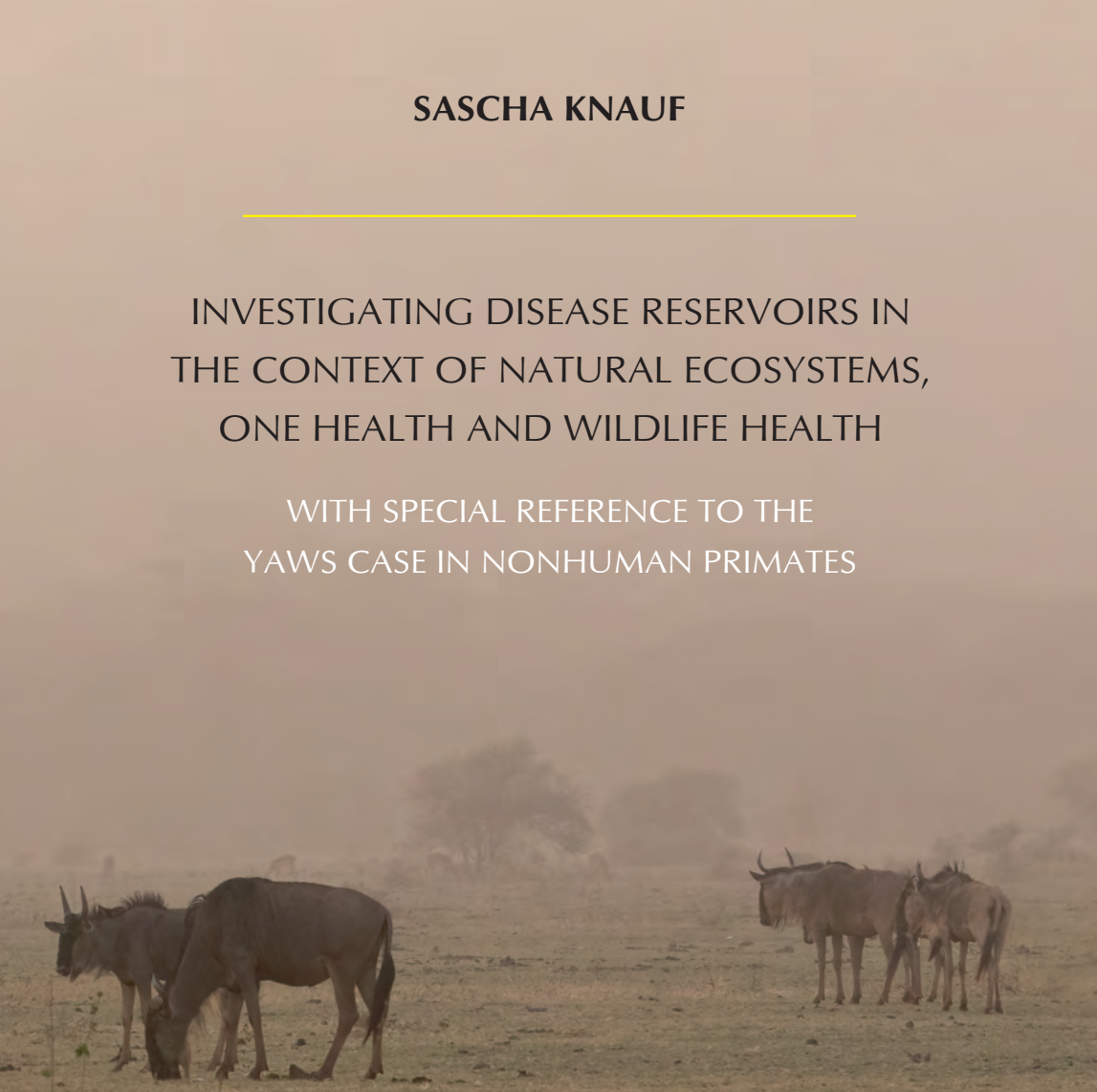
SASCHA KNAUF

INVESTIGATING DISEASE RESERVOIRS IN THE CONTEXT OF NATURAL ECOSYSTEMS, ONE HEALTH AND WILDLIFE HEALTH

WITH SPECIAL REFERENCE TO THE YAWS CASE IN NONHUMAN PRIMATES



HABILITATIONSSCHRIFT
ZUR ERLANGUNG DER LEHRBEFÄHIGUNG FÜR
DIE FACHGEBIETE WILDLIFE HEALTH UND ONE HEALTH
IM FACHBEREICH VETERINÄRMEDIZIN DER
JUSTUS-LIEBIG-UNIVERSITÄT GIEßEN

édition scientifique
VVB LAUFERSWEILER VERLAG

1. Auflage 2021

1st Edition 2021

Printed in Germany

édition scientifique
VVB LAUFERSWEILER VERLAG

STAUFENBERGRING 15, 35396 GIESSEN, GERMANY
Tel: 0641-5599888 Fax: 0641-5599890
email: redaktion@doktorverlag.de

www.doktorverlag.de

Justus-Liebig-Universität Gießen

Fachbereich Veterinärmedizin

Investigating Disease Reservoirs in the Context of Natural Ecosystems, One Health and Wildlife Health

With Special Reference to the Yaws Case in Nonhuman Primates

Habilitationsschrift

zur Erlangung der Lehrbefähigung für

die Fachgebiete Wildlife Health und One Health

im Fachbereich Veterinärmedizin der

Justus-Liebig-Universität Gießen

vorgelegt von

Tierarzt **Sascha Knauf**, PhD

Gießen 2021

Live at its best is an adventure,

a voyage of discovery.

- Peter C. Doherty

To Yvonne, Ella Siv Aina,

Richard and Darwin

“Ich erkläre: Ich habe die Habilitationsschrift selbstständig und nur mit den Hilfen angefertigt, die ich in der Arbeit angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei der Erstellung der Habilitationsschrift und bei den von mir durchgeführten und in der Arbeit erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der ‘Satzung der Justus-Liebig-Universität zur Sicherung guter wissenschaftlicher Praxis’ umschrieben sind, eingehalten.”

Sascha Knauf

Explanatory remarks

The thesis submitted consists of 19 original publications and additional five reviews have been incorporated. The main publication has been included. Where it applies, supplementary datasets and materials can be found at the respective Journal's repository as indicated in the publication.

This thesis includes five first authorships, ten last authorships, and 14 intermediate authorships. Details about the respective contribution can be found under 12. Author Contributions.

CONTENTS

1. Introduction

Current evidence suggests that preserving intact ecosystems and their endemic biodiversity should generally reduce the prevalence of infectious diseases (Johnson et al., 2019; Keesing et al., 2010). The world human population is, however, expected to increase to 9.6 billion in 2050 (Gerland et al., 2014) which furthers habitat destruction and competition for resources. Accelerating population growth rates are seen in areas that are classified as biodiversity hotspots, which are areas especially rich in endemic species (Cincotta et al., 2000). Yet, the global rise in human infectious disease outbreaks that are frequently associated with zoonotic pathogens (Jones et al., 2008; Smith et al., 2014) is a symptom of a growing human-wildlife-livestock interface (Karesh et al., 2012). There is, therefore, an increased risk of disease transmission from wildlife to humans and livestock and vice versa. Wildlife diseases can have a significant economic impact and many of them have the potential to become a major public health concern (Daszak et al., 2000).

Although human-introduced diseases can likewise become a significant health threat to wild animals (Dobson and Foufopoulos, 2001), this thesis will focus largely on diseases in wildlife that are relevant to human health. The first six chapters provide a framework for the different aspects that are pertinent to understand and investigate diseases in the context of natural ecosystems. Subsequently, chapter 7 applies this knowledge to investigate *Treponema pallidum* infection in wild nonhuman primates (NHPs) as a potential source for human yaws infection.

2. Basic Principles of Wildlife Disease Ecology

2.1 Natural Ecosystems and the Host-Pathogen Equilibrium

An ecosystem is defined as "all the living things in an area and the way they affect each other and the environment" (https://dictionary.cambridge.org/dictionary/english/ecosystem).

Microorganisms are an integral part of natural ecosystems and contribute to biodiversity. A microbe that is capable of causing host damage is defined as a pathogen (Casadevall and Pirofski, 1999). Yet, the combined biomass of pathogens within an ecosystem can exceed the biomass of top predators (Kuris et al., 2008). Ecosystems are balanced, which means there is a host-pathogen equilibrium within a given community of organisms. Intact ecosystems are therefore characterized through relative stability in diversity with gradual changes through ecological succession. Similar to predators, pathogens can regulate the trophic cascade (Buck and Ripple, 2017; Hudson et al., 2006; Roche and Guégan, 2011) and therefore contribute to the health of ecosystems. The disruption of the host-pathogen equilibrium can impact disease dynamics within a given host community. The effect becomes visible e.g., in the absence of predators where increasing numbers of herbivores are first bottom-up regulated through the availability of resources. The increase in population density leads to increased contact rates and subsequent exposure and transmission rates. At this stage, the population is at large regulated by virulent pathogens. A practical example of such an effect is tuberculosis caused by *Mycobacterium bovis* in wild boar (*Sus scrofa*). In the absence of an apex-predator, here the grey wolf (*Canis lupus*), tuberculosis prevalence has been estimated four-fold higher than in areas where the wolf prompted a top-down regulation of the trophic cascade (Tanner et al., 2019). The reestablishment of a top-down regulated trophic cascade, therefore, contributes to

the reduction of prevalence and subsequently possible disease spillovers from wildlife reservoir systems to livestock and humans.

2.2 The Evolution of Virulence

A microbe that is capable of causing host damage is defined as a pathogen and virulence is defined as the microbe's capacity to damage its host (Casadevall and Pirofski, 1999). The impact that a pathogen can have on the population level is largely based on its impact on the individual level, which in turn is influenced by the pathogen's virulence factors. Virulence factors are defined as components of a pathogen that damage the host (Casadevall and Pirofski, 1999). Factors can include toxins, bacterial capsules to produce endospores, adherence factors, serum or antimicrobial resistance factors, or enzymes that enable fibro-lysis. In bacteria, these factors may be coded on plasmids that can be easily transferred between different organisms.

Hosts and their pathogens are connected through an evolutionary arms race, that often strives to reduce susceptibility on the host's side and to maintain infectiousness on the pathogen's side. These processes can become visible when hosts and pathogens coevolve (Burmeister et al., 2016; Gómez et al., 2014). After a long history of coevolution, herpesviruses, for example, cause little harm to their natural host (Adler et al., 2017). However, transmission to non-adapted hosts can occur with often fatal consequences as shown for the macacine herpesvirus 1 (Huff and Barry, 2003) or the suid herpesvirus 1 (Sawitzky, 1997). The different *Leptospira* spp. are another example of pathogen-host adaptation. The *Leptospira* genus has co-evolved with its maintenance host, e.g. bats (Lei and Olival, 2014) where the pathogen causes almost no pathogenicity (Gomes-Solecki et al., 2017).

Understanding the evolution and mechanisms of virulence can help to identify pathogen candidates that have the highest potential to impact population health. Contrary to the idea of co-evolution and the host-pathogen equilibrium, some pathogens have maintained their virulence over millennia (Hoeprich, 1989). Although these pathogens have adapted to their host, they are associated with high pathogenicity. Coevolution and host-adaptation, therefore, do not always predict mild disease in a host-pathogen system (Anderson and May, 1981; Ewald, 1998, 1995, 1996, 1991a, 1991b, 1983; Frank, 1996; Lenski and May, 1994; Levin and Pimentel, 1981; May and Anderson, 1983). Pathogen competition within a single host likely selects those genotypes that reproduce more rapidly and therefore increases virulence. On the downside, pathogens that are directly transmitted benefit from the selection of benign genotypes, as mild diseases ensure stable contact rates between infected and susceptible hosts. Since the costs of high host mortality are relatively low for pathogens that can survive in the environment until they can invade and infect a new host (Ewald, 1994, 1987), it is suggested that pathogen survival in the environment is linked to increased virulence (Walther and Ewald, 2004). As a proof of concept, Walther and Ewald found a significant correlation between durability and mean mortality rates (as a measure for virulence) in respiratory tract pathogens of humans (Walther and Ewald, 2004). While the authors were not able to elaborate cause and effect, their results are supportive for the sit-and-wait hypothesis (Ewald, 1994, 1987) that argues for increased virulence in pathogenic microbes that have high tenacity. As a consequence, environmental contamination contributes to a reservoir system of pathogens that developed the capability to survive in the environment. An example is the bacterium *Bacillus anthracis*, which causes fatal anthrax disease in wildlife, livestock and humans (Hoffmann et

al., 2017). The bacterium produces highly resistant endospores that can stay infective for years (Nicholson et al., 2000).

2.3 Ecological Implications of Pathogens That Manipulate their Host

In addition to virulence factors, microbes have developed a range of strategies to manipulate their hosts for their advantage and dispersal (Heil, 2016). The majority of manipulating pathogens use a trophic mode of transmission in preference of a horizontal transmission mode (Heil, 2016). Manipulation of the host can be as 'simple' as inducing hypersensitivity in the upper airways to promote cuffing and sneezing and subsequent spread of the pathogen as it is seen in respiratory diseases (Richard et al., 2017) or the painlessness of skin ulcers at the genital to maintain sexual activity as seen in syphilis caused by the bacterium *Treponema pallidum* ssp. *pallidum* (*TPA*) (Lafond and Lukehart, 2006). More subtle and complex manipulation tactics of pathogens involve gradient changes in behaviour. A paramount example of such a host manipulation is the infection with the protozoon *Toxoplasma gondii*. It has been reported, that infected prey species increase their exposure to feline predators (Poirotte et al., 2016), are sexually more attractive to susceptible hosts (Dass et al., 2011), and (even infected humans) are more risk-taking (Flegr, 2007; Johnson et al., 2018). Thus, *T. gondii* manipulates its host in ways that are beneficial for the completion of its prey-predator driven lifecycle. Yet, host manipulation can also be indirect. Pathogens of high virulence, such as *B. anthracis*, rapidly kill their host and create conditions that do not favour the bacterium's survival in its vegetative form. Instead, endospores of high environmental resistance are produced (sit-and-wait pathogen) (Dittmann et al., 2015). The decomposing body of the host provides a substantial nutritional resource for plant growth, which attracts susceptible herbivorous species to visit sides of high *B. anthracis* contamination (Turner et al., 2014). This boosts host exposure rates and subsequently increases the incidence rate of anthrax in the herbivorous host.

On the population scale, the manipulation of the host's phenotype can significantly impact population health. In geladas (*Theropithecus gelada*), monkeys endemic to the alpine grasslands of Ethiopia, we could show that infection with tapeworms of the genus *Taenia* and the resulting coenurosis (geladas function as the intermediate host) adversely impact gelada survival and reproduction of a course of at least two host generations (Nguyen et al., 2015). In adult geladas, infection increased mortality rates and interbirth intervals. Offspring of females with coenurosis had a reduced chance of survival and adult males were less likely to gain reproductive control over units of females.

When ecologically important hosts are the target of infection, effects become clear on the ecosystem level. The nematode *Physocepahlus sexalatus* has been used as an example to illustrate these processes (Weinersmith and Faulkes, 2014). The parasite infects dung beetles of the genus *Phanaeus* and causes reduced uptake of faeces which is its main source of food. High disease prevalence in dung beetles, therefore, leaves a higher rate of faecal-boli undigested. Since dung beetles are known to reduce the number of infectious helminths in faecal material through digestion (Nichols and Gómez, 2014), high numbers of undigested faeces in the environment increases the risk for all other host species within the ecosystem to get into contact with potentially infectious material.

Examples of host manipulation and phenotypic changes are numerous. Interactions between the different pathogens in an ecosystem, but also within a single host, challenge our

understanding of diseases. Due to the complexity of natural ecosystems, such knowledge is often missing, although it is crucial for the understanding of the pathogen's biology, population and community ecology (Weinersmith and Faulkes, 2014).

2.4 Host Factors that Influence Disease Susceptibility and Transmission

To be successful, pathogens need to be able to invade, adapt and survive in their host(s). Several host factors can influence disease susceptibility, exposure and infection (DeCandia et al., 2018). These factors are responsible for the degree to which a host can avoid infection. Factors can include genotype, nutritional status, social behaviour and the immune system (Katz et al., 2014). In general, there are two lines of host defence. The first and most rigid defence line that a pathogen needs to overcome to convert an exposed and susceptible individual into an infected host is the behavioural defence. If for example in sexually transmitted diseases a healthy subject avoids contact by not mating with an infected partner, it will not acquire the disease or in the case of sit-and-wait pathogens, behavioural defence could keep susceptible hosts away from areas with high environmental pathogen contamination.

The second line of host defence is the host's immunological reaction to an invading pathogen. In mammals, the immune system consists of two main branches. The innate immune response is the more ancient branch and has an invertebrate origin (Leulier et al., 2003). It develops within hours after invasion of an infectious organism, but is of limited specificity and exhibits only limited immunological memory for the pathogen that was previously encountered (Koenderman et al., 2014). Independent of the respective route of pathogen invasion (skin, mucosa, respiratory or gastrointestinal), the pathogen is at first confronted with a continuous epithelial layer that it needs to actively (virulence factors) or passively (e.g., skin abrasions) overcome. Further host factors that challenge the pathogen's entry into the host organism are for example the pH (e.g., in the stomach (Smith, 2003)), the production of lysozyme containing mucus (Gerson et al., 2000) and ciliary surface activity (e.g., respiratory tract (Bustamante-Marin and Ostrowski, 2017)). Pathogens that overcome these physical and chemical barriers are subsequently confronted with some antimicrobial factors. These factors include complement, lysozyme, interferons and lactoferrin. Cellular response to the invading pathogen comes from leucocytes. Monocytes, macrophages, neutrophils and dendritic cells can phagocytize and enzymatically digest microbes. Different kinds of toxic chemicals are produced and released by activated effector cells such as neutrophilic granulocytes in response to a pathogen (Kobayashi et al., 2018). Through cytokine release, leucocytes can orchestra a pro-inflammatory response that attracts more effector cells to the side of inflammation (Zhang and An, 2007) and which helps to control the spread of the pathogen within a host. The innate immunity is inherited and germline-encoded (Janeway and Medzhitov, 2002; Kumar et al., 2009). Toll-like receptors (TRLs) are an evolutionarily ancient recognition and signaling system with a diverse set of receptor proteins devoted to the recognition of molecular patterns that are not found in normal vertebrates (e.g., lipoproteins, zymosan of yeast, flagellin, dsRNA, ssRNA, or unmethylated CpG DNA) (Murphy et al., 2008a). TRL-receptors are expressed on sentinel cells such as macrophages and dendritic cells and among others mediate cytokine release. They recognize patterns that are characteristic of components of pathogenetic microorganisms but the receptor recognition is not as specific as in the case of the antigen immune receptors of the adaptive immune system (Akira and Takeda, 2004).

A hallmark of the adaptive immune system, which is the second branch of the mammalian immune system, is its specific recognition of antigen and the immunological memory function (Bonilla and Oettgen, 2010). B-lymphocytes are responsible for the production of antigen-specific antibodies (humoral immunity), whereas the T-lymphocytes are important regulators for cell-mediated immunity. Both, T- and B-lymphocytes, are the only cells in the body which are capable of recognizing specific antigenic determinates of foreign antigen (immunological memory function).

The impact of commensal microorganisms on disease susceptibility is yet not fully understood, but evidence accumulates that commensal bacteria act with the host to prevent colonization by invasive pathogens (Abt and Pamer, 2014; Buffie et al., 2015; Dennison et al., 2014). The human gut alone is colonized by an estimated number of 10^{14} microorganisms (Gill et al., 2006). The different niches within a host are ecosystems on their own and microorganisms within those ecosystems compete and cooperate for resources (McNally and Brown, 2015; van de Wijgert, 2017). They share and change their environment to a remarkable extent and are therefore recognized as prodigious 'nice constructors' (McNally and Brown, 2015). Similarly, the vaginal environment in healthy women is dominated by benign microorganisms (lactobacilli) that help to lower the pH and protect against sexually transmitted diseases including human immunodeficiency virus (HIV) acquisition (van de Wijgert, 2017). Due to the relatively recent innovations in sequencing techniques, microbiome analyses are becoming more and more popular, which will help to gain further insight into the host's microbiome in health and diseases.

2.4.1 Factors Determining Human-to-Human Transmissibility of Zoonotic Pathogens via Contact

The following review provides further details on factors that promote human-to-human pathogen transmission. It underlines the need to comparatively study common and specific factors that promote transmission as these factors are critical to assess the risk for zoonotic pathogens to establish a human-to-human cycle and to mitigate measures that are needed to control disease outbreaks.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

Factors determining human-to-human transmissibility of zoonotic pathogens via contact

Mathilde Richard[1], Sascha Knauf[2], Philip Lawrence[3,4], Alison E Mather[5], Vincent J Munster[6], Marcel A Müller[7], Derek Smith[8] and Thijs Kuiken[1]

The pandemic potential of zoonotic pathogens lies in their ability to become efficiently transmissible amongst humans. Here, we focus on contact-transmitted pathogens and discuss the factors, at the pathogen, host and environmental levels that promote or hinder their human-to-human transmissibility via the following modes of contact transmission: skin contact, sexual contact, respiratory contact and multiple route contact. Factors common to several modes of transmission were immune evasion, high viral load, low infectious dose, crowding, promiscuity, and co-infections; other factors were specific for a pathogen or mode of contact transmission. The identification of such factors will lead to a better understanding of the requirements for human-to-human spread of pathogens, as well as improving risk assessment of newly emerging pathogens.

Addresses
[1] Department of Viroscience, Postgraduate School Molecular Medicine, Erasmus MC, Rotterdam, The Netherlands
[2] Work Group Neglected Tropical Diseases, German Primate Center, Leibniz-Institute for Primate Research, Göttingen, Germany
[3] Université de Lyon, UMRS 449, Laboratoire de Biologie Générale, Université Catholique de Lyon – EPHE, Lyon 69288, France
[4] Molecular Basis of Viral Pathogenicity, International Centre for Research in Infectiology (CIRI), INSERM U1111 – CNRS UMR5308, Université Lyon 1, Ecole Normale Supérieure de Lyon, Lyon 69007, France
[5] Department of Veterinary Medicine, University of Cambridge, Cambridge, UK
[6] Virus Ecology Unit, Laboratory of Virology, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rocky Mountain Laboratories, Hamilton, MT, USA
[7] Institute of Virology, University of Bonn Medical Center, Bonn, Germany
[8] Department of Zoology, University of Cambridge, Cambridge, UK

Corresponding author: Kuiken, Thijs (t.kuiken@erasmusmc.nl)

Current Opinion in Virology 2017, **22**:7–12

This review comes from a themed issue on **Emerging viruses: intraspecies transmission**

Edited by **Ron** and **Linfa**

http://dx.doi.org/10.1016/j.coviro.2016.11.004

Introduction

Infectious diseases are the second leading cause of death worldwide after cardiovascular diseases [1]. More than half of the known pathogens that are able to infect humans are of zoonotic origin [2]. Once a zoonotic pathogen has crossed the species barrier by infecting humans, its success in the human population will depend on whether or not it can acquire the ability of sustained human-to-human (H2H) transmissibility. A better understanding of the factors that determine this ability would help to prevent the emergence or re-emergence of infectious diseases in the human population.

Transmission of infectious pathogens amongst humans can occur via multiple routes: airborne (aerosols and respiratory droplets) route, faecal-oral route, contact route or vector-borne route. In this review, we focused on pathogens that are transmitted via direct or indirect contact as their main or substantial routes of transmission. Pathogens that are mainly transmitted via the faecal-oral and food-borne routes — which also are types of contact transmission — were excluded because they are discussed elsewhere in this issue. Direct contact transmission requires physical contact between an infected person and a susceptible person and the transfer of pathogens via touching, sexual contact, or contact with bodily fluids or lesions. Indirect contact refers to the infection of a susceptible person via a contaminated surface. We divided contact transmission into four modes: skin, sexual, respiratory and multiple. We used the following examples to illustrate these four modes of contact transmission: *Treponema pallidum pertenue* (TPE) for skin contact transmission, human immunodeficiency virus type 1 (HIV-1) for sexual contact transmission, coronaviruses (CoV) for respiratory contact transmission and Ebola virus for contact transmission via multiple routes. For each of these pathogens and their specific mode of transmission, we identified the factors, at the level of the pathogen, host or environment that promoted or hindered their ability of sustained H2H transmissibility.

Skin contact transmission

The spirochete bacterium *Treponema pallidum* (ssp. *pertenue*; TPE) causes yaws. Another subspecies (ssp. *pallidum*) that causes syphilis is not further discussed here. Yaws primarily affects the skin, bones and cartilages of children in hot and humid areas of Africa and Asia and the

Pacific region. The main sources of infection are direct contact with skin ulcers.

Although TPE is traditionally considered to exclusively infect humans, it has recently been identified in African nonhuman primates. The fact that human and simian TPE strains share a high degree of genetic and functional similarity suggests that African nonhuman primates may serve as a reservoir for human infection and highlights the potential for zoonotic transmission [3].

TPE has obviously acquired the ability of sustained H2H transmissibility. There are several pathogen factors that may contribute to this ability. The spirochete evades the immune response by antigenic variation and abrogation of opsonizing antibodies [4–6], allowing it to survive permanently in the infected host. It reaches high loads in skin ulcers and the infectious dose is low. Besides entering a new host via cuts or abrasions, TPE uses peptides of the outer membrane to attach to host surface proteins [7] and is able to penetrate healthy mucous membranes [8]. Host factors favouring TPE transmission include crowded living conditions. High humidity and temperature are environmental factors that increase TPE survival outside the host. Finally, lack of surveillance and inadequate health care favour the persistence and spread of human yaws in affected countries.

Sexual transmission

The retrovirus HIV-1 is the archetypal example of a sexually transmitted human pathogen. Although HIV-1 can be contracted by sexual, percutaneous and perinatal routes, nearly 70% of infections worldwide result from heterosexual intercourse [9]. HIV-1 is the causative agent of Acquired Immune Deficiency Syndrome (AIDS) in humans, characterised by severe depletion of memory CD4+ T-lymphocytes early following infection, leading ultimately to immunodeficiency and death due to opportunistic infections and rare diseases [10]. Sexual transmission involves the transfer of virus particles or infected cells present in contaminated genital secretions or blood from an infected person to the mucosa of a susceptible host [11•]. Following transmission, the successfully transmitted founder virus population is established in CD4+ cells in mucosa/submucosa, draining lymphatics, gut-associated lymphoid tissue and systemic lymphatic tissues. Viraemia follows and increases exponentially as a result of massive virus replication in gut associated and other peripheral lymphoid tissue [11•].

HIV-1 likely originated from nonhuman primates at some time in the twentieth century. The most genetically similar and related primate lentivirus described to date is the simian immunodeficiency virus (SIV) found in chimpanzees in central Africa (SIVcpz) [12]. The majority of nonhuman primate species appear afflicted with a single strain of SIV that is mostly non-pathogenic in its natural host.

Important restriction factors by which infected hosts control lentiviral infection are tetherin [13], APOBEC3G [14] and TRIM5α [15]. The ability of the most prevalent strain of HIV-1, the M strain, to overcome such restriction factors is believed to have been critical to establish an infection in humans and to allow sustained H2H transmission, leading to the current global pandemic [16]. In comparison, the reduced abilities of other HIV-1 strains (N, O and P) and HIV-2 to counteract these restriction factors [17–19] may partly explain why they were not able to spread so effectively within the human population. Other pathogen factors that have contributed to the 'success' of HIV-1 M strain as a human pathogen, despite its relatively low infectivity (risk estimate of 1 in 1000 exposures for heterosexual transmission; [9]), include its extraordinary propensity to evolve its genome through recombination and low-fidelity replication, allowing immune and therapeutic escape [20], the nature of its long, 'latent', often sub-clinical infection, during which patients can transmit the virus [21], and high viral load. Host factors that favour transmission are the presence of other sexually transmitted diseases [9], as well as promiscuous sexual behaviour.

Respiratory contact transmission

Of the six known human (CoV), severe acute respiratory syndrome CoV (SARS-CoV) [22] and Middle East respiratory syndrome CoV (MERS-CoV) [23] are responsible for high morbidity and mortality in infected individuals. The other four human CoV (HCoV-229E, NL63, OC43, HKU1) have low pathogenicity and are associated with seasonal common colds [24].

The zoonotic origin of four out of six human CoV has been elucidated. HCoV-229E, SARS-CoV and MERS-CoV originate from bats and HCoV-OC43 from bovids, whereas animal ancestors for HCoV-HKU1 and NL63 are still to be found [25•]. The common cold CoV likely emerged a long time ago in the human population, as reflected by a global distribution and a high prevalence in humans [26]. In contrast, intermediate host species such as Himalayan palm civets [27] and dromedaries [28••] likely played a role in the recent introduction of SARS-CoV and MERS-CoV, respectively, in the human population.

SARS-CoV and MERS-CoV are predominantly transmitted via direct H2H contact, droplets and fomites [29–31], and have not (yet) established long-term and sustained H2H transmission. Virus replication occurs mainly in the lower respiratory tract (LRT) in type II pneumocytes and alveolar macrophages [32,33••,34]. Replication in the LRT may be explained by the protein expression profile of the respective receptors, the exopeptidases

angiotensin-converting enzyme 2 (ACE2) in case of SARS-CoV [35] and dipeptidyl peptidase 4 (DPP4) for MERS-CoV [36,37]. In addition, antiviral immunity of the epithelium may reduce viral replication in the upper respiratory tract (URT) [38]. For contact transmission of SARS-CoV and MERS-CoV between humans, the quantity of infectious particles seems to be an important factor as high viral loads in patients facilitated H2H transmission [39,40]. Pronounced stability of infectious CoV on surfaces for up to several days [41] could also explain fomite-related transmissions, a phenomenon that may contribute to superspreading events [39,42].

In contrast to SARS-CoV and MERS-CoV, the four common cold CoV are predominantly droplet-transmitted [43] and efficiently H2H transmissible. Virus replication occurs mainly in the central and upper parts of the respiratory tract. For HCoV-229E, this may be explained in part by the abundant expression of its entry receptor, aminopeptidase N), on non-ciliated cells of the bronchial epithelium [44]. However, although HCoV-NL63 uses the same entry receptor as SARS-CoV (ACE2) [45], it replicates mainly in the URT, perhaps because it uses additional attachment factors like heparan sulfate proteoglycans [46], or because of other as yet unknown viral replication-related or immune-related factors.

Comparison of these two groups of CoV suggests that URT replication and droplet transmission (common cold CoV) is more advantageous for sustainable H2H transmission than LRT replication and direct contact transmission (SARS-CoV and MERS-CoV). Replication in the URT, as well as transmission via the respiratory route, are also factors that favour the efficient H2H transmission of human influenza viruses, as compared to zoonotic avian influenza viruses [47].

In conclusion, high expression levels of a suitable receptor molecule in the URT combined with efficient and probably well-balanced viral countermeasures against local immunity may be major pathogen factors for zoonotic CoV to attain successful and sustained H2H transmission, as exemplified by the common cold CoV.

Multiroute transmission

Of the five known Ebola virus species, four are known to cause disease in humans [48]. Infection with Ebola virus causes Ebola virus disease (EVD), which is an acute systemic illness with a high case fatality rate [49].

Of all potential transmission routes, direct contact with patients or bodily fluids from these patients, as well as contact with contaminated surfaces or materials, is considered the most important [50,51]. Ebola virus has indeed been isolated from several bodily fluids such as blood, breast milk and semen of infected patients. In addition, Ebola virus RNA has been detected in sweat, tears, stool, on skin and from vaginal and rectal samples [52]. During the 2014 outbreak, several researchers speculated about the potential for airborne transmission of Ebola virus [53]. However, the majority of EVD patients in previous outbreaks were infected by contact transmission and all EVD outbreaks, including the 2014 epidemic, have been contained without measures against airborne transmission in the general population. This suggests that extensive airborne transmission is unlikely and of limited epidemiological importance. The 2014 EVD outbreak revealed the potential for Ebola virus to be sexually transmitted. Infectious virus and Ebola virus RNA have been detected in semen from male EVD survivors as long as 70 and 270 days, respectively, after recovery from the initial infection [54•]. This detection of infectious Ebola virus and viral RNA months after recovery from EVD highlights the potential for Ebola virus to seed new outbreaks after patients with clinical EVD are no longer present and an area is declared Ebola virus free.

The putative animal reservoirs for Ebola viruses are bats, and zoonotic transmission is thought to occur either by direct contact with bats, or via indirect transmission by contact with bats, or via indirect transmission by contact with other infected wildlife species, such as gorillas, chimpanzees or duikers, which — like humans — are affected by Ebola virus [55].

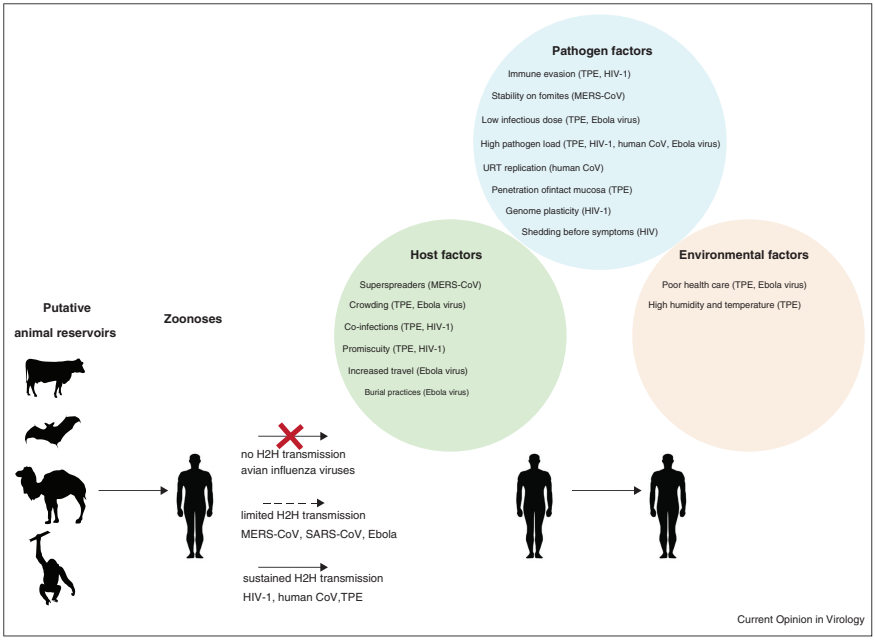
Amongst the pathogen factors that promote H2H transmission of Ebola virus is the high virus load in secreted bodily fluids combined with a very low infectious dose, as low as 10 plaque forming units as measured in experimental infection studies in nonhuman primates [56]. Amongst host factors, ancestral funeral and burial practices of deceased EVD patients, in which levels of Ebola virus remain high after death, have been identified as a major source of human infection [57]. Moreover, the 2014 outbreak of EVD in West Africa, caused by the Zaire Ebola virus, has shown, for the first time, the ability of Ebola virus to cause a long-term large-scale epidemic with sustained H2H transmission [58••]. In addition to EVD cases in Guinea, Sierra Leone and Liberia, travel-associated cases with subsequent nosocomial transmission have been reported in Mali, Nigeria and the United States. The 2014 Ebola virus strains were relatively closely related to viral strains from the previous two Zaire Ebola virus outbreaks in Democratic Republic of Congo, and, although the evolution rate of the genome of the Ebola virus during the 2014 outbreak was higher than the between-outbreak rate, the virus did not change substantially [58••,59••]. The clinical course, for example, incubation time, symptoms and development of the disease, as well as the transmissibility of the virus (R_0, basic reproductive number) were not different from those in past outbreaks of Ebola virus. Most likely, the unprecedented epidemic of Ebola in 2014 was the result of a

combination of human behavioural and societal factors [60]. Firstly, West African countries never had experienced an EVD outbreak before, other than a single case of Taï Forest ebolavirus infection in the 1980s in Ivory Coast. In addition, Guinea, Sierra Leone and Liberia are amongst the poorest countries in the world, with impaired public health infrastructures. Moreover, compared to previous outbreaks of Ebola virus, the virus was not confined to remote and rural areas and the outbreak spread into large population centres, such as Monrovia, Conakry and Freetown. The spatial connectivity provided by roads and a travelling population allowed for the rapid dissemination of Ebola virus over these three countries, before a targeted international response was initiated.

Concluding remarks

Despite being categorised under one heading, contact-transmitted pathogens may differ substantially in their specific modes of transmission: via skin, via genital mucosa, via respiratory mucosa, or via several of these modes. Nonetheless, several factors were identified that were common amongst at least two modes of transmission. Therefore, it is important to identify both factors promoting H2H transmission that are common amongst contact-transmitted pathogens and factors that are specific for each mode of contact transmission (Figure 1). Common pathogen factors were immune evasion, high viral load, and low infectious dose. Common host factors were crowding, promiscuity, and co-infections. Other factors were specific to one of the modes of transmission and the

Figure 1



Factors, at the pathogen, host and environmental levels, that promote human-to-human contact transmission of human pathogens of zoonotic or putative zoonotic origin. The transmissibility of pathogens of zoonotic origin determines their pandemic potential. Common factors, as well as specific factors, that promoted the transmissibility via contact amongst humans of the following pathogens via the following routes are described and categorised under pathogen, host and environmental factors: *Treponema pallidum pertenue* for skin contact transmission, human immunodeficiency virus type 1 for sexual contact transmission, coronaviruses for respiratory contact transmission and Ebola virus for contact transmission via multiple routes. The pathogen to which these factors refer is indicated between brackets. *Abbreviations*: H2H: human-to-human; TPE: *Treponema pallidum pertenue;* CoV: coronavirus; MERS-CoV: Middle East respiratory syndrome CoV; SARS-CoV: severe acute respiratory syndrome CoV; HIV-1: human immunodeficiency virus type 1.

pathogen described. Specific factors that may be critical for efficient H2H transmission are high viral load in skin lesions for skin contact transmission, promiscuous sexual behaviour for sexual contact transmission, URT replication and a switch from contact to aerosol transmission for respiratory contact transmission and burial practices for the transmission of Ebola virus. Identification of these factors is critical to assess the risk of contact-transmitted zoonotic pathogens gaining efficient H2H transmissibility, and to implement mitigation measures and large scale prevention campaigns in case of outbreaks.

Acknowledgements

The expert workshop was financially supported by European FP7 programme ANTIGONE (ANTIcipating the Global Onset of Novel Epidemics, project number 278976). MR's research is partly supported by NIAID/NIH contract HHSN272201400008C. SK's research is partly supported by grants of the German Research Foundation (DFG): KN1097/3-1 and KN1097/4-1. AEM is supported by a Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/M014088/1. VJM is supported by the Division of Intramural Research of the NIAID/NIH. Marcel A Müller is supported by the Zoonoses Anticipation and Preparedness Initiative (ZAPI project; IMI Grant Agreement n° 115760 granted to Christian Drosten).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. WHO: *The top 10 causes of death*. URL: http://www.who.int/mediacentre/factsheets/fs310/en/

2. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P: **Global trends in emerging infectious diseases**. *Nature* 2008, **451**:990-993.

3. Knauf S, Liu H, Harper KN: **Treponemal infection in nonhuman primates as possible reservoir for human yaws**. *Emerg Infect Dis* 2013, **19**:2058-2060.

4. LaFond RE, Molini BJ, Van Voorhis WC, Lukehart SA: **Antigenic variation of TprK V regions abrogates specific antibody binding in syphilis**. *Infect Immun* 2006, **74**:6244-6251.

5. Alderete JF, Baseman JB: **Surface-associated host proteins on virulent *Treponema pallidum***. *Infect Immun* 1979, **26**:1048-1056.

6. Cox DL, Chang P, McDowall AW, Radolf JD: **The outer membrane, not a coat of host proteins, limits antigenicity of virulent *Treponema pallidum***. *Infect Immun* 1992, **60**:1076-1083.

7. Baseman JB, Hayes EC: **Molecular characterization of receptor binding proteins and immunogens of virulent *Treponema pallidum***. *J Exp Med* 1980, **151**:573-586.

8. Riviere GR, Thomas DD, Cobb CM: **In vitro model of *Treponema pallidum* invasiveness**. *Infect Immun* 1989, **57**:2267-2271.

9. Shaw GM, Hunter E: **HIV transmission**. *Cold Spring Harb Perspect Med* 2012:2.

10. Williams KC, Burdo TH: **HIV and SIV infection: the role of cellular restriction and immune responses in viral replication and pathogenesis**. *APMIS* 2009, **117**:400-412.

11. • Ronen K, Sharma A, Overbaugh J: **HIV transmission biology: translation for HIV prevention**. *AIDS* 2015, **29**:2219-2227.
In this very complete review, several aspects of the biology of HIV transmission and their implications for existing prevention and control measures are reviewed.

12. Sharp PM, Hahn BH: **Origins of HIV and the AIDS pandemic**. *Cold Spring Harb Perspect Med* 2011, **1** a006841.

13. Neil SJ, Zang T, Bieniasz PD: **Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu**. *Nature* 2008, **451**:425-430.

14. Sheehy AM, Gaddis NC, Choi JD, Malim MH: **Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein**. *Nature* 2002, **418**:646-650.

15. Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J: **The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys**. *Nature* 2004, **427**:848-853.

16. Sauter D, Schindler M, Specht A, Landford WN, Munch J, Kim KA, Votteler J, Schubert U, Bibollet-Ruche F, Keele BF *et al.*: **Tetherin-driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV-1 strains**. *Cell Host Microbe* 2009, **6**:409-421.

17. Bush S, Tebit DM: **HIV-1 Group O origin, evolution, pathogenesis, and treatment: unraveling the complexity of an outlier 25 years later**. *AIDS Rev* 2015, **17**:147-158.

18. de Silva TI, Cotten M, Rowland-Jones SL: **HIV-2: the forgotten AIDS virus**. *Trends Microbiol* 2008, **16**:588-595.

19. Sauter D, Hue S, Petit SJ, Plantier JC, Towers GJ, Kirchhoff F, Gupta RK: **HIV-1 Group P is unable to antagonize human tetherin by Vpu, Env or Nef**. *Retrovirology* 2011, **8**:103.

20. Boutwell CL, Rolland MM, Herbeck JT, Mullins JI, Allen TM: **Viral evolution and escape during acute HIV-1 infection**. *J Infect Dis* 2010, **202(Suppl. 2)**:S309-S314.

21. Morse SS: **Factors in the emergence of infectious diseases**. *Emerg Infect Dis* 1995, **1**:7-15.

22. Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA *et al.*: **Identification of a novel coronavirus in patients with severe acute respiratory syndrome**. *N Engl J Med* 2003, **348**:1967-1976.

23. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA: **Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia**. *N Engl J Med* 2012, **367**:1814-1820.

24. Lau SK, Woo PC, Yip CC, Tse H, Tsoi HW, Cheng VC, Lee P, Tang BS, Cheung CH, Lee RA *et al.*: **Coronavirus HKU1 and other coronavirus infections in Hong Kong**. *J Clin Microbiol* 2006, **44**:2063-2071.

25. • Drexler JF, Corman VM, Drosten C: **Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS**. *Antiviral Res* 2014, **101**:45-56.
Drexler *et al*. provide phylogenetic evidence for a zoonotic origin of four out of six human coronaviruses: HCoV-229E, SARS-CoV and MERS-CoV originate from bats, and HCoV-OC43 from bovines.

26. Perlman S, Netland J: **Coronaviruses post-SARS: update on replication and pathogenesis**. *Nat Rev Microbiol* 2009, **7**:439-450.

27. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C *et al.*: **Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor**. *Nature* 2013, **503**:535-538.

28. •• Haagmans BL, Al Dhahiry SH, Reusken CB, Raj VS, Galiano M, Myers R, Godeke GJ, Jonges M, Farag E, Diab A *et al.*: **Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation**. *Lancet Infect Dis* 2014, **14**:140-145.
This study provides virological and serological evidence of MERS-CoV in camels. The sequences of the viruses obtained from nasal swabs of camels were similar to that of two human isolates from the same farm, suggesting the possibility of a recent outbreak affecting both humans and camels.

29. WHO: *First data on stability and resistance of SARS coronavirus compiled by members of WHO laboratory network*. URL: http://www.who.int/csr/sars/survival_2003_05_04/en/

30. WHO: *SARS (Severe Acute Respiratory Syndrome)*. URL: http://www.who.int/ith/diseases/sars/en/

31. WHO: *Middle East respiratory syndrome coronavirus (MERS-CoV)*. URL: http://www.who.int/mediacentre/factsheets/mers-cov/en/

32. Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H: **Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis**. *J Pathol* 2004, **203**:631-637.

33. •• Meyerholz DK, Lambertz AM, McCray PB Jr: **Dipeptidyl Peptidase 4 distribution in the human respiratory tract: implications for the Middle East Respiratory Syndrome**. *Am J Pathol* 2016, **186**:78-86.

In this study, the authors studied the localization of the receptor of the MERS-CoV, DPP4, in the human respiratory tract. DPP4 was found preferentially in alveolar regions, in several different cells types. Moreover, authors found that patients with respiratory co-morbidities exhibited an increased expression of DPP4, explaining why pre-existing pulmonary diseases could predispose people to MERS-CoV morbidity and mortality.

34. Ng DL, Al Hosani F, Keating MK, Gerber SI, Jones TL, Metcalfe MG, Tong S, Tao Y, Alami NN, Haynes LM *et al.*: **Clinicopathologic, immunohistochemical, and ultrastructural findings of a fatal case of Middle East respiratory syndrome coronavirus infection in the United Arab Emirates, April 2014**. *Am J Pathol* 2016, **186**:652-658.

35. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC *et al.*: **Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus**. *Nature* 2003, **426**:450-454.

36. Raj VS, Mou H, Smits SL, Dekkers DH, Muller MA, Dijkman R, Muth D, Demmers JA, Zaki A, Fouchier RA *et al.*: **Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC**. *Nature* 2013, **495**:251-254.

37. Bertram S, Glowacka I, Muller MA, Lavender H, Gnirss K, Nehlmeier I, Niemeyer D, He Y, Simmons G, Drosten C *et al.*: **Cleavage and activation of the severe acute respiratory syndrome coronavirus spike protein by human airway trypsin-like protease**. *J Virol* 2011, **85**:13363-13372.

38. Kindler E, Jonsdottir HR, Muth D, Hamming OJ, Hartmann R, Rodriguez R, Geffers R, Fouchier RA, Drosten C, Muller MA *et al.*: **Efficient replication of the novel human betacoronavirus EMC on primary human epithelium highlights its zoonotic potential**. *MBio* 2013, **4**:e00611-e00612.

39. Korea Centers for Disease C: **Prevention Middle East respiratory syndrome coronavirus outbreak in the Republic of Korea, 2015**. *Osong Public Health Res Perspect* 2015, **6**:269-278.

40. Corman VM, Albarrak AM, Omrani AS, Albarrak MM, Farah ME, Almasri M, Muth D, Sieberg A, Meyer B, Assiri AM *et al.*: **Viral shedding and antibody response in 37 patients with Middle East respiratory syndrome coronavirus infection**. *Clin Infect Dis* 2016, **62**:477-483.

41. van Doremalen N, Bushmaker T, Munster VJ: **Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions**. *Euro Surveill* 2013:18.

42. Lee SS, Wong NS: **Probable transmission chains of Middle East respiratory syndrome coronavirus and the multiple generations of secondary infection in South Korea**. *Int J Infect Dis* 2015, **38**:65-67.

43. Callow KA, Parry HF, Sergeant M, Tyrrell DA: **The time course of the immune response to experimental coronavirus infection of man**. *Epidemiol Infect* 1990, **105**:435-446.

44. van der Velden VH, Wierenga-Wolf AF, Adriaansen-Soeting PW, Overbeek SE, Moller GM, Hoogsteden HC, Versnel MA: **Expression of aminopeptidase N and dipeptidyl peptidase IV in the healthy and asthmatic bronchus**. *Clin Exp Allergy* 1998, **28**:110-120.

45. Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pohlmann S: **Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry**. *Proc Natl Acad Sci U S A* 2005, **102**:7988-7993.

46. Milewska A, Zarebski M, Nowak P, Stozek K, Potempa J, Pyrc K: **Human coronavirus NL63 utilizes heparan sulfate proteoglycans for attachment to target cells**. *J Virol* 2014, **88**:13221-13230.

47. Richard M, Fouchier RA: **Influenza A virus transmission via respiratory aerosols or droplets as it relates to pandemic potential**. *FEMS Microbiol Rev* 2016, **40**:68-85.

48. Kuhn JH, Becker S, Ebihara H, Geisbert TW, Johnson KM, Kawaoka Y, Lipkin WI, Negredo AI, Netesov SV, Nichol ST *et al.*: **Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations**. *Arch Virol* 2010, **155**:2083-2103.

49. Feldmann H, Geisbert TW: **Ebola haemorrhagic fever**. *Lancet* 2011, **377**:849-862.

50. Dowell SF, Mukunu R, Ksiazek TG, Khan AS, Rollin PE, Peters CJ: **Transmission of Ebola hemorrhagic fever: a study of risk factors in family members, Kikwit, Democratic Republic of the Congo, 1995. Commission de Lutte contre les Epidemies a Kikwit**. *J Infect Dis* 1999, **179(Suppl. 1)**:S87-S91.

51. Judson S, Prescott J, Munster V: **Understanding ebola virus transmission**. *Viruses* 2015, **7**:511-521.

52. Bausch DG, Towner JS, Dowell SF, Kaducu F, Lukwiya M, Sanchez A, Nichol ST, Ksiazek TG, Rollin PE: **Assessment of the risk of Ebola virus transmission from bodily fluids and fomites**. *J Infect Dis* 2007, **196(Suppl. 2)**:S142-S147.

53. Osterholm MT, Moore KA, Kelley NS, Brosseau LM, Wong G, Murphy FA, Peters CJ, LeDuc JW, Russell PK, Van Herp M *et al.*: **Transmission of Ebola viruses: what we know and what we do not know**. *MBio* 2015, **6**:e00137.

54. • Uyeki TM, Erickson BR, Brown S, McElroy AK, Cannon D, Gibbons A, Sealy T, Kainulainen MH, Schuh AJ, Kraft CS *et al.*: **Ebola virus persistence in semen of male survivors**. *Clin Infect Dis* 2016, **62**:1552-1555.

In this study, the semen of five male Ebola virus disease survivors was tested for virus presence after recovery. Viral RNA could be detected in the semen up to 290 days after the onset of the symptoms and virus culture from semen sample was possible until 72 days after illness onset.

55. Olival KJ, Hayman DT: **Filoviruses in bats: current knowledge and future directions**. *Viruses* 2014, **6**:1759-1788.

56. Reed DS, Lackemeyer MG, Garza NL, Sullivan LJ, Nichols DK: **Aerosol exposure to Zaire ebolavirus in three nonhuman primate species: differences in disease course and clinical pathology**. *Microbes Infect* 2011, **13**:930-936.

57. Curran KG, Gibson JJ, Marke D, Caulker V, Bomeh J, Redd JT, Bunga S, Brunkard J, Kilmarx PH: **Cluster of Ebola Virus Disease Linked to a Single Funeral — Moyamba District, Sierra Leone, 2014**. *Morb Mortal Wkly Rep* 2016, **65**:202-205.

58. •• Holmes EC, Dudas G, Rambaut A, Andersen KG: **The evolution of Ebola virus: insights from the 2013–2016 epidemic**. *Nature* 2016, **538**:193-200.

This review addresses very well several aspects of Ebola virus biology during the last 2013–2016 outbreak, such as the controversy about evolution rates, the lack of changing in route of transmission or the sexual transmission of the virus.

59. •• Carroll MW, Matthews DA, Hiscox JA, Elmore MJ, Pollakis G, Rambaut A, Hewson R, Garcia-Dorival I, Bore JA, Koundouno R *et al.*: **Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa**. *Nature* 2015, **524**:97-101.

Carroll *et al*. retraced the genetic evolution of the Ebola virus during the last 2014 outbreak in West Africa by sequencing 179 full genomes of viruses isolated from patients using next generation sequencing. They traced back the origin and time of transmission of the virus in the different affected countries and found that the outbreak has resulted in multiple lineages.

60. WHO: *Factors that contributed to undetected spread of the Ebola virus and impeded rapid containment*. URL: http://www.who.int/csr/disease/ebola/one-year-report/factors/en/.

3. The Concept of One Health

The idea of One Health (Figure 1) was conceptualized in 2014 (Destoumieux-Garzón et al., 2018). The model echoes our understanding of disease ecology and is widely accepted in public health (Kahn et al., 2008). Of fundamental importance to the concept is that the majority of diseases can infect both, humans and animals alike. The health of humans is therefore linked to the health of animals and the environment.

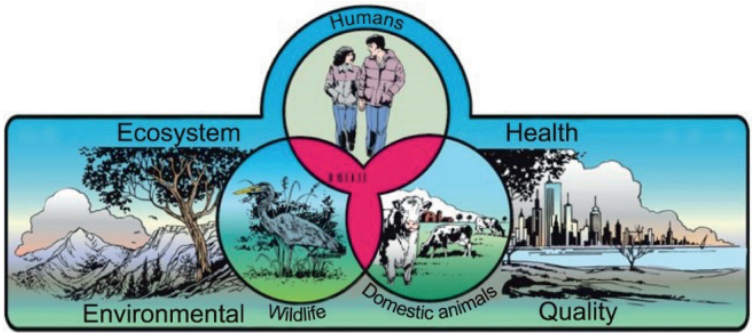


Figure 1. The figure illustrates the concept of One Health, where human health is interconnected with animal and environmental health. Pathogens and their hosts are integral part of natural ecosystems. Figure designed by John M. Evans, USGS. ©2017 The Korean Society of Veterinary Science (Sleeman et al., 2017).

Most pathogens are generalists and infect a wider range of host species and only a few have evolved high host-specificity (Taylor et al., 2001; Woolhouse and Gowtage-Sequeria, 2005). It is therefore not surprising that an estimated 60% of human diseases are of zoonotic origin (Taylor et al., 2001) and that even some of the exclusive human pathogens have an equally zoonotic origin (Wolfe et al., 2007). A paramount example for the latter is HIV that derived from at least two NHP-to-human spill-overs of the simian immunodeficiency virus (SIV) (Hirsch et al., 1995). Per definition, a zoonosis requires at least three species: the pathogen, a human being and an animal (Karesh et al., 2012). The World Health Organization defines a zoonosis as a "disease and infection which is *naturally* transmitted between vertebrate animals and man" (Shakespeare, 2009). The latter definition therefore not only provides a directionality, but it also includes the important aspect that disease transmission must happen under natural conditions. This implies that animal infection experiments alone are not sufficient to demonstrate the functionality of a disease reservoir in the wild (Hallmaier-Wacker et al., 2017). Knowledge about reservoir host systems is central to our understanding of infectious diseases (Delahay et al., 2009). Despite numerous efforts, until today, smallpox which had no nonhuman reservoir is the single disease that has been eradicated in humans (Fenner et al., 1988).

4. The Emergence of Diseases and Reservoir Systems

The emergence of infectious diseases is generally the result of rapid change in the ecology of the host, the pathogen or a combination of both (Daszak et al., 2000). The global increase in anthropogenic factors such as land-use, resource exploitation, urbanization, the introduction of new species and climate change impact ecosystems that have evolved over millions of years. The continuous growth of the human population in our modern world has led to an unprecedented growth of the human-wildlife(-livestock) interface (Anand and Radhakrishna, 2017; Hassell et al., 2017; Nyhus, 2015), which provides the basis for new and re-emerging diseases (Hassell et al., 2017).

Historically host-pathogen studies have addressed a single host-single pathogen system, that in the majority of cases is at best overly simplistic and at worst wrong. Although the conceptual framework for the definition of a disease reservoir has been defined multiple times (Ashford, 2003; Haydon et al., 2002; Viana et al., 2014), empirical characterization of zoonotic reservoir systems remains challenging (Hallmaier-Wacker et al., 2017). There are two essential requirements that a (wildlife) reservoir system must fulfil until it can be accepted as a functional disease reservoir (Hallmaier-Wacker et al., 2017). (A) The pathogen must be maintained in the reservoir system and (B) there must be a feasible transmission route that allows the pathogen to be transmitted from the reservoir system to the target group, which is an explicitly defined population of interest (e.g., humans). Both requirements, the pathogen maintenance in the reservoir and the feasible transmission route, must be demonstrated multiple times over a reasonable time to demonstrate the stability of the reservoir system. Furthermore, the respective pathogen that is maintained in the reservoir system must have a high genetic and functional similarity to the pathogen that is found to infect the target group. The functional similarity (the pathogen in the reservoir system is functionally able to infect the target host) is of particular importance since a single point mutation in an otherwise identical organism could potentially result in changes of virulence in the target group host. This is for example discussed in the context of the Ebola virus (Bedford and Malik, 2016).

Communicable diseases require feasible transmission routes and therefore spatial connectivity between the reservoir system and the target group. This becomes clear in vector-borne diseases. Even if the identical pathogen has been demonstrated in the nonhuman reservoir and the target group, the potential reservoir system would be functionally irrelevant in the absence of the vector. Moreover, proof of pathogen involvement in a transmission pathway needs to be demonstrated as well as proof of the viability of the pathogen during transmission. The capability of survival in the external environment differs from pathogen to pathogen (Walther and Ewald, 2004) and some microbes cannot maintain viability outside their host e.g., the syphilis and yaws bacterium *Treponema pallidum* (Willcox and Guthe, 1966b). As a consequence, nucleic acid amplification tests (NAAT) that are commonly applied to detect the pathogen on the RNA or DNA level, are not predictive for its viability and thus infectiousness for a new host (Pillai, 1997).

The aforementioned arguments demonstrate, that infectious diseases must be evaluated within their natural environment. Neither laboratory experiments nor isolated cases alone are informative to describe the functionality of a disease reservoir. A sophisticated understanding of reservoir systems is key for the monitoring of potential and known disease reservoirs and to detect disease emergencies in early stages (Karesh et al., 2012).

4.1 Disease Reservoirs: From Conceptual Frameworks to Applicable Criteria

Central to any disease eradication program is the question of whether or not a pathogen has a nonhuman reservoir (Hallmaier-Wacker et al., 2017). There is a tendency in published scientific data to exaggerate positive scientific results and to hype certain areas of science (Caulfield et al., 2016; Vinkers et al., 2015). This has resulted in an inflationary use of the term 'reservoir of infection' across the different research fields. The one-time discovery of a zoonotic pathogen in a wildlife species is not enough evidence to claim a disease reservoir as both, the maintenance of the pathogen in the reservoir system and the feasible transmission route must be demonstrated. The following publication aimed to discuss and define the requirements for a more standardize acceptance of a disease reservoir across the different disciplines in One Health. It provides a simple and functional oriented framework that can be applied at any population-level including multi-species systems and meta-populations.

OPEN

Emerging Microbes & Infections (2017) 6, e79; doi:10.1038/emi.2017.65
www.nature.com/emi

REVIEW

Disease reservoirs: from conceptual frameworks to applicable criteria

Luisa K Hallmaier-Wacker[1,2], Vincent J Munster[3] and Sascha Knauf[1]

Central to the One Health approach and any disease eradication program is the question of whether a pathogen has a non-human reservoir. Despite well-established conceptual frameworks that define a reservoir of infection, empirical characterization of reservoirs often remains controversial, challenging and sometimes misleading. What is essentially missing are applicable requirements that standardize the use of the term 'reservoir of infection' across multiple disciplines. We propose an empirical framework, considering maintenance and feasible transmission of a pathogen, to standardize the acceptance of a disease reservoir across multiple disciplines. We demonstrate the intended use of these requirements by applying them to different diseases that are known to infect both humans and animals.
Emerging Microbes & Infections (2017) **6,** e79; doi:10.1038/emi.2017.65; published online 6 September 2017

Keywords: disease eradication; infection; infectious diseases; interdisciplinary; one health; multidisciplinary

A RESERVOIR NEEDS TO MAINTAIN THE PATHOGEN AND HAVE A FEASIBLE TRANSMISSION ROUTE

The high prevalence of infectious agents of zoonotic and anthropozoonotic origin pose a major health threat to both human and animal populations. A conceptual framework for understanding a reservoir of infection has been established through various studies that have emphasized different aspects of zoonotic diseases.[1–4] However, empirical characterization of reservoirs often remains controversial and challenging. The most applicable and accepted way to investigate and define a reservoir emphasizes the annotation of a target group (Figure 1), which is an explicitly defined population of interest in a dynamic and heterogeneous landscape (for example, humans at the livestock–wildlife–human interface).[4,5] According to Haydon *et al.*,[4] the target group is a matter of definition and may therefore be disconnected from the ecological reality. The target group provides a directionality to the study of a reservoir system. All other susceptible populations (non-target populations), which directly or indirectly connect epidemiologically to the target (Figures 1 and 2), can be part of the potential reservoir.[4] For a non-target population to be considered an accepted functional reservoir, maintenance of a single pathogen in the population needs to be shown in combination with a feasible transmission route between the target and non-target populations.[4]

Although the conceptual framework of a disease reservoir is already well-defined, applicable requirements for an evidence-based rejection or acceptance of a reservoir are currently missing. In particular, interdisciplinary standards on genetic and functional similarities of reservoir and human isolates of pathogens are nonexistent. Considering the increase in interdisciplinary research, we see the need to critically discuss and standardize the use of the term 'reservoir of infection' across different research fields to oppose the tendency of published scientific data to exaggerate positive results and hype certain areas of science.[6,7] Although we do not claim absolute standardization of empirical requirements to accept a reservoir across disciplines, we present a framework to serve as a basis for a pending discussion in the growing One Health community. The simplicity and functional orientation of the presented framework allows for straightforward application but does not negate more complex populations, as the same principles can be applied to multi-species systems and metapopulations (Figure 2).

According to the accepted definition of a reservoir proposed by Haydon *et al.*,[4] we discuss the requirements in two parts: the pathogen's maintenance in a potential population or community followed by a discussion on proof of a feasible transmission route. Although the two components are addressed separately, only together they demonstrate the existence of a functional reservoir.

PROOF OF PATHOGEN MAINTENANCE IN A POTENTIAL RESERVOIR

Increases in technological advancements (for example, next-generation sequencing) and vast quantities of available data have not led to concrete applicable criteria when examining the capacity of a pathogen to be maintained in a population. Recognizing both the ethical limitations in regards to animal testing[8] and the advances in the molecular detection of pathogens, we propose the following criteria to demonstrate the maintenance of a pathogen in a population: (i) a high-genetic similarity of the pathogen found in the reservoir system, (ii) a high degree of functional similarity (infectivity and viability), and

[1]Work Group Neglected Tropical Diseases, Pathology Unit, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, Göttingen 37077, Germany; [2]Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, Göttingen 37077, Germany and [3]Virus Ecology Unit, Laboratory of Virology, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rocky Mountain Laboratories, Hamilton, MT 59840, USA
Correspondence: S Knauf
E-mail: sknauf@dpz.eu
Received 21 March 2017; revised 20 June 2017; accepted 28 June 2017

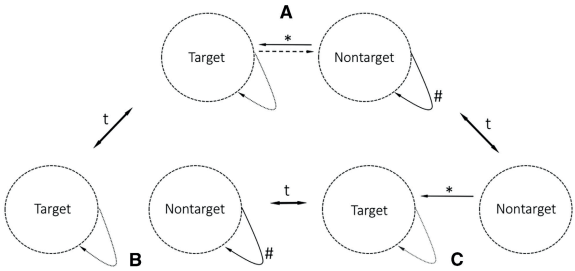


Figure 1 Three scenarios describing the dynamics of a simple reservoir system. (**A**) Pathogen maintenance in the non-target population and feasible transmission route towards the target population. Only this constellation fulfills the requirements of a functional reservoir system. (**B**) Pathogen maintenance in the non-target but no feasible transmission route towards the target population. This is a likely situation whether contact rates between the non-target and target populations are below the threshold. (**C**) No pathogen maintenance in the non-target, but a feasible transmission route exists. An example of the effect of a vaccination strategy in the non-target population. The dynamic of the system is indicated by arrows associated with a 't' (time factor). #Maintenance, *feasible transmission, solid arrows = obligatory, broken line = optional.

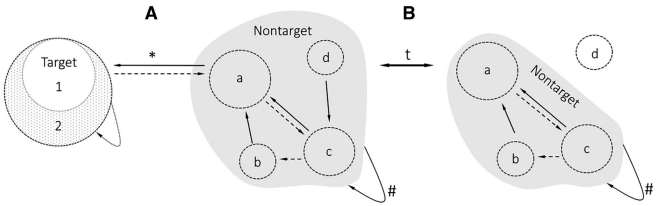


Figure 2 The simplicity and functional orientation of the presented framework allows for straightforward application but does not negate more complex populations. The same principles apply to multi-species systems and metapopulations. The defined target group may be adjusted based on interest and may therefore include metapopulations (targets 1 and 2). The non-target group increases in complexity due to the inclusion of multiple populations (a–d). (**A**) Similarly, to a simple reservoir system, all susceptible populations that connect to the target either (a) directly or (b–d) indirectly are part of the non-target population. (**B**) Temporal shifts in the ecological landscape of the non-target population may lead to the (d) exclusion of populations either due to lack of connectivity or susceptibility. The dynamic of the system is indicated by arrows associated with a 't' (time factor). #Maintenance, *feasible transmission, solid arrows = obligatory, broken line = optional.

(iii) a longitudinal approach that considers the factor of time (Table 1). Owing to the functional orientation of the requirements and for simplicity, all entities involved in the biological lifecycles of a parasite (for example, primary and intermediate hosts) should be considered a single functional unit. Appropriate sequence and functional analysis of a pathogen isolated multiple times from a potential reservoir should be required to prove that a pathogen is maintained in a population. The ability to quickly and cheaply sequence whole genomes has allowed for better genetic resolution.[49,50] Sequence data can be used to examine similarity in the pathogen between a potential reservoir and a target. However, mutation rates vary significantly between pathogens[51,52] and the threshold for sequence and functional similarity must be individually defined and accepted by the scientific community. A single-nucleotide difference can potentially result in a loss of infectivity, for example, when important invasion mechanisms are affected (receptor affinity). In bacteria, investigations can be further complicated by plasmids that can be exchanged and mutated over time.[53] A high amount of phylogenetic relatedness of pathogens isolated from the non-target and target populations does not provide sufficient evidence for the involvement of a pathogen and its ability to infect both groups. Importantly, DNA-based analyses only provide information on the functional potential of a pathogen and must not reflect the gene-expression within a host.[54] For example, the bacterium *Treponema paraluiscuniculi* (which causes syphilis in rabbits), is over 99% identical on the basis of the whole genome to the human pathogen *T. pallidum* (which causes human treponematosis), but does not infect humans.[55] As phylogenetic information fails to reflect the downstream effects of mutations, proof that a pathogen can proliferate in the potential reservoir is required.[56] Information on the transcriptome and proteome of bacteria or the phenotype of viruses are necessary to see the effect of mutations on pathogen viability.[57] There are different ways to test for the functional ability of a pathogen in different species. Owing to the ethical concerns, cell and tissue assays have been increasingly used in therapeutic research instead of animal models.[8] Although these assays are limited in their conclusiveness, they can provide important insight into the molecular mechanisms involved. For example, the failure to infect primary tissue culture from rhesus macaques with human immunodeficiency virus 1 (HIV-1) demonstrates that non-human primates were unlikely to act as a maintenance population (Table 1).[58] In some instances, for example, with uncultivable bacteria such as *Treponema pallidum*, it may be necessary to use animal models to examine the functionality of a pathogen within a potential reservoir species. Knowledge of the biology of the pathogen is essential to properly define a sequence and functional similarity threshold for a particular reservoir system.

Table 1 Applicable requirements that need to be fulfilled for the acceptance of a disease reservoir and their exemplary use in selected diseases that are known to infect humans and animals

				Maintenance in NT		Feasible transmission route			Time factor	
Pathogen	Target	Non-target	Main transmission route	High-genetic similarity	Functional similarity	Spatial and temporal connectivity	Pathogen involvement	Maintaining pathogen viability	Example of longitudinal study	Refs.
Influenza A virus (H1N1)	Human	Swine	Aerosol	X	X	X	X	X	9	10–12
MERS-Coronavirus	Human	Camel	Direct contact	X	X	X	X	(X)	13	14–16
Brucella melitensis (localized brucellosis)	Human	Sheep	Food-borne	X	X	X	X	X	17	18,19
Immunodeficiency virus	Human	NHP	Direct contact	(X)	NP	X	NP	NP	N/A	20
Treponema pallidum pertenue (yaws)	Human	NHP	Direct contact/vector	X	(X)	X	(X)	NP	N/A	21–23
Mycobacterium bovis (bovine tuberculosis)	Human	Cattle	Food-borne/aerosol	X	X	X	X	X	24	25,26
Rabies virus	Human	Fox	Bite	X	X	X	X	X	27	28–30
Echinococcus multilocularis (alveolar echinococcosis)	Human	Fox	Oral/fecal	X	X	X	X	X	31	32,33
Hantavirus	Human	Rodent	Aerosol	X	X	X	X	X	34	35
Ebola virus	Human	Bats	Contact/aerosol	X	NP	X	(X)	NP	N/A	36–38
Zika virus	Human	NHP	Vector	(X)	NP	X	X	NP	39	40–42
Borrelia burgdorferi (borreliosis)	Human	Wildlife	Vector	X	X	X	X	X	43	44–46
Yellow fever virus	Human	NHP	Vector	X	X	X	X	X	47	48

Abbreviations: not available, N/A; non-human primate, NHP; non-target, NT; not provided/no current evidence, NP; evidence, X; partial evidence, (X).
Classical reservoir systems fulfill all requirements proposed in this study.

When examining pathogen maintenance, a longitudinal approach is required to consider the dynamics of a potential reservoir system, including the influence of genetic variation in any given population. Defining a population that was infected at a single time point as a maintenance population for a pathogen is based on assumptions and is therefore speculative. Sero-prevalence surveys are an attractive way to detect the presence of pathogens in a population, as it indicates that an immunocompetent subject was in contact with the pathogen.[1] However, only longitudinal studies with adequate sampling regimes (multiple sampling) to test for antibodies against a pathogen can provide information on the timing or frequency of infection, both of which are important for reservoir studies.[59] Furthermore, cross-reactivity and erroneous assays can lead to false-positive results. For more diffuse reservoir systems, including multi-species compositions where the diversity of host susceptibility (at the individual, species or population level) protects against widespread infection (dilution effect),[60] a longer time frame must be applied. This guarantees a more accurate understanding of the maintenance within a population (for example, Ebola[36]).

PROOF OF FEASIBLE TRANSMISSION ROUTE

Maintenance of a pathogen in a population alone does not provide sufficient proof that a functional reservoir exists. A connection between the target and the non-target populations must be established;[4] otherwise the non-target population remains a maintenance population with the potential to be a reservoir. Therefore, the determination of a feasible and somewhat permanent transmission route between the non-target and target populations is key to identifying a reservoir system (Figure 1). For multi-species reservoir systems, the transmission route between the target and non-target populations may be indirect (Figure 2, connection between b and target), possibly incorporating different hierarchical levels of a non-target community.[4,61] The type of transmission route dictates the form of evidence needed to prove that a feasible transmission route exists between the reservoir and target. For simplicity, we define vectors as part of the transmission route, although under certain circumstances (for example, permanency or substantial amplification in the vector), they may act as part of the non-target community.[61] Four basic requirements need to be met to make a compelling argument for the existence of a feasible transmission route: (i) spatial (direct or indirect) and temporal connectivity between the reservoir system and the target population, (ii) pathogen involvement in this feasible transmission route, (iii) proof of viability of the pathogen during the proposed transmission route and (iv) a longitudinal approach that requires the isolation of a pathogen multiple times in a given transmission route (Table 1).

To prove the feasibility of a transmission route, direct or indirect spatial connectivity as well as temporal overlap between the non-target and target populations must be present. Connectivity measurements depend on the type of transmission route; for example, direct contact transmission requires overlapping territory. Computational tools can help determine the necessary overlap in a population by modeling the transmission across an affected population.[62] In addition to spatial and temporal overlap, the involvement of the pathogen in the particular transmission route needs to be shown, which again requires long-term field projects. In the case of Lyme disease caused by *Borrelia burgdorferi*, nucleic acids from the bacterium were detected in ticks using PCR.[44] However, the detection of DNA does not directly prove that transmission occurs. To gain further confidence that the transmission is feasible, it is therefore essential to show that the infectious organism remains viable during the proposed transmission

route.[45] This means that in addition to PCR detection, the viable pathogen needs to be isolated during a transmission event, where the measure of viability depends on the type of pathogen. In airborne transmission, for example, environmental factors such as size of droplets, UV light and humidity can greatly influence the transmissibility of a virus (as reviewed in Tang[63]). If the amount of viable and therefore infectious organisms is below the infectious dose, the particular transmission route is unfeasible. Without a feasible transmission route between target and non-target populations, no functional reservoir exists. Furthermore, to include all parts of a reservoir population, long-term investigations must focus on the transmission between the non-target and target groups as well as feasible transmission within the non-target community.[61] Unconnected maintenance host populations may become a future reservoir through temporal shifts of the ecosystem.

CHALLENGES OF IMPLEMENTATION

Biological systems are dynamic and can change over time (Figure 1). Single transmission events do not confirm a reservoir of infection (for example, HIV,[20] Table 1). It is therefore important to show continuity and persistence in both maintenance and transmission, which can only be achieved through multiple and adequately timed (field) investigations. Well-designed intervention studies can be used as quasi-experiments to study a reservoir of infection but should not be used as a stand-alone test for the existence of a reservoir.[1] Despite sufficient planning, the cause and effect of intervention studies are often difficult to determine[1,64] and the removal of a pathogen from a particular ecosystem may cause unanticipated effects. A negative outcome does not necessary indicate the lack of a reservoir or transmission route.[64,65] Instead, it can show that the intervention may have been incomplete or that the complexity of a reservoir is not entirely understood.

Pathogens must be studied in the context of natural ecosystems. The complexity of reservoir systems increases as multiple non-target populations interact as an ecological entity, which is influenced by factors such as competition, co-existence or predation.[66] Furthermore, the artificial environment in a laboratory, which is often used to study the susceptibility of a species, differs substantially from a natural setting.[67] The use of laboratory animals or cell- and tissue-based assays can be advantageous when studying pathogenicity, but it cannot solely contribute to the understanding of the epidemiology of a pathogen, which is largely impacted by variables such as genetic diversity, co-infection, cross-protective immunity and spatial connectivity. As a consequence, any epidemiological model requires additional information on the geographic range and the ecological landscape.[68] This includes population densities and functional profiles of species that are involved in the reservoir system.[60,69] The importance of sample size in field studies and animal experiments cannot be stressed enough as it greatly affects the efficacy of analysis, especially in reservoirs with low-frequency crossover events.

Neither laboratory experiments, nor intervention studies, nor epidemiological models alone can provide a full understanding of a natural reservoir of infection. Only the combination of methods that are based on established and validated species-specific assays and technically sound field investigations can provide confidence that the pathogen is maintained in a non-target population and that a feasible transmission route exists. This, however, requires the political will and financial support to conduct long-term One Health studies to explore diseases in their natural context.

CONCLUSION

The term 'disease reservoir' should be used carefully and only if there is convincing evidence demonstrating the maintenance and a feasible transmission route of a particular pathogen (Figure 1). We propose overarching requirements that must be fulfilled to provide ample proof that a reservoir exists (Table 1). Classical reservoir systems (for example, Lyme disease caused by *Borrelia burgdorferi*) fulfill all of the requirements proposed in this study, whereas some well-known diseases, such as Ebola, need further research until a reservoir system can be accepted (Table 1). For the pathogens without an accepted reservoir, the framework introduced in this study also indicates the outstanding questions that future research should focus on to investigate the presence of a reservoir system. A broader expert-based multidisciplinary discussion is needed to develop standards for the diversity of pathogens.

ACKNOWLEDGEMENTS

VJM is supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH). SK received funding from the German Research Foundation (DFG): KN1097/3-1 and KN1097/4-1. We would like to thank the reviewer for addressing important aspects that clearly improved the proposed framework.

1 Viana M, Mancy R, Biek R *et al*. Assembling evidence for identifying reservoirs of infection. *Trends Ecol Evol* 2014; **29**: 270–279.
2 Drexler JF, Corman VM, Müller MA *et al*. Bats host major mammalian paramyxoviruses. *Nat Commun* 2012; **3**: 796.
3 Ashford R. When is a reservoir not a reservoir?. *Emerg Infect Dis* 2003; **9**: 1495–1496.
4 Haydon DT, Cleaveland S, Taylor LH *et al*. Identifying reservoirs of infection: a conceptual and practical challenge. *Emerg Infect Dis* 2002; **8**: 1468–1473.
5 Dunning JB, Stewart DJ, Danielson BJ *et al*. Spatially explicit population models: current forms and future uses. *Ecol Appl* 1995; **5**: 3–11.
6 Caulfield T, Sipp D, Murry CE *et al*. Confronting stem cell hype. *Science* 2016; **352**: 776–777.
7 Vinkers CH, Tijdink JK, Otte WM. Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. *BMJ* 2015; **351**: h6467–h6777.
8 Ferdowsian HR, Beck N. Ethical and scientific considerations regarding animal testing and research. *PLoS One* 2011; **6**: e24059.
9 Simon G, Larsen LE, Dürrwald R *et al*. European surveillance network for influenza in pigs: surveillance programs, diagnostic tools and Swine influenza virus subtypes identified in 14 European countries from 2010 to 2013. *PLoS One* 2014; **9**: e115815.
10 Brookes SM, Núñez A, Choudhury B *et al*. Replication, pathogenesis and transmission of pandemic (H1N1) 2009 virus in non-immune pigs. *PLoS One* 2010; **5**: e9068.
11 Steel J, Staeheli P, Mubareka S *et al*. Transmission of pandemic H1N1 influenza virus and impact of prior exposure to seasonal strains or interferon treatment. *J Virol* 2010; **84**: 21–26.
12 Maines TR, Jayaraman A, Belser JA *et al*. Transmission and pathogenesis of swine-origin 2009A (H1N1) influenza viruses in ferrets and mice. *Science* 2009; **325**: 484–487.
13 Müller MA, Meyer B, Corman VM *et al*. Presence of Middle East respiratory syndrome coronavirus antibodies in Saudi Arabia: a nationwide, cross-sectional, serological study. *Lancet Infect Dis* 2015; **15**: 559–564.
14 Mohd HA, Al-Tawfiq JA, Memish ZA. Middle East respiratory syndrome coronavirus (MERS-CoV) origin and animal reservoir. *Virol J* 2016; **13**: 87.
15 Van Doremalen N, Bushmaker T, Munster V. Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. *Euro Surveill* 2013; **18**: 20590.
16 Haagmans BL, van den Brand JM, Raj VS *et al*. An orthopoxvirus-based vaccine reduces virus excretion after MERS-CoV infection in dromedary camels. *Science* 2015; **351**: 77–81.
17 Dean AS, Crump L, Greter H *et al*. Global burden of human brucellosis: a systematic review of disease frequency. *PLoS Negl Trop Dis* 2012; **6**: e1865.
18 de Figueiredo P, Ficht TA, Rice-Ficht A *et al*. Pathogenesis and immunobiology of brucellosis: review of Brucella–Host Interactions. *Am J Pathol* 2015; **185**: 1505–1517.
19 Paixão TA, Roux CM, Den Hartigh AB *et al*. Establishment of systemic *Brucella melitensis* infection through the digestive tract requires urease, the type IV secretion system, and lipopolysaccharide O antigen. *Infect Immun* 2009; **77**: 4197–4208.
20 Gao F, Bailes E, Robertson DL *et al*. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* 1999; **397**: 436–441.

21 Knauf S, Dahlmann F, Batamuzi EK *et al*. Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. *PLoS Negl Trop Dis* 2015; **9**: e0003637.
22 Knauf S, Raphael J, Mitjà O *et al*. Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. *EBioMedicine* 2016; **11**: 85–90.
23 Harper KN, Knauf S. *Treponema pallidum* infection in primates: clinical manifestations, epidemiology, and evolution of a stealthy pathogen. In: Brinkworth JF, Pechenkina K (eds). *Primates, Pathogens, and Evolution*. New York, NY: Springer, 2013: 189–219.
24 Walker TM, Lalor MK, Broda A *et al*. Assessment of *Mycobacterium tuberculosis* transmission in Oxfordshire, UK, 2007–12, with whole pathogen genome sequences: an observational study. *Lancet Respir Med* 2014; **2**: 285–292.
25 Fitzgerald S, Kaneene J. Wildlife reservoirs of bovine tuberculosis worldwide: hosts, pathology, surveillance, and control. *Vet Pathol* 2013; **50**: 488–499.
26 Michel AL, Müller B, van Helden PD. *Mycobacterium bovis* at the animal-human interface: a problem, or not? *Vet Microbiol* 2010; **140**: 371–381.
27 Sidwa TJ, Wilson PJ, Moore GM *et al*. Evaluation of oral rabies vaccination programs for control of rabies epizootics in coyotes and gray foxes: 1995–2003. *J Am Vet Med Assoc* 2005; **227**: 785–792.
28 Rupprecht CE, Hanlon CA, Hemachudha T. Rabies re-examined. *Lancet Infect Dis* 2002; **2**: 327–343.
29 Hemachudha T, Laothamatas J, Rupprecht CE. Human rabies: a disease of complex neuropathogenetic mechanisms and diagnostic challenges. *Lancet Neurol* 2002; **1**: 101–109.
30 Chautan M, Pontier D, Artois M. Role of rabies in recent demographic changes in red fox (*Vulpes vulpes*) populations in Europe. *Mammalia* 2000; **64**: 391–410.
31 Schweiger A, Ammann RW, Candinas D *et al*. Human alveolar echinococcosis after fox population increase, Switzerland. *Emerg Infect Dis* 2007; **13**: 878–882.
32 Veit P, Bilger B, Schad V *et al*. Influence of environmental factors on the infectivity of *Echinococcus multilocularis* eggs. *Parasitology* 1995; **110**: 79–86.
33 Craig P. *Echinococcus multilocularis*. *Curr Opin Infect Dis* 2003; **16**: 437–444.
34 Kuenzi AJ, Douglass RJ, Bond CW *et al*. Long-term dynamics of Sin Nombre viral RNA and antibody in deer mice in Montana. *J Wildl Dis* 2005; **41**: 473–481.
35 Bi Z, Formenty PB, Roth CE. Hantavirus infection: a review and global update. *J Infect Dev Ctries* 2008; **2**: 3–23.
36 Leroy EM, Kumulungui B, Pourrut X *et al*. Fruit bats as reservoirs of Ebola virus. *Nature* 2005; **438**: 575–576.
37 Pourrut X, Kumulungui B, Wittmann T *et al*. The natural history of Ebola virus in Africa. *Microb Infect* 2005; **7**: 1005–1014.
38 Weingartl HM, Embury-Hyatt C, Nfon C *et al*. Transmission of Ebola virus from pigs to non-human primates. *Sci Rep* 2012; **2**: 811.
39 Roth A, Mercier A, Lepers C *et al*. Concurrent outbreaks of dengue, chikungunya and Zika virus infections-an unprecedented epidemic wave of mosquito-borne viruses in the Pacific 2012-2014. *Euro Surveill* 2014; **19**: 20929.
40 Favoretto S, Araujo D, Oliveira D *et al*. First detection of Zika virus in neotropical primates in Brazil: a possible new reservoir. *BioRxiv* 2016, doi: http://www.biorxiv.org/content/early/2016/04/20/049395.
41 Dick G. Zika virus (II). Pathogenicity and physical properties. *Trans R Soc Trop Med Hyg* 1952; **46**: 521–534.
42 Marano G, Pupella S, Vaglio S *et al*. Zika virus and the never-ending story of emerging pathogens and transfusion medicine. *Blood Transfus* 2016; **14**: 95–100.
43 Ogden NH, Bouchard C, Kurtenbach K *et al*. Active and passive surveillance and phylogenetic analysis of *Borrelia burgdorferi* elucidate the process of Lyme disease risk emergence in Canada. *Environ Health Perspect* 2010; **118**: 909–914.
44 Barbour AG, Maupin GO, Teltow GJ *et al*. Identification of an uncultivable *Borrelia* species in the hard tick *Amblyomma americanum*: possible agent of a Lyme disease-like illness. *J Infect Dis* 1996; **173**: 403–409.
45 Moriarty TJ, Norman MU, Colarusso P *et al*. Real-time high resolution 3D imaging of the lyme disease spirochete adhering to and escaping from the vasculature of a living host. *PLoS Pathog* 2008; **4**: e1000090.
46 Tilly K, Rosa PA, Stewart PE. Biology of infection with *Borrelia burgdorferi*. *Infect Dis Clin North Am* 2008; **22**: 217–234.
47 Almeida MA, Cardoso JDC, dos Santos E *et al*. Surveillance for yellow fever virus in non-human primates in Southern Brazil, 2001–2011: a tool for prioritizing human populations for vaccination. *PLoS Negl Trop Dis* 2014; **8**: e2741.
48 Monath TP. Yellow fever: an update. *Lancet Infect Dis* 2001; **1**: 11–20.
49 Roetzer A, Diel R, Kohl TA *et al*. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. *PLoS Med* 2013; **10**: e1001387.
50 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci* 1977; **74**: 5463–5467.
51 Dapp MJ, Heineman RH, Mansky LM. Interrelationship between HIV-1 fitness and mutation rate. *J Mol Biol* 2013; **425**: 41–53.
52 Taylor LH, Latham SM, Mark E. Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci* 2001; **356**: 983–989.
53 Okinaka R, Cloud K, Hampton O *et al*. Sequence, assembly and analysis of pX01 and pX02. *J Appl Microbiol* 1999; **87**: 261–262.
54 Miosge LA, Field MA, Sontani Y *et al*. Comparison of predicted and actual consequences of missense mutations. *Proc Natl Acad Sci* 2015; **112**: 5189–5198.
55 Šmajs D, Zobaníková M, Strouhal M *et al*. Complete genome sequence of *Treponema paraluiscuniculi*, strain Cuniculi A: the loss of infectivity to humans is associated with genome decay. *PLoS One* 2011; **6**: e20415.
56 Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. *Genomics* 2008; **92**: 255–264.
57 Prakash T, Taylor TD. Functional assignment of metagenomic data: challenges and applications. *Brief Bioinform* 2012; **13**: 711–727.
58 Hofmann W, Schubert D, LaBonte J *et al*. Species-specific, postentry barriers to primate immunodeficiency virus infection. *J Virol* 1999; **73**: 10020–10028.
59 Hoye BJ, Munster VJ, Nishiura H *et al*. Reconstructing an annual cycle of interaction: natural infection and antibody dynamics to avian influenza along a migratory flyway. *Oikos* 2011; **120**: 748–755.
60 Roche B, Rohani P, Dobson AP *et al*. The impact of community organization on vector-borne pathogens. *Am Nat* 2012; **181**: 1–11.
61 Zeppelini CG, de Almeida AMP, Cordeiro-Estrela P. Zoonoses as ecological entities: a case review of plague. *PLoS Negl Trop Dis* 2016; **10**: e0004949.
62 Kao RR, Haydon DT, Lycett SJ *et al*. Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends Microbiol* 2014; **22**: 282–291.
63 Tang JW. The effect of environmental parameters on the survival of airborne infectious agents. *J R Soc Interface* 2009; **6**: 737–746.
64 Dietze R, Barros GB, Teixeira L *et al*. Effect of eliminating seropositive canines on the transmission of visceral leishmaniasis in Brazil. *Clin Infect Dis* 1997; **25**: 1240–1242.
65 Ezenwa VO, Jolles AE. Opposite effects of anthelmintic treatment on microbial infection at individual versus population scales. *Science* 2015; **347**: 175–177.
66 Menge BA, Sutherland JP. Species diversity gradients: synthesis of the roles of predation, competition, and temporal heterogeneity. *Am Nat* 1976; **110**: 351–369.
67 McCafferty J, Mühlbauer M, Gharaibeh RZ *et al*. Stochastic changes over time and not founder effects drive cage effects in microbial community assembly in a mouse model. *ISME J* 2013; **7**: 2116–2125.
68 Lindenfors P, Nunn CL, Jones KE *et al*. Parasite species richness in carnivores: effects of host body mass, latitude, geographical range and population density. *Glob Ecol Biogeogr* 2007; **16**: 496–509.
69 Giardina AR, Schmidt KA, Schauber EM *et al*. Modeling the role of songbirds and rodents in the ecology of Lyme disease. *Can J Zool* 2000; **78**: 2184–2197.

5. Molecular Methods to Investigate Disease Epidemiology in Wildlife

5.1 Serological Tests

Many pathogens induce an antibody response in their host (adaptive immune system; chapter 2.4). As a consequence, circulating antibodies can be measured from collected blood samples (Murphy et al., 2008b) and in some cases other materials such as faecal material (Neel et al., 2010; Reed et al., 2014) or urine samples (e.g., echinococcosis (Chirag et al., 2015) and *Helicobacter pylori* (Shimoyama et al., 2017)). In contrast to visualization, culture and nucleic acid amplification, the presence of antibodies against a specific pathogen does only permit the conclusion that the host has been exposed to the respective pathogen or an anti-genetically related pathogen (Gardner et al., 1996). Antibodies are characterized through their specificity, amount, isotype or class, and affinity (Murphy et al., 2008b) and these parameters impact serological test performance. The sensitivity and specificity of the serological test are critical parameters that define the confidence in the test results. The number of antibodies can be measured on a continuous scale, although most studies make use of an ordinary scaling (according to the dilution of the tested sera e.g., 1:16). A cut-off value that delineates the positive from the negative test outcome needs to be defined (Gardner et al., 1996). If the cut-off value is set too high, a high number of false-negative reactions must be expected. On the contrary, an excessively low cut-off value results in a high number of false-positive test reactions (Gardner et al., 1996). Although antibodies are highly specific for their corresponding antigen (Murphy et al., 2008b), immunoglobulins produced in response to anti-genetically related pathogens are in many cases cross-reactive. Examples are zika and dengue virus sero cross-reactivity (Dejnirattisai et al., 2016), the cross-reactivity among herpes viruses (Yeo et al., 1981) and the inability to differentiate infection with *T. pallidum* subspecies causing syphilis (ssp. *pallidum*), yaws (ssp. *pertenue*) and bejel (ssp. *endemicum*) based serology (Antal et al., 2002).

Samples taken at different time intervals may show the rise and fall of titres, which provides valuable information on the time of infection as shown for experimentally infected deer mouse (*Peromycus maniculatus*) (Botten et al., 2000). Negative test results must be expected in animals that acquired infection recently. In this case, the host had not enough time to produce antibodies. This can be a limitation for the serological diagnosis of highly virulent pathogens that cause peracute death. On the other hand, the detection of antibodies is associated with a considerable risk of false-positive test results when animals got previously vaccinated (Barasona et al., 2019) or when very young animals are tested that passively acquired maternal antibodies as shown for lyssavirus infection in African bats (Hayman et al., 2018). Despite these limitations, serological tests are (still) an important tool for epidemiological surveys (Metcalf et al., 2016), which is largely based on the fact that methods to detect antibodies (e.g., enzyme-linked immunosorbent assay (ELISA)) are generally robust, associated with low expenditures as well as they can be automated to allow high-throughput testing.

5.2 Nucleic Acid Amplification Tests

Most of the information on living organisms is coded in deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA polymers consist of unique sequences of the four different DNA nucleotides (Adenosine (A), Guanine (G), Cytosine (C) and Thymine(T)) (Watson and Crick, 1953). In 1977, DNA sequencing became available (Maxam and Gilbert, 1977; Sanger et al.,

1977a) and since then has revolutionized infectious disease research. It provided access to information (genetic code) of pathogens that previously were inaccessible. Until today, the determination of nucleotide sequences with high accuracy is of utmost importance. In particular in disease diagnostics and molecular epidemiology, Sanger sequencing remains an important method (Heather and Chain, 2016). Fortunately, automated systems made it possible to move from examining a single gene locus to a genome-wide approach (Hunkapiller et al., 1991; Smith et al., 1986).

5.2.1 Molecular Typing

The ability to detect differences in the genetic code of a pathogen enables us to reliably and reproducibly discriminate different strains within a single species or subspecies (Pérez-Losada et al., 2017). Molecular typing methods, therefore, provide powerful tools to investigate the epidemiological context of infectious diseases, locally and globally (Foley et al., 2009). There are different approaches to strain-type pathogens. Outdated, but still in use is the multilocus-enzyme electrophoresis (Tibayrenc, 2009) where multiple gene loci are amplified and subsequently digested with restriction enzymes. The resulting restriction fragment length polymorphisms were visualized on an agarose gel and the information was then used to distinguish the different types. Some of the current typing systems in pathogens reflect the transition from the isoenzyme era to the use of modern sequence-based methods. While the 1998 established molecular subtyping system for the syphilis bacterium makes use of the amplification of the acid-repeat protein (*arp*) gene that is known to have a varying number of 60-bp repeats in combination with the restriction fragment length polymorphism of *T. pallidum* repeat (*tpr*) genes (Pillay et al., 1998), it was later on modified with an additional gene target sequence at the *tp0548* locus (Marra et al., 2010). Due to the ubiquitous availability of sequencing facilities and the challenges associated with the amplification of repeat patterns and the homologous *tpr*-gene family in *T. pallidum*, current systems make use of multi-locus sequence typing (MLST) (Godornes et al., 2017). Briefly, MLST examines nucleotide variation in multiple sequence parts of the pathogen's genome (Jolley and Maiden, 2014; Maiden, 2006; Maiden et al., 2013; Pavón and Maiden, 2009; Pérez-Losada et al., 2013) and was first introduced in 1998 to identify clones within populations of pathogenic microorganisms (Maiden et al., 1998). Sequence variations within a species or subspecies are assigned as distinct alleles and for each analysed organism, the alleles at any of the target genes are summarized to define the sequence haplotype. At the core of the MLST is the selection of suitable gene targets. Suitable loci need to be a single copy (Pérez-Losada et al., 2017) and have a varying degree of genetic drift (variable region) that is flanked by conserved parts of the genome to allow universal primer binding (Foley et al., 2009). The different candidate genes should be evenly distributed across the genome and should not be part of a homologous gene family. Since recombination is a general problem for typing systems in prokaryotes, target loci ideally focus on housekeeping genes that code for proteins with known function (Glaeser and Kampfer, 2015) and which are, to a lesser degree, affected by diversifying selection (Pérez-Losada et al., 2017). Unfortunately, the in silico identification of MLST candidate loci requires a substantial number of reference genomes. Once suitable target genes are identified, primers need to be designed and tested in a representative number of known strains (Pérez-Losada et al., 2017). To overcome the difficulties that are associated with low copy numbers and

restricted amounts of sample material, which is often the case when working with wildlife samples, nested PCR can be used to amplify the respective target loci.

While MLST can help to identify transmission events at the human-wildlife-livestock interface (Belkum et al., 2007; Janowicz et al., 2018; Kingry et al., 2016), there are considerable problems associated with the change of target genes or the use of multiple systems. Any changes or differences between different MLST systems will annihilate comparability that is required for the long-term study of infectious diseases (Pérez-Losada et al., 2017). MLST systems can be adjusted for resolution and can focus on virulence-associated genes (MVLST) (Chen et al., 2007; Zhang et al., 2004), polymorphic repeat sequences (VNTR) (Djouaka et al., 2018), the variable number of tandem repeats (MLVA) (Lee et al., 2019), ribosomal genes (rMLST) (Jolley et al., 2012) or genes of the core-genome (cgMLST) (Lee et al., 2019).

5.2.2 Metataxonomics

Metataxonomics is a term that describes a high-throughput sequencing (HTS) approach to characterize the microbiota in a given sample (Marchesi and Ravel, 2015). The approach uses marker sequences in genes that are omnipresent in a respective group of microorganisms (e.g., 16S rRNA gene in bacteria and archaea (Edgar, 2018a; Xue et al., 2018) or ITS rRNA gene in fungi (Edgar, 2018a)). Unlike all other microbiota, viruses do not share a single gene (Sullivan, 2015), which limits the use of metataxonomics in this group of pathogens. The term metataxonomics is frequently confused with the term metagenomics, which described a shotgun sequencing approach to investigate the microbiota in a given sample (Marchesi and Ravel, 2015). Metataxonomics (and metagenomics) can be a useful tool to investigate disease outbreaks of unknown aetiology. The results provide important information that can be used to guide further investigations such as target enrichment and subsequent whole genome sequencing. Technically, the metataxonomics approach follows the HTS methodology described under 5.2.3 with the difference that includes an amplicon-based library preparation, which makes the DNA fragmentation step obsolete.

5.2.2.1 A Metataxonomic Tool to Investigate the Diversity of *Treponema*

The following study established a targeted metataxonomic approach that was designed and validated to investigate the diversity of *Treponema* in complex biological samples. Through in silico analysis of the information content of the hypervariable regions of the 16S rRNA gene (V2-V8) we could show that the genus *Treponema* generally achieved a high operational taxonomic unit (OUT) resolution. Several validation steps show the usefulness of the assay, which was consequently tested in samples of Gilbert's potoroos (*Potorous gilbertii*) that are infected with a, yet, unclassified *Treponema* species.

frontiers in Microbiology

METHODS
published: 10 September 2019
doi: 10.3389/fmicb.2019.02094

A Metataxonomic Tool to Investigate the Diversity of *Treponema*

Luisa K. Hallmaier-Wacker[1,2], Simone Lüert[1,2], Sabine Gronow[3], Cathrin Spröer[3], Jörg Overmann[3,4], Nicky Buller[5], Rebecca J. Vaughan-Higgins[6] and Sascha Knauf[1]†*

[1] Neglected Tropical Diseases Work Group, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany, [2] Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany, [3] Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany, [4] Department of Microbiology, Braunschweig University of Technology, Braunschweig, Germany, [5] Animal Pathology – Bacteriology Laboratory, Department of Primary Industries and Regional Development, South Perth, WA, Australia, [6] Department of Conservation Medicine, College of Veterinary Medicine, School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA, Australia

OPEN ACCESS

Edited by:
Yasir Muhammad, King Abdulaziz University, Saudi Arabia

Reviewed by:
Naruo Nikoh, The Open University of Japan, Japan
Yuki Chan, The University of Hong Kong, Hong Kong

***Correspondence:**
Sascha Knauf
sknauf@dpz.eu

†Present address:
Sascha Knauf, Division of Microbiology and Animal Hygiene, University of Göttingen, Göttingen, Germany

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *13 June 2019*
Accepted: *26 August 2019*
Published: *10 September 2019*

Citation:
Hallmaier-Wacker LK, Lüert S, Gronow S, Spröer C, Overmann J, Buller N, Vaughan-Higgins RJ and Knauf S (2019) A Metataxonomic Tool to Investigate the Diversity of Treponema. Front. Microbiol. 10:2094. doi: 10.3389/fmicb.2019.02094

The genus *Treponema* contains a number of human and animal pathogenic as well as symbiotic bacteria that are found in vastly different anatomical and environmental habitats. Our understanding of the species range, evolution, and biology of these important bacteria is still limited. To explore the diversity of treponemes, we established, validated, and tested a novel metataxonomic approach. As the informative nature of the hypervariable regions of the 16S rRNA gene differ, we first analyzed each variable region independently. Considering the *in silico* results obtained, we established and validated the sequencing of the V4-region of the 16S rRNA gene using known mixtures of *Treponema* species as well as a selected number of clinical samples. The metataxonomic approach was able to identify *Treponema* to a near-species level. We demonstrate that using a spirochete-specific enrichment, our method is applicable to complex microbial communities and large variety of biological samples. The metataxonomic approach described provides a useful method to unravel the full diversity and range of *Treponema* in various ecosystems.

Keywords: metagenomics, metataxonomics, one health, spirochete, 16S rRNA, *Treponema*, marsupial, *Potorous*

INTRODUCTION

Spirochaetes, a phylum of spiral-shaped bacteria, range from pathogenic (e.g., *Treponema pallidum*) to symbiotic (e.g., *Sphaerochaeta coccoides*) to free-living (e.g., *Exilispira thermophile*) species (Paster, 2001). The ability of spirochetes to inhabit vastly anatomical and ecological habitats is remarkable and indicates a high diversity of the bacterial members of this phylum (Paster, 2001). Until recently, spirochetes were predominantly discovered and subsequently characterized using cultivation, microscopical, or serological approaches. These techniques make it difficult and sometimes impossible to characterize not-yet-cultivated species, to identify species in multiple-spirochete infections, or to discover commensal microbes. The advent of cultivation-independent molecular techniques (e.g., nucleic acid amplification technology) has allowed for a broader detection of *Treponema* in different biological niches. To date, the 16S rRNA phylogenetic marker gene has been particularly instrumental in the detection of *Treponema* diversity (Pace, 1997). Based on defined similarity thresholds the 16S rRNA sequences can be grouped into phylotypes. For

example, using a clonal 16S rRNA gene library and subsequent Sanger sequencing, the termite (*Reticulitermes flavipes*) gut was found to harbor more than 67 different treponemal phylotypes (Lilburn et al., 1999) and the human oral cavity up to 23 different treponemal clusters (Choi et al., 1994).

More recently, 16S rRNA gene-based metataxonomic studies have moved from clonal libraries to high-throughput sequencing approaches. Single hypervariable regions of the 16S rDNA have been used to examine different microbiomes (Kozich et al., 2013) and have identified treponemes in many ecological niches (Hong et al., 2012; Klitgaard et al., 2014; Rodriguez-R et al., 2015; Clayton et al., 2018; Hicks et al., 2018). For example, the gut microbiome of wild western lowland gorillas (*Gorilla gorilla gorilla*) (Hicks et al., 2018) and other nonhuman primates (Clayton et al., 2018) harbors multiple operational taxonomic units (OTUs) corresponding to the genus *Treponema*. Yet, conventional data analysis pipelines used in microbiome studies still do not allow for species-level characterization (Schloss et al., 2009; Caporaso et al., 2010). Taxonomic classification for many bacterial genera is restricted by the limited sequence differences in the 16S rRNA gene (Wang et al., 2007). Rossi-Tamisier et al. (2015) showed that spirochetes, in particular *Treponema* and *Spirochaeta*, have an exceptionally large variability in the 16S rRNA gene. For *Treponema*, only 2.1% of the analyzed 16S rRNA sequences fell within the recommended similarity threshold (95–98.7%) (Rossi-Tamisier et al., 2015).

To explore the range and diversity of *Treponema*, we established, validated, and tested a newly designed spirochete-specific metataxonomic approach that utilizes the 16S rRNA gene. Based on the known variability of the 16S rRNA gene, we hypothesized that a single hypervariable region of this gene provides a good target for a metagenomics-based assay to examine the diversity of *Treponema* in various biological sample types.

MATERIALS AND METHODS

In silico Analysis

Based on the nomenclature of Bergey's Manual of Systematic Bacteriology, we selected all bacterial species that are classified within the phylum *Spirochaetes* (Paster, 2001). Subsequently, a representative 16S rRNA gene sequence corresponding to each *Spirochaetes* bacterial species was retrieved from the GenBank database[1]. Where possible, sequences were chosen with maximal length and no ambiguous bases. Sequences shorter than 1,250 base pairs and/or containing more than two ambiguous bases were not included in the dataset even if no other sequence of the bacterial species was available (**Supplementary Table S1**).

The Perl-based high-throughput software tool V-Xtractor was used to locate the hypervariable regions (V2–V8) of the 16S rRNA sequences using Hidden Markov Models (Hartmann et al., 2010). Subsequently, the sequences of each variable region were analyzed using the mothur software package (v.1.41.1) (Schloss et al., 2009). In an initial step, identical sequences were removed using the unique.seq command. Then, the SILVA bacterial reference database (Quast et al., 2012) was utilized to align the sequences [align.seqs command using kmer searching (8mers) and Needleman–Wunsch pairwise alignment method]. OTU clustering was performed for distance threshold ranging from 0.01 to 0.10 at increments of 0.01 (cluster.split command with the OptiClust algorithm) (Westcott and Schloss, 2017).

[1]https://www.ncbi.nlm.nih.gov

Spirochete Mock Community

The spirochete mock community comprised an equal mixture of 19 strains of the phylum *Spirochaetes*. Single bacterial DNA isolates were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). DNA from rabbit inoculated *T. pallidum* subsp. *pertenue* strain Gauthier (referred to as *T. pallidum* throughout the manuscript) was obtained from David Šmajs, Department of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic. The 19 *Spirochaetes* species which were used in this study, including the cultivation method, DSMZ reference number, 16S rRNA gene copy number, genome size, and NCBI reference, are shown in **Supplementary Table S2**.

The DNA of the cultured spirochetes obtained from DSMZ was quantified using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific). *T. pallidum* DNA, due to the rabbit background DNA from *in vivo* inoculation experiments, was quantified using an established TaqMan PCR (qPCR) targeting the *polA* gene with slight modifications as described previously (Knauf et al., 2018). Based on the DNA content, genome size and 16S rRNA gene copy number, the 19 spirochetes were mixed together at equimolar (even) ribosomal RNA operon counts per organism. The final spirochete mock community contained 100,000 16S rDNA copies/µl of each species. All dilutions were made using Microbial DNA-Free water (Qiagen GmbH). Suitable precautions were taken during all sample handling and processing to avoid microbial contamination.

Treponema Mock Communities

In addition to the spirochete mock community, we created three bacterial DNA validation sets to evaluate the intra-metagenomic assays performance. *T. pallidum* DNA was quantified using TaqMan PCR as described above. For the first validation set, the stock of *T. pallidum* (50,000 16S rRNA copies) was used to make a 10-fold dilution series. The dilutions of *T. pallidum* DNA were subsequently mixed with bacterial DNA contained no *Spirochaetes* [Microbial Mock Community, HM-280, Biodefense and Emerging Infectious Research (BEI) Resources, Manassas, VA, United States] (**Supplementary Table S3**). The second validation set was a mixture of *T. pallidum* and *T. denticola* in different ratios (**Supplementary Table S4**). The final ratios of the *T. pallidum* to *T. denticola* were 1:100, 1:10, 1:1, 10:1, and 100:1. The third validation set was a 10-fold serial dilution series of *T. pallidum* starting at 50,000 copies of 16S rRNA gene. Dilutions for all validation sets were made using Microbial DNA-Free water (Qiagen GmbH).

Spirochete 16S Ribosomal RNA Gene Enrichment

Spirochete-selective primers were used to enrich spirochetal DNA (Dewhirst et al., 2010). The primers F24 (5′-GAGTTTGATYMTGGCTCAG-3′) and M98 (5′-GTTACGACTTCACCCYCCT-3′) were used to amplify a ∼1,450 bp fragment of the 16S rRNA gene covering the V1–V9 region. This first PCR step was performed in triplicates using the Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific), which has been validated for the use in microbiome studies (Hallmaier-Wacker et al., 2018). PCR reactions consisted of 12.5 µl of 2× PCR master mix, 9.5 µl of Microbial DNA-Free water (Qiagen GmbH), 1.0 µl of each primer (0.5 mM each, Metabion), and 1 µl of template in a total reaction volume of 25 µl. PCR cycling conditions comprised of a pre-denaturation step of 30 s at 98°C, followed by either 20 or 35 cycles of 98°C for 10 s, 57°C for 15 s and 72°C for 120 s, and a final 10 min extension step at 72°C. A 16S rRNA amplification control sample (blank controls; Microbial DNA-Free water) was included. Subsequently, PCR triplicates were pooled before library preparation.

Analysis of the V4 Region of the 16S rRNA Gene After an Initial Spirochete-Specific Amplification Step

A pre-test to re-amplify the V3, V4, and V6 regions was performed to identify the most suitable variable regions. The V4 region was selected, as the V3 and V6 region primers demonstrated technical issues to evenly amplify the variable regions. A modular, two-step PCR process was used to specifically re-amplify the V4-region of the 16S rRNA gene and prepare the samples for sequencing on the MiSeq platform. In the first step, the V4 region of the 16S rRNA gene was amplified using TruSeq adaptor-tailed universal primers 515F and 806R. The primers 515F-TruSeq (5′-ACACTCTTTCCCTCCACGACGCTCTTCCGCTCTGTGTGCCAGCMGCCGCGGTAA-3′) and 806R-TruSeq (5′-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCCGGACTACHVGGGTWTCTAAT-3′) were composed of the universal primer targeting the V4 region (Caporaso et al., 2011) followed by a linker and the TruSeq adaptor (Illumina, Inc.). Amplification was performed in triplicates and each 25.0 µl reaction contained 1.0 µl of PCR product of the enrichment step, 12.5 µl of 2× Phusion Hot Start II High-Fidelity PCR Master Mix (Thermo Fisher Scientific), 9.5 µl of Microbial DNA-Free water (Qiagen GmbH), and 1.0 µl of each V4-targeting 16S primer (0.5 mM each, Metabion). The cycling conditions were as follows: a pre-denaturation step of 30 s at 98°C, followed by 20 cycles of 98°C for 10 s, 55°C for 15 s and 72°C for 60 s, and a final 10 min extension step at 72°C. To monitor contamination, the blank control of the enrichment step was included as a 16S rDNA amplification control.

In the second-step PCR reaction, sample-specific Illumina indices and flow cell adapters were added in an indexing reaction. Illumina i7 and i5 indices were added to each amplicon using the indexing primer P5 (5′-AATGATACGGCGACCACCGAG ATCTACAC-[i5-INDEX]-ACACTCTTTCCCTACACGACGCTC-3′) and indexing primer P7 (5′-CAAGCAGAAGACGGCATACGAGAT-[i7-INDEX]-GTGACTGGAGTTCAGAC GTGT-3′). Amplification was performed in a 50.0 µl reaction containing 2.0 µl of PCR product from the first-step, 25.0 µl of 2× KAPA HiFi HotStart ReadyMix (KAPA Biosystems), 21.0 µl of Microbial DNA-Free water (Qiagen GmbH), and 1.0 µl of each Truseq index primer (0.5 mM each, Metabion). The cycling conditions were as follows: a pre-denaturation step of 3 min at 98°C, followed by eight cycles of 98°C for 20 s, 62°C for 30 s and 72°C for 30 s, and a final 5 min extension step at 72°C. To monitor overall contamination, the blank control of the first-step PCR reaction was included as a 16S rRNA gene amplification control.

V4-Region 16S rDNA Amplification Without an Initial Spirochete-Specific Amplification Step

For comparison the initial enrichment PCR was not performed on a sample of the spirochete mock community and a sample of validation set 1 (5,000 16S rDNA copies of *T. pallidum*). For these two samples the first-step V4-targeting PCR reaction was performed for additional 15 cycles (total of 35 cycles). A 16S rRNA gene amplification control was included for this altered procedure. All other conditions were kept identical.

Applications to Clinical Samples

The applicability of the metataxonomic approach was tested on extracted DNA from genital swabs of Gilbert's potoroo (*Potorous gilbertii*), a small marsupial found in Western Australia (Vaughan et al., 2009). For more information on sample processing see the **Supplementary Materials**.

MiSeq Library Preparation and Pooling

After the indexing reaction, all amplicons were purified using 0.7× AMPure XP beads (Beckman Coulter), and quantified using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific) (**Supplementary Table S5**). The amplicon integrity was verified for a representative number of four samples using the BioAnalyzer 2000 (Agilent). Equimolar amounts (2 nM) of sample amplicons were pooled. For samples with <2 nM concentration, the maximum volume (5 µl) was pooled prior to sequencing. The Transcriptome and Genome Analysis Laboratory at the University of Goettingen performed the Illumina MiSeq 2 × 250 bp paired-end sequencing (Illumina V2 chemistry) run.

Data Processing and Analysis

Raw reads were processed using the mothur software package (version 1.41.1) (Schloss et al., 2009). Initial pre-processing and quality control were performed in accordance with the MiSeq SOP (Schloss et al., 2009). Briefly, paired-end reads were assembled using the make.contigs command. Subsequently, the screen.seqs command was used to trim sequences and filter out any sequences with ambiguous base calls. Identical trimmed sequences (unique.seq command) were aligned (align.seqs command) to the SILVA bacterial

TABLE 1 | Identifiable *in silico* OTUs for the different genus within the phylum of *Spirochaetes* at a 97% threshold.

Variable region	Total OTU	*Treponema* (*n* = 28)[#]	*Sphaerochaeta* (*n* = 4)	*Spirochaeta* (*n* = 16)	*Leptospira* (*n* = 21)	*Exilispira* (*n* = 1)	*Leptonema* (*n* = 1)	*Borrelia* (*n* = 30)	*Brachyspira* (*n* = 14)	*Spironema* (*n* = 1)
V2	69	25	3	14	11	1	1	5	8	1
V3	50	25	2	12	4	1	1	3	1	1
V4	53	24	3	13	3	1	1	4	3	1
V5	47	22	3	12	2	1	1	2	3	1
V6	53	24	4	12	6	1	1	2	3	1
V7	34	12	2	11	3	1	1	2	1	1
V8	47	23	2	11	5	1	1	1	2	1

[#]*(n) indicates the number of unique representative sequences in the in silico.fasta file for each genus.*

reference database (Quast et al., 2012). Poorly aligned sequences, chimeras [chimera.uchime command; UCHIME algorithm (Edgar et al., 2011)], and other erroneous non-bacterial sequences (remove.lineage command) were removed. The remaining sequences were classified using a Bayesian classifier implemented in mothur and OTUs were assigned based on a distance threshold of 0.03.

For the species-level classification, *Treponema*-classified sequences were extracted from the dataset using the get.lineage command. Using the *Treponema* sequence data in the *in silico* fasta file (**Supplementary Table S1**), a database was created using the create.database command. Using this database, the taxonomy of the filtered sequences was assigned using the classify.otu command.

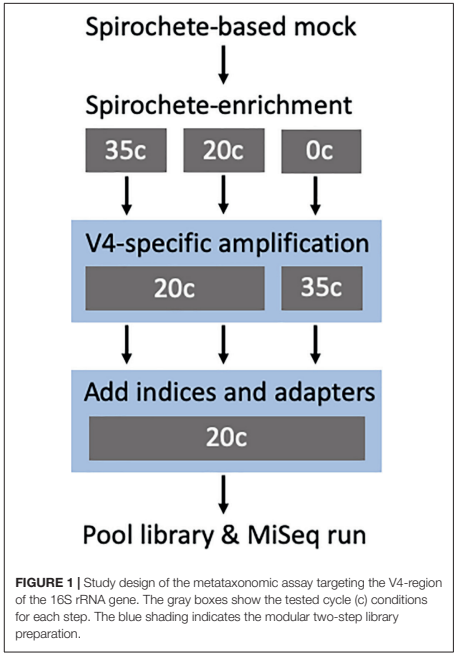


FIGURE 1 | Study design of the metataxonomic assay targeting the V4-region of the 16S rRNA gene. The gray boxes show the tested cycle (c) conditions for each step. The blue shading indicates the modular two-step library preparation.

Data Availability

All generated read files have been deposited in the NCBI Sequence Read Archive under the accession number PRJNA541286.

RESULTS

In silico Analysis of the Information Content of the V2–V8 Regions of the 16S rRNA Gene

We analyzed each hypervariable region (V2–V8) of the 16S rRNA gene for its potential to distinguish nine bacterial genera that make up the phylum *Spirochaetes*. In total, we analyzed the information content of the variable regions of 114 representative sequences *in silico* (**Supplementary Table S1**). Hypervariable regions V2–V8 were able to distinguish the nine bacterial genera at a similarity threshold of 97% (**Table 1**). The V2 region identified the largest total number of OTUs (n = 69) compared to all other tested regions (**Table 1**). Overall, the least number of OTUs were identified in the genera *Leptospira*, *Borrelia*, and *Brachyspira*. For the genus *Treponema* on the other hand, all variable regions with the exception of V7 were able to detect a high number of distinct OTUs (**Table 1**). Regions V2 and V3 were both able to detect 25 OTUs at a threshold of 97% in the *in silico* dataset containing 28 unique representative sequences. To examine the robustness of the *in silico* results for *Treponema*, we examined the identifiable OTUs at threshold cutoffs ranging from 90 to 99% (**Supplementary Figure S1**). For V2, V3, and V4 regions at 90% similarity threshold, >15 OTUs are distinguishable in the *Treponema* genus (**Supplementary Figure S1**).

V4-Region 16S rDNA Amplification of the Spirochete Mock Community

We tested three different amplification conditions targeting the V4-region of the 16S rRNA gene (**Figures 1**, **2**). All 16S rRNA gene amplification conditions were able to identify all seven

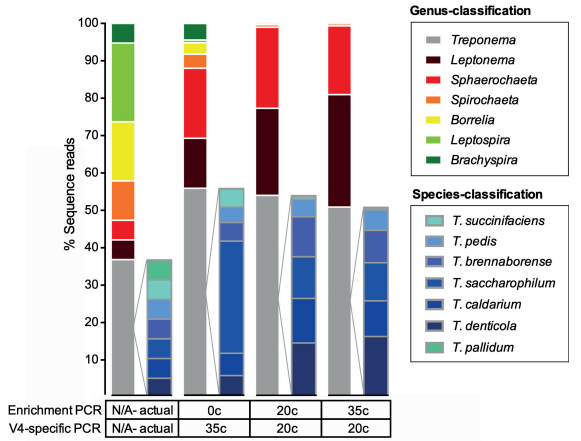


FIGURE 2 | Actual and observed proportions of a spirochete mock community to a genus and species-level. Relative abundance of OTUs in percentage of reads for a spirochete microbial mock community amplified using multiple methods. The spirochete enrichment and V4-region-specific PCR cycle numbers are shown below the taxa plots. The expected strain proportion (actual) of the spirochete mock community represents the theoretical composition and was corrected for differential copy number of the 16S rRNA gene. Species-level classification only shown for the genus *Treponema*.

genera that were included in the mock community samples (**Figure 2**). The amplification condition without spirochete-specific 16S rRNA gene enrichment differed less from the actual mixing proportion than the samples which were enriched (**Figure 2**). In all conditions, *Treponema*, *Leptonema*, and *Sphaerochaeta* were preferentially detected. For the spirochete enriched samples (20 cycles and 35 cycles), *Borrelia*, *Leptospira*, and *Brachyspira* made up <1% of the detected sequence reads (**Figure 2**). In addition to the genus-level identification, we classified the *Treponema* sequences on a species-level using a *Treponema*-specific database of the V4-region. At 97% similarity threshold, the database contains 24 OTUs of which 21 OTUs correspond to single species and 3 OTUs correspond to species-clusters (*denticola*-, *medium*-, and the *pallidum*-cluster) (**Supplementary Table S6**). All 16S rDNA amplification conditions were able to identify all seven species of *Treponema* in the mock community (**Figure 2**). Independent of amplification conditions, sequences corresponding to *T. pallidum*-cluster were amplified less efficiently.

Amplification control samples (blank samples) were included for each amplification method to test for contamination during the amplification process. The blank sample from the 20-cycle enrichment had the lowest amplicon quantity before sequencing as well as the lowest overall corresponding number of sequences reads (**Supplementary Figure S2** and **Supplementary Table S5**). Compared to the blank control enriched for 20 cycles, the control of the enrichment for 35 cycles had a 100× fold increase in sequence reads (**Supplementary Figure S2**). Despite an overall lower cycle count, the sequence reads corresponding to the nonenriched sample were as high as for the control enriched for 35 cycles. Unlike the enriched samples (20 cycle and 35 cycle), which detected minimal *Treponema* in the blank sample (<10 sequence reads), the non-enriched control sample detected 6,364 reads of *Treponema* (**Supplementary Figure S2**).

Intra-Metagenomic Assays Performance

The validation sample sets were used to assess the efficiency of the spirochete enrichment amplification, the effect of competition between two species, and the detection limit of the amplicon sequencing approach. The first validation set was a mixture of different concentrations of *T. pallidum* with a microbial mock community (HM-280). **Figure 3** shows that both 20 cycle- and 35 cycle-enrichment steps significantly improve the detection of *T. pallidum* compared to the unenriched sample at 5,000 16S rRNA gene copies of *T. pallidum*. As the input DNA of *T. pallidum* decreases, the dilution effect between *T. pallidum* to microbial mock community HM-280 can be visualized clearly (**Figure 3**). Four to six 16S rRNA gene copies of *T. pallidum* were detected for 35 cycle- (44,141 sequence reads) and 20 cycle-enrichment (2,224 sequence reads).

The second validation set was a mixture of *T. pallidum* and *T. denticola* in different ratios (see the section "Materials and Methods" for details). For all ratios, *T. denticola* outcompeted *T. pallidum* in detected sequence reads (**Table 2**). However, both species of *Treponema* were detected at all tested ratios (**Table 2**). The third validation set was a 10-fold serial dilution series of *T. pallidum*. Sequence reads were >9,000

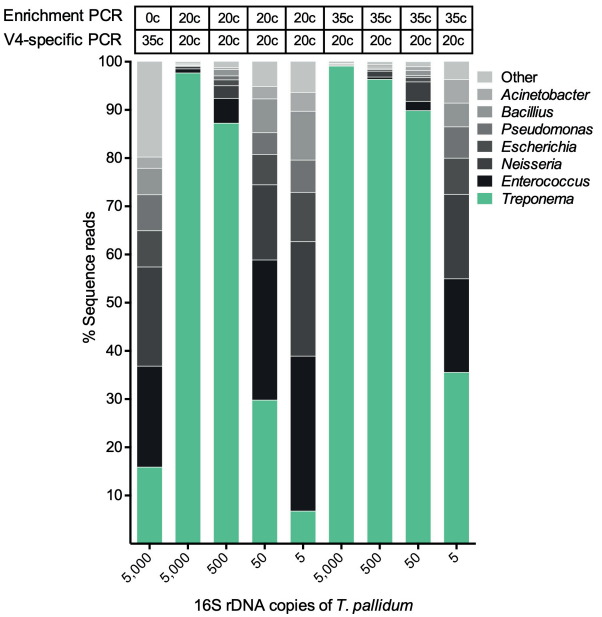


FIGURE 3 | The effect of spirochete enrichment on the detection limit of *T. pallidum*. Relative abundance of OTUs in percentage of reads for a *T. pallidum*-spiked microbial mock community (HM-280). The theoretically calculated amount of spike in *T. pallidum* ranging from 5 to 5,000 16S rDNA copies is shown along the *x*-axis. The biologically effective range of *T. pallidum* may differ from theoretically calculated amount (e.g., 5 copies = 4–6 copies). Cycle number for the spirochete enrichment and V4-region-specific PCR is shown above the taxa plots.

reads down to 500 16S rRNA gene copies of *T. pallidum* (**Figure 4**). At 50 16S rRNA gene copies, *T. pallidum* was detectable but overall sequence reads were markedly decreased (13,453 sequence reads). For the final two dilutions, total read numbers were <1,500 and *T. pallidum* sequence detection was analogous to the blank control (<10 sequence reads) (**Figure 4**).

Applications to Clinical Samples (Gilbert's Potoroo)

We examined samples from four Gilbert's potoroo which had been found to harbor a *Treponema* infection (Vaughan et al., 2009). Using the amplicon sequencing technique, we identified a *Treponema* species in all four clinical samples (**Supplementary Figure S3**). Sequence reads corresponding to the *Treponema* made up >75% of the total read count for Gilbert's potoroo samples No. 2–4 (**Supplementary Figure S3**). The *Treponema* sequences clustered into a single OTU, which cannot be identified using the *Treponema*-specific V4-region database at a 97% threshold identity.

TABLE 2 | Relative abundance of OTUs in percentage of reads for different proportions of *T. pallidum* and *T. denticola* 16S rRNA gene.

Ratio of *T. pallidum* and *T. denticola* 16S rRNA gene	% sequence reads for *T. pallidum* (read count)	% sequence reads for *T. denticola* (read count)
1:100	0.1 (37)	99.9 (81,032)
1:10	0.3 (350)	99.7 (112,402)
1:1	3.0 (3,108)	97.0 (100,507)
10:1	26.6 (31,077)	73.4 (85,750)
100:1	77.3 (74,179)	22.7 (21,809)

DISCUSSION

We used the *in silico* analysis to predict the informative nature of each 16S rDNA hypervariable region for different spirochetes. Our findings expand on the results of Rossi-Tamisier et al. (2015), indicating that three genera of spirochetes have low interspecies sequence similarity with the 16S rRNA gene, which makes it a suitable gene target for identification (**Table 1**). The

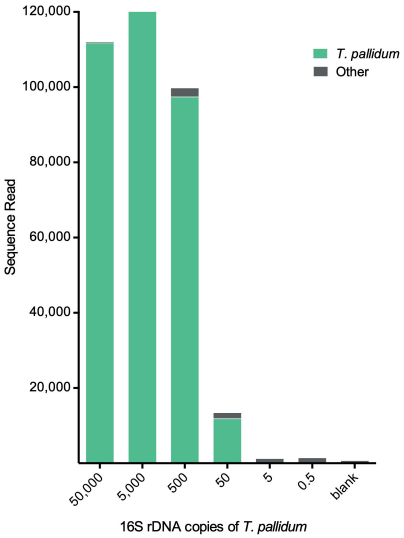


FIGURE 4 | Detection limit of the metagenomic approach for *Treponema*. Total sequence reads resulting from different input amounts of *T. pallidum*. Displayed amounts of *T. pallidum* (50,000–0.5 16S rDNA copies) represent theoretically calculated amounts. Biologically effective range may differ from theoretically calculated amount (e.g., 0.5 copies = 0–2 copies). Blank control represents the 16S rDNA amplification control using microbial DNA-free water as input. For these samples 20 cycles of enrichment PCR was followed by 20 cycles of V4-region-specific PCR (for more detail see the section "Materials and Methods").

V3 and V4 region have been previously described for their discriminatory power (Chakravorty et al., 2007; Yang et al., 2016; Graspeuntner et al., 2018). In a study of 110 bacterial species, V2 and V3 were the most suitable candidates (Chakravorty et al., 2007). Considering phylogenetic resolution, the variable regions 4, 5, and 6 have been previously identified as prime targets (Yang et al., 2016; Graspeuntner et al., 2018). Overall, the *in silico* analysis provided good initial data to efficiently design the *in vitro* experiments. It is, however, important to note that technical caveats of NGS sequencing must be considered prior to the selection of the most appropriate region (Kozich et al., 2013). For example, for the Illumina MiSeq Platform, paired-end sequencing can currently cover 300 base pairs. Considering the overall error rate of this platform [∼0.1–0.01% per base, depending on the data-filtering scheme (Meacham et al., 2011; Loman et al., 2012)], the ideal read length for a metataxonomic approach allows for full overlap of the two pair-end reads (Kozich et al., 2013). Based on our *in silico* results and pre-test using different primers, we selected the V4 region of the 16S rRNA gene for further *in vitro* testing.

A spirochete mock community of known species composition allowed for the systematic comparison between the different amplification methods (**Figure 1**; Brooks et al., 2015). Independent of the amplification method, our metagenomic approach was able to detect all seven genera of spirochetes in the mock community, as well as all species of *Treponema* (**Figure 2**). However, not all spirochetes were detected equally well with all amplification procedures (**Figure 2**). The spirochete-specific enrichment step, which was included for a better detection of spirochetes, led to the distortion of the actual proportions and favored *Treponema*, *Sphaerochaeta*, and *Leptonema* (**Figure 2**). The distortion of the bacterial profiles due to preferential amplification of multi-template PCR is a known phenomenon and a major limitation of 16S rRNA gene amplification that results from sub-optimal primer binding (Polz and Cavanaugh, 1998; Brooks et al., 2015; Hallmaier-Wacker et al., 2018). It has been previously shown that this distortion effect is not significantly influenced with decreasing the number of amplification cycles (Acinas et al., 2005; Sipos et al., 2007; Wu et al., 2010; Brooks et al., 2015). Similarly, our results did not remarkably change with an increased number of enrichment cycles (20 cycles vs. 35 cycles; **Figure 2**). Nevertheless, the use of unnecessary cycles should be avoided as it can lead to formation of unwanted side products such as chimeras (Ahn et al., 2012), as well as a higher risk of overamplifying reads that originate from contamination (blank controls; **Supplementary Figure S2**) (Salter et al., 2014).

To examine the benefits of the spirochete enrichment PCR (20 cycles and 35 cycles) on the detection limit of analysis, we tested the metataxonomic approach on mock communities that simulate bacterial proportions found in clinical samples (**Supplementary Table S3**). For these samples, the enrichment step critically improved the detection of *Treponema* at low copy numbers, thus indicating that enrichment is a useful tool for samples with low spirochete numbers (<5,000 16S rDNA copies) (**Figure 3**). We showed that five 16S rRNA gene copies of *T. pallidum* were detectable in a sample with 20 other bacterial species (even bacterial mock HM-280). Using serial dilutions, we were able to detect as little as 50 16S rRNA gene copies of *T. pallidum* using 20 cycle enrichment amplification (**Figure 4**). These data indicate a sensitivity of our assay that is comparable to standard TaqMan qPCR and which outcompetes the conventional 16S rRNA clonal approach (Leslie et al., 2007). Obtaining a high detection limit using a clonal approach is both time consuming and resource intensive (Leigh et al., 2010). On the other hand, Sanger sequence analysis of clone libraries provide greater phylogenetic resolution due to an increased read length, covering the full 16S rRNA gene (Leigh et al., 2010). The complex microbial communities present in many clinical samples is a frequent challenge in diagnostics and in these sample the occurrence of multi-*Treponema* species is not uncommon. For example, in oral syphilis patients, *T. pallidum* can be found in combination with *T. denticola* (Scott and Flint, 2005). We therefore tested the effect of competing species by simulating a co-infection of *T. pallidum* and *T. denticola* (**Supplementary Table S4**). Overall, the metataxonomic approach underestimated the ratio

of *T. pallidum* in the samples (**Table 2**). Amplicon sequencing was, however, sufficient to identify both species of *Treponema* at all tested ratios (**Table 2**). We note here that the metataxonomic approach does not accurately represent absolute abundance of different species (Widder et al., 2016). The used primers may have a significant effect in distorting tested ratios and thus alternative primer should be designed and evaluated for specific research questions. Additionally, quantitative techniques such as qPCR, flow cytometry, or fluorescence *in situ* hybridization (FISH) may be superior for evaluating known competing species (e.g., *T. pallidum and T. denticola*) (Props et al., 2017). Moreover, the metataxonomic approach should not be used for defining novel bacterial species even if species-level clustering is possible using the 16S rRNA sequence information (e.g., *Treponema*) (Tindall et al., 2010). Instead, 16S rRNA gene amplicon sequencing provides a qualitative view on the diversity of treponemes within a given DNA sample. For example, we used the metataxonomic approach to examine clinical samples of the Gilbert's potoroo that have been previously found to harbor a *Treponema* infection (Vaughan et al., 2009). As the potoroo clinical samples were associated with a polymicrobial environment and infection was believed to be chronic (Vaughan et al., 2009), we performed a 35 cycle-enrichment in order to detect low concentrations of spirochetes. We identified a single *Treponema* species in all tested samples of the four Gilbert's potoroos, which currently remains unclassified at a species level. The high percentage of *Treponema* in the detected samples (>75%; **Supplementary Figure S3**) indicates that the amplicon method is applicable for clinical samples and guides subsequent approaches that aim to fully characterize the discovered *Treponema* species. The results from the metataxonomic approach can be used to select most promising samples for whole-genome analysis (WGS), as well as provide a preliminary understanding of the possible phylogeny, which may assist in reference-based assembly (Wyres et al., 2014). Importantly, further WGS sequences of known and unknown *Treponema* are crucial to study the evolution and epidemiology of this ancient group of bacteria and to enhance future shotgun metagenomic studies. Currently, there is only a limited number of whole-genome sequences of *Treponema*, in particular the non-pathogenic species, due to the difficulty to culture many of the species [e.g., from the termite gut (Paster et al., 1996)].

CONCLUSION

We showed that the V4 region of the 16S rRNA gene is a valuable target to explore the diversity of *Treponema* in various biological sample types. To monitor the quality of each sequencing run, it is essential to including relevant controls with all clinical samples. When applied appropriately, the presented modular metataxonomic approach is broadly applicable as it requires only small amounts of bacterial DNA for the detection of a broad range of *Treponema* species.

DATA AVAILABILITY

The datasets generated for this study can be found in the NCBI Sequence Read Archive accession number PRJNA541286.

AUTHOR CONTRIBUTIONS

LH-W and SK conceived and designed the study. LH-W, SL, SG, CS, and SK performed the experiments in the laboratory. LH-W, SL, and SK analyzed the data. NB and RV-H contributed DNA from the Gilbert's potoroos. SG, CS, and JO contributed DNA samples of spirochetes for the mock sample. All authors contributed to the writing of the manuscript, read, reviewed, and approved the final manuscript.

FUNDING

SK received funding to conduct parts of this study by the German Research Foundation (DFG KN1097/3-2 and KN1097/7-1). The funders had no role in any part of this study.

ACKNOWLEDGMENTS

We thank the Biodefense and Emerging Infectious Research (BEI) Resources, NIAID, NIH for providing the cells from Microbial Mock Community (Even, HM-280) as part of the Human Microbiome Project. We thank David Šmajs of the Department of Biology, Faculty of Medicine at the Masaryk University for providing DNA from the *T. pallidum* subsp. *pertenue* strain Gauthier. Additionally, we thank Christian Roos (German Primate Center) and Fabian Ludewig (Transcriptome and Genome Analysis Laboratory at the University of Göttingen) for their assistance in optimizing the sequencing run. Finally, we thank Simone Severitt and Carola Berg (both DSMZ) for excellent technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.02094/full#supplementary-material

REFERENCES

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., and Polz, M. F. (2005). PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Appl. Environ. Microbiol.* 71, 8966–8969. doi: 10.1128/AEM.71.12.8966-8969.2005

Ahn, J.-H., Kim, B. Y., Song, J., and Weon, H. Y. (2012). Effects of PCR cycle number and DNA polymerase type on the 16S rRNA gene pyrosequencing analysis of bacterial communities. *J. Microbiol.* 50, 1071–1074. doi: 10.1007/s12275-012-2642-z

Brooks, J. P., Edwards, D. J., Harwich, M. D., Rivera, M. C., Fettweis, J. M., Serrano, M. G., et al. (2015). The truth about metagenomics: quantifying and

counteracting bias in 16S rRNA studies. *BMC Microbiol.* 15:66. doi: 10.1186/s12866-015-0351-6

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4516–4522. doi: 10.1073/pnas.1000080107

Chakravorty, S., Helb, D., Burday, M., Connell, N., and Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods* 69, 330–339. doi: 10.1016/j.mimet.2007.02.005

Choi, B. K., Paster, B. J., Dewhirst, F. E., Göbel, U. B. (1994). Diversity of cultivable and uncultivable oral spirochetes from a patient with severe destructive periodontitis. *Infect. Immun.* 62, 1889–1895.

Clayton, J. B., Gomez, A., Amato, K., Knights, D., Travis, D. A., Blekhman, R., Knight, R., et al. (2018). The gut microbiome of nonhuman primates: lessons in ecology and evolution. *Am. J. Primatol.* 30:e22867. doi: 10.1002/ajp.22867

Dewhirst, F. E., Chen, T., Izard, J., Paster, B. J., Tanner, A. C., Yu, W.-H., et al. (2010). The human oral microbiome. *J. Bacteriol.* 192, 5002–5017. doi: 10.1128/JB.00542-10

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Graspeuntner, S., Loeper, N., Künzel, S., Baines, J. F., and Rupp, J. (2018). Selection of validated hypervariable regions is crucial in 16S-based microbiota studies of the female genital tract. *Sci. Rep.* 8:9678. doi: 10.1038/s41598-018-27757-8

Hallmaier-Wacker, L. K., Lueert, S., Roos, C., and Knauf, S. (2018). The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. *Sci. Rep.* 8:6292. doi: 10.1038/s41598-018-24573-y

Hartmann, M., Howes, C. G., Abarenkov, K., Mohn, W. W., and Nilsson, R. H. (2010). V-Xtractor: an open-source, high-throughput software tool to identify and extract hypervariable regions of small subunit (16S/18S) ribosomal RNA gene sequences. *J. Microbiol. Methods* 83, 250–253. doi: 10.1016/j.mimet.2010.08.008

Hicks, A. L., Lee, K. J., Couto-Rodriguez, M., Patel, J., Sinha, R., Guo, C., et al. (2018). Gut microbiomes of wild great apes fluctuate seasonally in response to diet. *Nat. Commun.* 9:1786. doi: 10.1038/s41467-018-04204-w

Hong, P. Y., Li, X., Yang, X., Shinkai, T., Zhang, Y., Wang, X., et al. (2012). Monitoring airborne biotic contaminants in the indoor environment of pig and poultry confinement buildings. *Environ. Microbiol.* 14, 1420–1431. doi: 10.1111/j.1462-2920.2012.02726.x

Klitgaard, K., Nielsen, M. W., Ingerslev, H.-C., Boye, M., and Jensen, T. K. (2014). Discovery of bovine digital dermatitis-associated Treponema spp. in the dairy herd environment by a targeted deep-sequencing approach. *Appl. Environ. Microbiol.* 80, 4427–4432. doi: 10.1128/AEM.00873-14

Knauf, S., Lüert, S., Šmajs, D., Strouhal, M., Chuma, I. S., Frischmann, S., et al. (2018). Gene target selection for loop-mediated isothermal amplification for rapid discrimination of *Treponema pallidum* subspecies. *PLoS Negl. Trop. Dis.* 12:e0006396. doi: 10.1371/journal.pntd.0006396

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79, 5112–5120. doi: 10.1128/AEM.01043-13

Leigh, M. B., Taylor, L., Neufeld, J. D. (2010). "Clone libraries of ribosomal RNA gene sequences for characterization of bacterial and fungal communities," in *Handbook of hydrocarbon and Lipid Microbiology*, ed. K. N. Timmis (Berlin: Springer) 3969–3993. doi: 10.1007/978-3-540-77587-4_310

Leslie, D. E., Azzato, F., Karapanagiotidis, T., Leydon, J., and Fyfe, J. (2007). Development of a real-time PCR assay to detect Treponema pallidum in clinical specimens and assessment of the assay's performance by comparison with serological testing. *J. Clin. Microbiol.* 45, 93–96. doi: 10.1128/JCM.01578-06

Lilburn, T. G., Schmidt, T. M., Breznak, J. A. (1999). Phylogenetic diversity of termite gut spirochaetes. *Environ. Microbiol.* 1, 331–345. doi: 10.1046/j.1462-2920.1999.00043.x

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., et al. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* 30, 434–439. doi: 10.1038/nbt.2198

Meacham, F., Boffelli, D., Dhahbi, J., Martin, D. I., Singer, M., and Pachter, L. (2011). Identification and correction of systematic error in high-throughput sequence data. *BMC Bioinformatics* 12:451. doi: 10.1186/1471-2105-12-451

Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. *Science* 276, 734–740. doi: 10.1126/science.276.5313.734

Paster, B. J. (2001). "Phylum XV. *Spirochaetes*," in *Bergey's Manual of Systematic Bacteriology*, 2nd Edn, Vol. 4. eds N. R. Krieg, J. T. Staley, D. R. Brown, B. P. Hedlund, B. J. Paster, N. L. Ward, et al. (New York, NY: Springer) 471–566.

Paster, B. J., Dewhirst, F. E., Cooke, S. M., Fussing, V., Poulsen, L. K., and Breznak, J. A. (1996). Phylogeny of not-yet-cultured spirochetes from termite guts. *Appl. Environ. Microbiol.* 62, 347–352.

Polz, M. F., and Cavanaugh, C. M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Appl. Environ. Microbiol.* 64, 3724–3730.

Props, R., Kerckhof, F.-M., Rubbens, P., De Vrieze, J., Sanabria, E. H., Waegeman, W., et al. (2017). Absolute quantification of microbial taxon abundances. *ISME J.* 11, 584–587. doi: 10.1038/ismej.2016.117

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Rodriguez-R, L. M., Overholt, W. A., Hagan, C., Huettel, M., Kostka, J. E., and Konstantinidis, K. T. (2015). Microbial community successional patterns in beach sands impacted by the Deepwater Horizon oil spill. *ISME J.* 9, 1928–1940. doi: 10.1038/ismej.2015.5

Rossi-Tamisier, M., Benamar, S., Raoult, D., and Fournier, P. E. (2015). Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *Int. J. Syst. Evol. Microbiol.* 65, 1929–1934. doi: 10.1099/ijs.0.000161

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12:87. doi: 10.1186/s12915-014-0087-z

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Scott, C., and Flint, S. R. (2005). Oral syphilis—re-emergence of an old disease with oral manifestations. *Int. J. Oral Maxillofac. Surg.* 34, 58–63. doi: 10.1016/j.ijom.2004.01.029

Sipos, R., Székely, A. J., Palatinszky, M., Révész, S., Márialigeti, K., and Nikolausz, M. (2007). Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiol. Ecol.* 60, 341–350. doi: 10.1111/j.1574-6941.2007.00283.x

Tindall, B. J., Rosselló-Móra, R., Busse, H. J., Ludwig, W., and Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol.* 60, 249–266. doi: 10.1099/ijs.0.016949-0

Vaughan, R. J., Warren, K. S., Mills, J. S., Palmer, C., Fenwick, S., Monaghan, C. L., et al. (2009). Hematological and serum biochemical reference values and cohort analysis in the Gilbert's potoroo (*Potorous gilbertii*). *J. Zoo Wildl. Med.* 40, 276–288. doi: 10.1638/2008-0058.1

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Westcott, S. L., and Schloss, P. D. (2017). OptiClust, an improved method for assigning amplicon-based sequence data to operational taxonomic units. *mSphere* 2:e00073-17. doi: 10.1128/mSphereDirect.00073-17

Widder, S., Allen, R. J., Pfeiffer, T., Curtis, T. P., Wiuf, C., Sloan, W. T., et al. (2016). Challenges in microbial ecology: building predictive understanding of

community function and dynamics. *ISME J.* 10, 2557–2568. doi: 10.1038/ismej.2016.45

Wu, J. Y., Jiang, X. T., Jiang, Y. X., Lu, S. Y., Zou, F., and Zhou, H. W. (2010). Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. *BMC Microbiol.* 10:255. doi: 10.1186/1471-2180-10-255

Wyres, K. L., Conway, T. C., Garg, S., Queiroz, C., Reumann, M., Holt, K., et al. (2014). WGS analysis and interpretation in clinical and public health microbiology laboratories: what are the requirements and how do existing tools compare? *Pathogens* 3, 437–458. doi: 10.3390/pathogens3020437

Yang, B., Wang, Y., and Qian, P. Y. (2016). Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* 17:135. doi: 10.1186/s12859-016-0992-y

Conflict of Interest Statement: SG, CS, and JO were employed by the Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

CORRECTION
published: 08 November 2019
doi: 10.3389/fmicb.2019.02581

Corrigendum: A Metataxonomic Tool to Investigate the Diversity of *Treponema*

Luisa K. Hallmaier-Wacker[1,2], Simone Lüert[1,2], Sabine Gronow[3], Cathrin Spröer[3], Jörg Overmann[3,4], Nicky Buller[5], Rebecca J. Vaughan-Higgins[6] and Sascha Knauf[1†]*

[1] Neglected Tropical Diseases Work Group, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany, [2] Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany, [3] Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany, [4] Department of Microbiology, Braunschweig University of Technology, Braunschweig, Germany, [5] Animal Pathology – Bacteriology Laboratory, Department of Primary Industries and Regional Development, South Perth, WA, Australia, [6] Department of Conservation Medicine, College of Veterinary Medicine, School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA, Australia

Keywords: metagenomics, metataxonomics, one health, spirochete, 16S rRNA, *Treponema*, marsupial, *Potorous*

OPEN ACCESS

Edited by:
Yasir Muhammad,
King Abdulaziz University, Saudi Arabia

Reviewed by:
J. Christopher Fenno,
University of Michigan, United States

****Correspondence:***
Sascha Knauf
sknauf@dpz.eu

†Present address:
Sascha Knauf,
Division of Microbiology and Animal Hygiene, University of Göttingen, Göttingen, Germany

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *18 September 2019*
Accepted: *24 October 2019*
Published: *08 November 2019*

Citation:
Hallmaier-Wacker LK, Lüert S, Gronow S, Spröer C, Overmann J, Buller N, Vaughan-Higgins RJ and Knauf S (2019) Corrigendum: A Metataxonomic Tool to Investigate the Diversity of Treponema. Front. Microbiol. 10:2581. doi: 10.3389/fmicb.2019.02581

A Corrigendum on

A Metataxonomic Tool to Investigate the Diversity of *Treponema*
by Hallmaier-Wacker, L. K., Lüert, S., Gronow, S., Spröer, C., Overmann, J., Buller, N., et al. (2019). Front. Microbiol. 10:2094. doi: 10.3389/fmicb.2019.02094

In the original article, there was a mistake in **Figure 1** as published. The figure contains an error in the number of cycles for the "add indices and adapters" step. The step should read 8 cycles instead of 20 cycles. The corrected **Figure 1** appears below.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

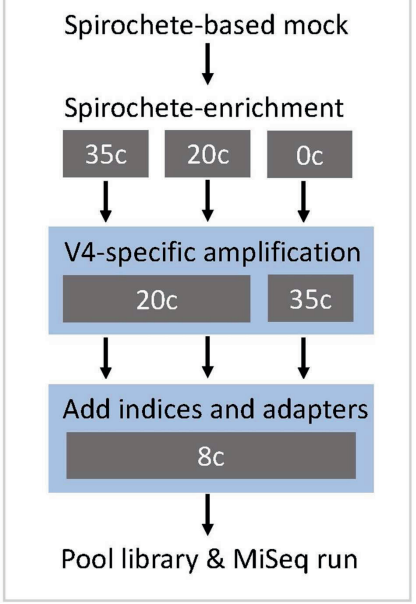


FIGURE 1 | Study design of the metataxonomic assay targeting the V4-region of the 16S rRNA gene. The gray boxes show the tested cycle (c) conditions for each step. The blue shading indicates the modular two-step library preparation.

5.2.3 High-Throughput Whole Genome Sequencing

The bacteriophage φX174 was the first pathogen whose genome was fully sequenced (5Mbp) (Maxam and Gilbert, 1977; Sanger et al., 1977b). This achievement was further surpassed with the sequencing of the much greater Epstein-Barr virus in 1984 (170 Mbp) (Baer et al., 1984). Since then whole genome sequencing (WGS) has become standard, and two decades after the Epstein-Barr virus was sequenced with Sanger technology, a new sequencing generation (HTS technology) was introduced to the market. This technique allowed researchers to push the boundaries of sequencing in terms of costs, speed, automatization of the remaining manual processes and sequencing accuracy (El-Metwally et al., 2014) to never-before-seen dimensions. The sequencing of the human genome is now possible within a single day and at costs as low as 1,000 USD, compared to the first human genome sequence that was generated over years and which took hundreds of millions of USDs (Collins and Hamburg, 2013). Since 2005 marked the starting point for the HTS technology (El-Metwally et al., 2014), different sequencing biochemistries (Roche (454) (Margulies et al., 2005), Illumina (Solex) (Bentley et al., 2008) and Life Technologies (ABI/SOLiD) (Shendure and Ji, 2008)) have been introduced. Today, Illumina's Sequencing-by-Synthesis (SBS) technology is the most widely used and accepted HTS biochemistry (Buermans and den Dunnen, 2014). Briefly, Illumina HTS workflows contain four steps: library preparation, cluster generation, sequencing, and data analysis (Illumina, 2017). DNA libraries are generated from random fragmentation of the target DNA (or cDNA) samples, either through physical shattering or enzymatic digest. Resulting DNA fragments are 5'- and 3'-adapter-ligated and subsequently, PCR amplified. In the next step the library is loaded into flow cells where target DNA is bound to surface-coated oligos complementary to the library adapters. Amplification of the bound DNA fragments generates clonal clusters (cluster generation) that are ready for sequencing. Illumina SBS uses a proprietary reversible terminator-based method that detects single nucleotides as they are incorporated into the DNA template strands. Resulting paired-sequence reads are adapter and low quality trimmed, length and duplicate-sequence filtered, as well as error corrected before they can be reference-, de novo-, or a combination of both, aligned.

A major weakness of the Illumina HTS technology is that DNA libraries are length restricted. It seems bizarre to fragment high-quality DNA into read pieces of e.g. 250 base pairs (bps), just to put those pieces back into place to align the full genome. Especially areas of repetitive sequence parts or genomes that contain a homologous gene family such as the *tpr*-gene family in *T. pallidum*, are prone to errors and fail to be adequately aligned. A relatively new HTS sequencing technology often referred to as the third-generation of sequencing, is the Oxford Nanopore Technology (ONT). It was introduced to the market in 2014 (Lu et al., 2016) and holds promise to overcome some of the Illumina HTS limitations. The MinION device is a pocked sized (10 x 3 x 2 cm, 90 g) sequencing platform that is equipped with nanopore flow cells. It can be run in virtually any setting from the polar region to tropical Africa with only limited technical requirements (Johnson et al., 2017; Pomerantz et al., 2018). When DNA molecules are guided through the flow cells, changes in the magnitude of the current in the nanopores code for the four different nucleotides (A, G, C and T). A major advantage over the high-performance Illumina read technology is Nanopore's read length profile that has a mean read length of 10 kb. Unfortunately, the technology is currently limited by its high sequencing

error rate (Ashton et al., 2015; Ip et al., 2015; Laver et al., 2015; Lu et al., 2016). Even with the rolling circle to concatemeric consensus (R2D2) method, the median accuracy does not exceed 94% (Volden et al., 2018). However, in unknown pathogen species (in the absence of suitable reference genomes) or genomes that are difficult to whole-genome sequence (homologous genes, long repetitive regions), the ONT can be used to create a draft genome that is subsequently used as reference for complementary Illumina reads, that have an expected final sequencing error rate of only ~1%. This way, previously ONT-introduced sequencing errors will be counterbalanced and the final genome assembly will have high accuracy. However, it can be expected that Nanopore technology further improves its read accuracy and that future studies can go without complementary Illumina sequencing for high-accuracy WGS. A current alternative to ONT is the single molecule, real-time (SMRT) sequencing technology that uses a DNA polymerase driven uninterrupted template-directed synthesis in the presence of distinguishable fluorescently labeled deoxyribonucleoside triphosphates (dNTPs) (Eid et al., 2009). According to Pacific Biosciences (PacBio), a company that has commercialized the technology, SMRT sequencing provides long read lengths (>50 kb) without compromising throughput or accuracy. SMRT is reported to achieve a median accuracy of 98.9% and performs well through AT- or GC-rich regions as well as repetitive sequences, long homopolymers, and palindromic sequences. While the pre-run cost of flow-cells and reagents of PacBio and ONT are comparable, the PacBio Sequal sequencer significantly exceeds the costs of the ONT. Nevertheless, the high accuracy of SMRT is currently unmatched by ONT, even when a modern method such as R2D2 is used (Volden et al., 2018).

5.2.4 DNA target enrichment

Despite the recent achievements in sequencing technologies, one aspect remains a great challenge for the WGS of uncultivable pathogens (e.g., *T. pallidum*) in complex clinical or environmental samples: low quantity of target DNA is often paired with high amounts of background DNA. To enhance the ratio of target DNA to background DNA, and thus to enhance HTS-performance there are target-unspecific and -specific enrichment methods available.

Base methylation as a DNA modification is found in all living organisms (Cheng, 1995). In prokaryotes, its main role is to protect the host DNA from enzymatic digest and to a lesser degree to regulate DNA replication (Jeltsch, 2002; Marinus and Løbner-Olesen, 2014). In eukaryotes, it has several functions, such as gene regulation and genomic imprinting (Jeltsch, 2002). C^5-methyl-cytosine (m5C) is the only known covalent modification of the DNA in eukaryotes (Jeltsch, 2002) but is, to a lesser extent, also present in prokaryotic DNA (reviewed in (Sánchez-Romero et al., 2015)). Bacterial genomes, also, contain N^4-methyl-cytosine (m4C) and N^6-methyl-adenine (m6A) (Hattman, 2005; Jeltsch, 2002; Ratel et al., 2006). Commercially available DNA enrichment methods for the unspecific enrichment of bacterial DNA make use of these differences in DNA methylation. The Looxter® Enrichment kit (Analytic Jena) uses the specific affinity of the CXXC finger protein 1 (*CFP1*) for non-methylated CpG-dinucleotides (Voo et al., 2000) and is thus able to bind bacterial and fungal DNA (Sachse et al., 2009). Other systems (e.g., NEBNext® Microbiome DNA Enrichment Kit) are designed to bind the CpG-methylated (m5C) DNA and therefore selectively bind the background DNA of the mammalian host in clinical samples (Feehery et al., 2013). Currently,

the most promising technique for the unselective enrichment of prokaryotic DNA in samples with a high background DNA of mammalian hosts is the *DpnI* method published by Barnes et al. (Barnes et al., 2014). The method is currently not commercially available and unlike the other methods that rely on m5C methylation at CpG sides, it uses a restriction endonuclease-mediated DNA enrichment approach. *DpnI* is a methyl-directed endonuclease that is highly specific for methylated adenine residues (Lacks and Greenberg, 1975; Vovis and Lacks, 1977). The enzyme activity has been suppressed through manipulation of the reaction conditions so that the enzyme binds, but not digests the DNA (Barnes et al., 2014).

Target enrichment via hybridization-based capture is the most specific and powerful tool available today to selectively enrich HTS sequencing libraries. Custom-adapted myBaits® target capture kits (Arbor Biosciences) are commercially available. After the submission of reference sequences, the manufacturer designs and produces RNA probes of 80-bp length that are complementary to the target sequence(s). During in-solution hybridization, target DNA will bind to the complementary RNA baits, which will subsequently be coupled with magnetic beats. Unbound DNA is washed out and bait captured DNA can be amplified and subsequently high-throughput sequenced. The myBaits® system was for example successfully used to whole-genome sequence *T. pallidum* strains of NHP origin (Chapter 7.3.4) (Knauf et al., 2018).

5.3 The Importance of Assay Validation and Standardization

The validation and standardization of assays used to diagnose and characterize pathogens increase reproducibility and therefore the ability of an assay to deliver trustful results. In a nutshell, validation is a process that determines the integrity and fitness of an assay (OiE, 2019). It guarantees that a test is working accurately, that it is (species) specific and reproducible. Few assays that are available for disease testing have been sufficiently validated for the use in wildlife (Gardner et al., 1996). However, knowledge of the analytic and diagnostic performance of assays used to investigate infectious diseases is of fundamental importance, in particular when the results are used for decision making and disease management. Due to the limit of detection (a measure of assay sensitivity), low copy numbers of a pathogen in a given sample can result in false-negative test results. Furthermore, degraded samples might interfere with the technical requirements of an assay. An example of this would be the lack of amplification in long-range PCR methods that use DNA extracted from faecal samples, which is generally highly degraded (Chiou and Bergey, 2018). In addition to the host species specificity, assay development and optimisation efforts must pay due attention to the sample materials that are used to detect a pathogen (Stallknecht, 2007). Figure 2 illustrates a flowchart that has been published by the World Organisation for Animal Health (OiE, 2019). It demonstrates the assay development and validation pathways that need to be considered when implementing a new diagnostic test. While the extent of assay validation may differ according to the intended use and availability of appropriate control samples, assay validation is a principal component of any wildlife disease investigation. Due to differences in selectivity, the extent to which a method can accurately quantify the pathogen in the presence of interferents or degradants (OiE, 2019), assays that work at the first stage (Figure 2 analytic characteristics) do not necessarily perform well in the second stage of the assay validation (Figure 2: diagnostic characteristics). In the latter, the assay is tested using clinical samples.

Generally, test results from weakly validated assays should be confirmed by at least one independent assay before they are reported.

The diagnostic sensitivity (proportion of samples from known infected animals that test positive) and diagnostic specificity (proportion of samples from known uninfected animals that test negative) of a test are the primary performance indicators established during the validation process (Office International des Epizooties, 2019). The reporting of these statistical parameters requires accuracy and is dependent on an adequate sample size of known positive and negative reference samples. In many cases, however, sample availability is limited in wildlife research and a lower confidence interval for the diagnostic sensitivity and specificity must be accepted but needs to be reported along with the study outcome.

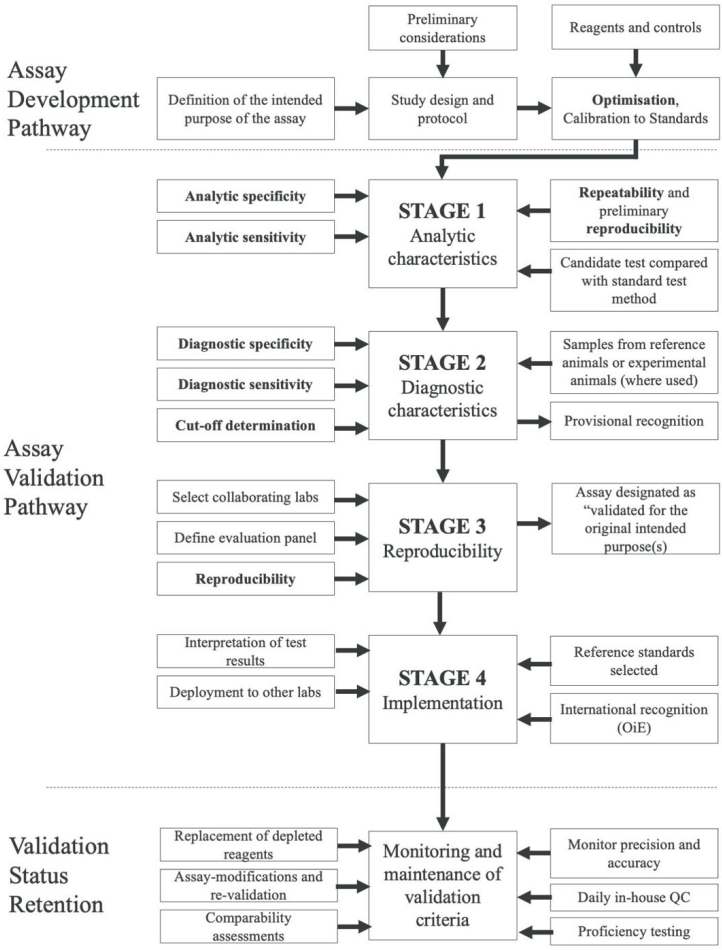


Figure 2. Proposed assay development and validation pathways with assay validation criteria highlighted in bold. The figure has been reproduced and modified from © OIE 2018 (World Organisation for Animal Health), reproduction of Fig. 1 given in chapter 1.1.2 Principles and methods of validation of diagnostic assays for infectious diseases of the Manual of Diagnostic Tests and Vaccines for Aquatic Animals, available at

https://www.oie.int/index.php?id=2439&L=0&htmfile=chapitre_validation_diagnostics_assays.htm.

The magnitude of technological advances during the last two decades, coupled with new and efficient diagnostic tests has revolutionized the molecular detection of pathogens. On the downside, the increase in assay sensitivity has magnified laboratory contamination problems. In particular microbiome-HTS approaches are prone to laboratory contamination, sometimes to an extent where microorganisms are introduced through ultrapure water systems (Laurence et al., 2014). Validation efforts, therefore, need to standardize the test performance across the different species and sample types. This includes all sample storage media and assay test components that are required to run the reaction (Brooks et al., 2015; Hallmaier-Wacker et al., 2018) (Chapter 5.3.4). An appropriate number of negative control samples needs to be included. It is of utmost importance to understand the technical limitations and (unavoidable) error rates associated with any given assays. Data-rich methods are further prone to interpretation errors (Tripathi et al., 2016) since the use of different algorithms and databases can affect the results (López-García et al., 2018; Xue et al., 2018). Another challenge to the standardization of microbial sequence data is the definition of sequence thresholds that enable the taxonomic classification of microbial sequence data based on sequence similarity (Cohan, 2002; Fraser et al., 2009; Rosselló-Mora and Amann, 2001; Staley, 2006). Yet, this kind of classification struggles with databases that do not cover the whole range of microorganisms (and will likely never do). The advanced field of metataxonomics, therefore, applies binning of sequences into e.g., OTUs (Blaxter et al., 2005; Schloss et al., 2009) or zero-radius OTUs (Edgar, 2018b; Xue et al., 2018). For 16S rRNA gene sequences, this provides more accuracy and makes data comparable across the different species without the need to have the orthologous sequence in a given database such as SILVA or GreenGenes (Schloss, 2016). However, there is an ongoing debate on the robustness of a generalized cut-off value and this is not restricted to the 16S rRNA data (Konstantinidis et al., 2006). The 97% threshold, which is commonly used in data that originate from 16S rRNA amplicon sequencing has been questioned, as it creates a bias in diversity metrics, which makes it difficult to compare results across different studies in particular when different sets of primers have been used (Mysara et al., 2017). To recognize the different evolutionary rates within the different variable regions of the 16S rRNA gene, Mysara et al. developed a methodology that calculates dynamic and evolutionary compliant OTU clustering (Mysara et al., 2017). For obvious reasons, it will be impossible to achieve consensus in every methodology that is used to diagnose and characterize pathogens. It is therefore in the interest of standardization to share raw research data along with the presentation of analysis outcomes (Piwowar, 2011).

The following publications, provide examples for the assay validation pathway (5.3.1 and 5.3.3) and the assay development pathway (5.3.2).

5.3.1 Validation of Serological Tests for the Detection of Antibodies Against *Treponema pallidum* in Nonhuman Primates

This study has been conducted to validate the usefulness of commercially available serological tests for the identification of anti-*T. pallidum* antibodies in baboons (*Papio* sp.). The outcome of the study generated confidence in reporting serological test results from naturally infected wild NHPs and paved the way for the use of the validated serological test to screen for naturally occurring infection of NHPs with *T. pallidum*.

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Validation of Serological Tests for the Detection of Antibodies Against *Treponema pallidum* in Nonhuman Primates

Sascha Knauf[1]*, Franziska Dahlmann[1], Emmanuel K. Batamuzi[2], Sieghard Frischmann[3], Hsi Liu[4]

1 German Primate Center, Pathology Unit, Work Group Neglected Tropical Diseases, Göttingen, Germany, **2** Sokoine University of Agriculture, Faculty of Veterinary Medicine, Department of Surgery and Theriogenology, Morogoro, Tanzania, **3** Mast Diagnostica GmbH, Reinfeld, Germany, **4** National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Diseases Control and Prevention, Atlanta, Georgia, United States of America

* sknauf@dpz.eu

OPEN ACCESS

Citation: Knauf S, Dahlmann F, Batamuzi EK, Frischmann S, Liu H (2015) Validation of Serological Tests for the Detection of Antibodies Against *Treponema pallidum* in Nonhuman Primates. PLoS Negl Trop Dis 9(3): e0003637. doi:10.1371/journal.pntd.0003637

Editor: Mathieu Picardeau, Institut Pasteur, FRANCE

Received: November 21, 2014

Accepted: February 22, 2015

Published: March 24, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: SK received funding and support for the field work at Lake Manyara National Park from German Academic Exchange Program (D/06/43974, https://www.daad.de/de/), Christian Vogel Fond (2006, http://www.gf-primatologie.de/), World Association Zoos and Aquariums (07002, http://www.waza.org/en/site/home), University of Leipzig, Justus Liebig University of Giessen, Scil animal care company, Telinject, Translogistic, and Wuppertal Zoo. The funders had no role in study design, data

Abstract

There is evidence to suggest that the yaws bacterium (*Treponema pallidum* ssp. *pertenue*) may exist in non-human primate populations residing in regions where yaws is endemic in humans. Especially in light of the fact that the World Health Organizaiton (WHO) recently launched its second yaws eradication campaign, there is a considerable need for reliable tools to identify treponemal infection in our closest relatives, African monkeys and great apes. It was hypothesized that commercially available serological tests detect simian anti-*T. pallidum* antibody in serum samples of baboons, with comparable sensitivity and specificity to their results on human sera. Test performances of five different treponemal tests (TTs) and two non-treponemal tests (NTTs) were evaluated using serum samples of 57 naturally *T. pallidum*-infected olive baboons (*Papio anubis*) from Lake Manyara National Park in Tanzania. The *T. pallidum* particle agglutination assay (TP-PA) was used as a gold standard for comparison. In addition, the overall infection status of the animals was used to further validate test performances. For most accurate results, only samples that originated from baboons of known infection status, as verified in a previous study by clinical inspection, PCR and immunohistochemistry, were included. All tests, TTs and NTTs, used in this study were able to reliably detect antibodies against *T. pallidum* in serum samples of infected baboons. The sensitivity of TTs ranged from 97.7-100%, while specificity was between 88.0-100.0%. The two NTTs detected anti-lipoidal antibodies in serum samples of infected baboons with a sensitivity of 83.3% whereas specificity was 100%. For screening purposes, the TT Espline TP provided the highest sensitivity and specificity and at the same time provided the most suitable format for use in the field. The enzyme immune assay Mastblot TP (IgG), however, could be considered as a confirmatory test.

collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: SF is employed by Mast Diagnostica GmbH, a company that sells treponemal tests (Serodia TP-PA, Espline TP, Mastafluor FTA-ABS IgG, and Mastablot TP IgG) that were used in this study. ALL.DIAG S.A.S (http://www.alldiag.com/) provided Syphilitop Optima tests. This does not alter our adherence to all PLOS policies on sharing data and materials.

Author Summary

The success of any disease eradication campaign depends on considering possible non-human reservoirs of the disease. Although the first report of *T. pallidum* infection in baboons was published in the 1970's and the zoonotic potential was demonstrated by inoculation of a West African simian strain into humans, nonhuman primates have not yet been considered as a possible reservoir for re-emerging yaws in Africa. Simian strains are genetically most closely related to the strains that cause yaws in humans. The identification of baboons as a reservoir for human infection in Africa would be revolutionary and aid important aspects to yaws eradication programs. Reliable serological tests and a useful standardized test algorithm for the screening of wild baboon populations are essential for studying potential transmission events between monkeys and humans.

Introduction

Treponema pallidum is the bacterium that causes venereal syphilis (ssp. *pallidum*) and the non-venereal diseases yaws (ssp. *pertenue*) and endemic syphilis (ssp. *endemicum*) in humans [1]. The spirochete is able to cause a life-long chronic infection in untreated individuals [2] and elicits a strong adaptive immune response against a wide array of antigens [3–4] with strong serum IgM and IgG response [5–7] towards a number of lipoproteins (e.g., Tp15, 17, and 47), endoflagellar proteins (e.g., FlaA, FlaB1, 2, and 3), and the Tpr family proteins (e.g., TprA-TprL) [6]. Furthermore, infection-related cellular damage is known to induce the production of non-treponemal antibodies mainly directed against cardiolipids [8,9].

Recently, we have reported that *T. pallidum* can infect large numbers of African monkeys and great apes [10]. To date, all simian isolates seem to be closely related to *T. pallidum* ssp. *pertenue*, the pathogen causing human yaws [11,12] and at least the Fribourg-Blanc simian strain, which was isolated from a baboon in Guinea [13], has the potential to cause sustainable infection in humans [14]. Thus, there is evidence to suggest that yaws exists in non-human primate populations residing in regions where humans are also infected [15]. The clinical manifestations in non-human primates (NHPs) however, show regional differences. While West African simian strains of *T. pallidum* mostly cause no clinical signs [16], gorillas in the Republic of the Congo show yaws-like lesions [17] and baboons in East Africa are known to develop severe genital ulceration [11,18]. However, independent of the clinical manifestations simian strains induce a pronounced serological response in the respective host [10], which may be used to screen and identify host populations for their potential as a natural reservoir.

In the context of the possible zoonotic potential of simian strains [14], the identification and knowledge of a nonhuman reservoir for *T. pallidum* is crucial to disease elimination or eradication efforts and could help to identify hot spots for potential simian-to-human disease transmission. There is therefore considerable need to validate treponemal tests (TTs) and non-treponemal (NTTs) for their use in NHPs. Due to the close relationship of simian and human treponemes [12], we hypothesized that A) commercially available serological tests are able to detect simian anti-*T. pallidum* IgM and IgG in serum samples of baboons, a NHP species with high infection rates and B) that the serological tests will be equally reliable in terms of sensitivity and specificity in baboon sera compared to the human sera.

Materials and Methods

Ethics statement

Baboon serum samples were taken in accordance with the Tanzania Wildlife Research Institute's Guidelines for Conducting Wildlife Research (2001) and with permission of Tanzania National Parks (TNP/HQ/E.20/08B) as well as Commission for Science and Technology in Tanzania (2007-56-NA-2006-176). The committee of Tanzania National Parks and Tanzania Wildlife Research Institute approved sample collection. Baboon serum samples from the German Primate Center were granted from the institute's bio bank and originated from healthy animals that were sampled during post-mortem examination. The Animal Welfare and Ethics Committee of the German Primate Center approved the use of samples for this study.

Study site and animals

In a previous study, we were able to detect *T. pallidum* infection in wild olive baboons (*Papio anubis*) at Lake Manyara National Park in Tanzania [18]. Although the isolated strain is most closely related to *T. pallidum* ssp. *pertenue* [11], the pathogen causes severe genital ulceration. Diagnosis was based on gross pathology, histology, and molecular biological tests. The latter included quantitative [19] and qualitative PCR [20], targeting the *polA* gene of *T. pallidum*. DNA was extracted from skin tissue samples [18]. Data and corresponding serum samples that were constantly stored at -80°C of 57 untreated baboons from this study were available for analysis in 2013. An additional set of 11 serum samples of healthy captive olive baboons from the German Primate Center were included as negative control. The extent of genital ulceration was used to classify and group animals as clinically healthy, initially-, moderately-, or severely-infected (Fig. 1). It is not known whether simian infection develops in stages similar to human infection.

Treponemal tests (TTs)

***Serodia TP-PA* (Fujirebio Diagnostics Inc., Malvern, PA, USA; Cat. No. 201326).** The test uses sensitized colored gelatin particles as carriers of the *T. pallidum* (Nichols Strain) antigen and is run in microtiter plate reaction wells (high-binding and U-shaped, Cat. No. 650061, Greiner bio-one, Frickenhausen, Germany). All steps and control standards followed the manufacturer's protocol. Serum was not heat pre-treated, but sera that showed agglutination with unsensitized and sensitized gelatin particles were re-tested with a pre-absorption step as recommended by the manufacturer. A quantitative analysis was carried out using an initial titration scheme (1:5 to 1:20,480). Serum samples showing high reaction titers of $\geq$ 1:20,480 were re-tested until an end-point with negative reaction was reached. Results were interpreted in accordance with the manufacturer's protocol. Two independent investigators read each test result on the day of testing and in addition 24 hours later. Performance characteristics as communicated by the manufacturer are summarized in Table 1. All baboon serum samples were tested at least twice. In this study and in accordance with the European Guidelines on the Management of Syphilis [21–23], *T. pallidum* particle agglutination assay (TP-PA) was used as the initial screening test and gold standard to be compared to all other test performances.

***Espline TP* (Fujirebio Diagnostics Inc., Malvern, PA, USA; Cat. No. 219126).** This screening test comes in a ready-to-use cassette format and is licensed for use in serum or plasma samples. It is based on immunochromatography for the detection of IgM and IgG class antibodies to *T. pallidum*. The recombinant antigens Tp47 (0.7 μg/cassette), Tp15-17 fusion protein (0.4 μg/cassette), alkaline phosphatase labeled TP recombinant antigen (Tp 47, 17.0 ng/cassette), and alkaline phosphatase labeled TP recombinant antigen (Tp15-17, 14.8 ng/

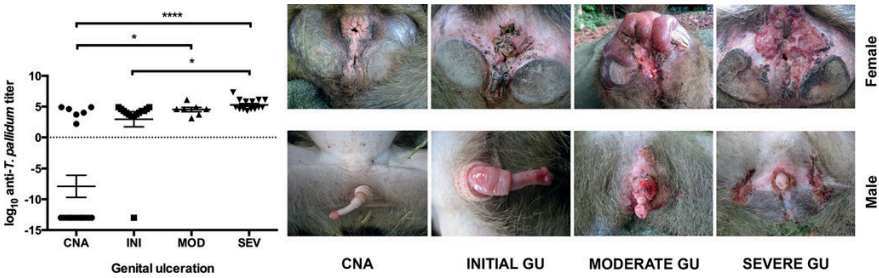


Fig 1. Multiple comparisons of log anti-*T. pallidum* titers in 4 groups with a different stage of genital ulceration in baboons ([CNA] = clinically non-affected (n = 20), [INI] = initial (n = 14), [MOD] = moderate (n = 7), and [SEV] = severe genital ulceration (n = 16); for stage definition see [18]), GU = genital ulceration. Anti-*T. pallidum* antibody quantification was investigated using the Serodia TP-PA. Kruskal-Wallis test using Dunn's correction for multiple comparison: CNA vs. SEV mean rank diff. = -30.04, $p \leq 0.0001$; CNA vs. MOD mean rank diff. = -19.95, $p \leq 0.05$; INI vs. SEV mean rank diff. = -17.56, $p \leq 0.05$. (mean ± SEM).

doi:10.1371/journal.pntd.0003637.g001

cassette) were coated on the reaction membrane. The assay was operated and interpreted according to the manufacturer's instructions. Test results were considered valid when the internal reference line was present. All baboon serum samples were tested at least twice and were not heat pre-treated.

***Syphilitop Optima* (ALL. DIAG S.A.S., Strasbourg, France, Cat. No. 5480).** According to the manufacturer's protocol, this immunochromatographic test must only be used on serum. Its membrane is partially coated with a mixture of different *T. pallidum*-specific antigens (17 kDa and 47kDa). Serum migrates along the test strip through chromatography. IgM and IgG class anti-*T. pallidum* antibodies bind to the antigen(s), indicated by the appearance of purple-colored indicator bands. Only tests that showed an internal reference band were considered for valid test interpretation. All baboon serum samples were tested at least twice and were not heat pre-treated.

***Mastafluor FTA-ABS* IgM and IgG (Mast Diagnostica, Reinfeld, Germany; Cat. No. 6305222).** The Fluorescence-*T. pallidum*-Antibody-Absorption-Test (FTA-ABS) kit for IgM and IgG was used to test serum samples of baboons. Samples were not heat pre-treated. Depending on the conjugate antibody, it was possible to distinguish between IgM or IgG class antibodies. The test procedure followed the manufacturer's protocol.

***Mastablot TP* (Mast Diagnostica, Reinfeld, Germany).** The immunoblots Mastablot TP IgM (Cat. No. 6653M24) and IgG (Cat. No. 6653G08) were used to detect anti-*T. pallidum* antibodies in all serum samples of baboons. Both blots use nitrocellulose strips coated with p15,

Table 1. Performance characteristics of the serological tests used in this study, as reported by the manufacturer. Sen = Sensitivity, Spec = Specificity, n.p. = not provided.

	Serodia TP-PA (n)	Espline TP (n)	Syphilitop Optima (n)	Mastafluor FTA-ABS IgG (n)	Mastablot TP IgG (n)	VDRLCHECK CHARBON/ RPR (n)	RPR-100 (n)
Sen	100.0% (145)	100.0% (145)	100.0% (103)	n.p.	n.p.	100.0% (50)	n.p.
Spec	100.0% (935)	99.7% (932)	95.0% (100)	n.p.	n.p.	100.0% (100)	n.p.

doi:10.1371/journal.pntd.0003637.t001

p17, TmpA and p47. Each immunoblot was used according to the manufacturer's guidance. As in all tests, serum was used without heat pre-treatment.

Non-treponemal tests (NTTs)

***VDRLCHECK CHARBON/RPR* (ALL. DIAG S.A.S., Strasbourg, France; Cat. No. 5474).** This macroscopic non-treponemal flocculation test for the detection, and to a certain extent quantification of anti-lipoidal antibodies was used in all serum samples of baboons. Tests were repeated twice and serum samples were not heat pre-treated. Serum was mixed with the VDRLCHECK CHARBON/RPR reagent and allowed to react for eight minutes. In case anti-lipoidal antibodies were present, black macroscopic visible floccules were visible. The test procedure and test interpretation was operated according to the manufacturer's protocol. To avoid false-negative test results due to the prozone phenomenon [24], critical test results were retested with 1:8 dilution.

***RPR-100* (Biorad, Marnes, France; Cat. No. 72505).** Similar to the VDRLCHECK CHARBON/RPR Test, the Rapid Plasma Reagin (RPR) test kit was used in serum samples of baboons to detect qualitative and semi-quantitative IgM and IgG antibodies against lipoidal material that originates from host cell damage or lipoprotein-like particles of the spirochete. Available samples were tested twice. The test procedure and interpretation of results were based on the manufacturer's protocol. The presence of antibodies resulted in macroscopic visible agglutination of carbon-particles coated with a mix of liquid antigens. Carbon particles are dispersed in a medium that contains not otherwise specified substances to eliminate unspecific reaction. To avoid false-negative test results due to the prozone phenomenon [24], critical test results were retested with 1:8 dilution.

At least one researcher and an experienced technician read the test results of each test. Results of each test were blinded and were communicated independently. If not otherwise specified, each test was performed on a single day. Performance characteristics of all tests, as reported by the manufacturer, are summarized in Table 1.

Definition of test results

Generally, the definition of a test result was based on the individual's overall infection status (Table 2). Details of infection status including genital ulceration status and each test's interpretation can be found in the Supporting Information (S1 Table).

True positive (Tpos). For TTs, a true positive test result was assumed, when the result coincided with the majority of all other serological test results of TTs, which were positive (3 out of 5, excluding tests to detect anti-*T. pallidum* IgM), a positive outcome of the skin tissue PCR, and/or immunohistochemistry (IHC). NTTs were considered Tpos when molecular biological tests indicated the presence of *T. pallidum* and/or when the test result correlated to the overall outcome of the TT results.

False positive (Fpos). A false positive test result of TTs was presumed when the outcome of a test was contrary to all other test results of the TTs, which were negative (3 out of 5, excluding tests to detect anti-*T. pallidum* IgM), as well as a negative skin tissue PCR and IHC result. NTTs were considered Fpos when molecular identification and the second NTT were negative.

True negative (Tneg). A TT result was defined as true negative when it conformed to the majority of all other TT results (3 out of 5, excluding tests to detect anti-*T. pallidum* IgM) along with a negative outcome of the skin tissue PCR and IHC. NTTs were defined Tneg, when both NTTs were non-reactive, while the overall TT result could be positive or negative.

False negative (Fneg). False negative TT results were documented when a result was contrary to all other test results of the serological tests, which were positive (3 out of 5, excluding

Table 2. Definitions used to determine the infectious stage of baboons.

Test		Results				
Genital Ulceration	/	+/-	+/-	-	+/-	-
IHC	/	+ (1 of 2)	-	-	-	-
PCR						
Treponemal Tests	/	+ (1 of 2)	-	-	-	-
Serodia TP-PA	3 of 5 tests	+/-	+	-	+	-
Espline TP						
Syphilitop Optima	3 of 5 tests	+/-	+	-	+	-
Mastafluor FTA-ABS IgG	3 of 5 tests	+/-	+	-	+	-
Mastablot TP IgG	3 of 5 tests	+/-	+	-	+	-
Non-Treponemal Tests	3 of 5 tests	+/-	+	-	+	-
VDRLCHECK CHARBON/RPR	/	+/-	+	+	-	-
RPR-100		+/-	+	+	-	-
Infection Status	/	**+**	**+**	**-**	**+**	**-**

The definition of infection status was based on the outcome of serological (this study) and molecular biological tests [18, S1 Table].

doi:10.1371/journal.pntd.0003637.t002

tests to detect anti-*T. pallidum* IgM), along with a positive skin tissue PCR and/or IHC result. NTTs were considered Fneg, when molecular biological tests indicated the presence of the spirochete and one of the second NTTs, VDRL or RPR, became reactive.

Statistics

Statistical analyses were performed using Prism 6.0 (GraphPad Software). Results of the TTs were first compared to the result of the Serodia TP-PA and second to the consensus of infection status, as it is defined in Table 2, and which takes into account the appearance of clinical symptoms (genital health status), IHC and skin tissue PCR results of the same animals as published elsewhere [18]. While the TP-PA was used as the gold standard for TTs, we compared results to the baboon's infection status for further verification of test results and accuracy. With regard to NTTs it was assumed that a significant proportion of tests might become nonreactive in chronically infected baboons, as it is described for untreated human syphilis infection [25–27]. Test performances of the NTTs were therefore evaluated exclusively on the basis of the consensus of infection.

A non-parametric test for nominal scale data, the two-tailed Fisher's Exact Test, was used to compare the proportions among the serological tests, skin tissue PCR results and clinical signs of infection as well as for the analysis of sensitivity and specificity of the serological tests.

Sensitivity. Proportion of actual positives that are correctly identified as true positives: Tpos/(Tpos + Fneg).

Specificity. Proportion of negative test results that are correctly identified as being negative: Tneg/(Tneg + Fpos).

Efficiency. Proportion of true test results within the overall sample size for the test: (Tpos + Tneg)/n total.

Positive and negative predictive value. Probability of a positive/negative test result to be true positive/negative: Tpos/(Tpos + Fpos) and Tneg/(Tneg + Fpos).

Endpoint titers of exponential scale were $\log_{10}$-transformed to reduce variance. In case of non-Gaussian distribution and $\log_{10}$-transformation, zero-titers were converted into 10e-14. Normal distribution was tested using the D'Agostino & Pearson omnibus normality test and

the Shapiro-Wilk normality test. Antibody titers were analyzed using four different baboon categories: clinically non-affected, initial, moderate, and severe stage of the genital ulceration as it was documented in the field [18]. The non-parametric ANOVA using Krustal-Wallis test was applied to the $\log_{10}$-transformed data sets of the antibody titers. Each mean rank of a genital ulceration stage (clinically non-affected [CNA], initial [INI], moderate [MOD], and severe genital-ulcerated [SEV]) was compared to mean rank of every other genital ulceration stage. Dunn's correction for multiple comparisons (significance without confidence intervals) was applied to the test. In all tests, $p \leq 0.05$ was considered statistically significant.

Results

Five treponemal and two non-treponemal tests were evaluated

Tables 3 and 4 summarize sample size, proportions, and performance characteristics of the serological tests that were used in the 57 baboon serum samples from Lake Manyara National Park and an additional set of 11 serum samples from olive baboons of the German Primate Center in Germany. The latter were included for the purpose of additional negative control.

Treponemal and non-treponemal tests detect antibodies against *T. pallidum* in serum samples of infected baboons

When comparing TT performances with the TP-PA, we observed test sensitivity in the range of 91.3–100%, and specificity ranging from 94.7–96.0% (Table 3). When test results were compared to the consensus of all test results including PCR, however, the observed sensitivity of the TTs was in the range of 97.7–100%; whereas the specificity reduced slightly to the range of 88.0–100% (Table 4). This reduction was almost exclusively caused by the test performance of Syphilitop Optima (Table 4). NTT performances were not compared to TP-PA results since positive TT results in untreated, chronically infected patients do not necessarily predict reactivity of the corresponding NTT [25–27]. However, both NTTs used in this study reliably detected anti-lipoidal antibodies in serum (Table 4) or plasma samples (S2 Table) of baboons. When infection status was considered in the context of all test results including PCR analysis, NTT

Table 3. Comparison of treponemal serological tests with the results of the Serodia TP-PA.

Assay and Result	Serodia TP-PA		p-Value	% Sensitivity (95% CI)	% Specificity (95% CI)	% Pos. pred. value (95% CI)	% Neg. pred. value (95% CI)
	Pos.	Neg.					
Treponemal Tests							
Espline TP							
Positive	42	1	< 0.0001	97.7 (0.877–0.999)	96.0 (0.797–0.999)	97.7 (0.877–0.999)	96.0 (0.797–0.999)
Negative	1	24					
Syphilotop Optima							
Positive	42	4	< 0.0001	91.3 (0.792–0.976)	95.5 (0.772–0.999)	97.7 (0.877–0.999)	84.0 (0.639–0.955)
Negative	1	21					
Mastafluor FTA-ABS IgG							
Positive	42	1	< 0.0001	97.7 (0.877–0.999)	96.0 (0.796–0.999)	97.7 (0.877–0.999)	96.0 (0.796–0.999)
Negative	1	24					
Mastablot TP IgG							
Positive	41	0	< 0.0001	100.0 (0.914–1.000)	94.7 (0.740–0.999)	97.6 (0.874–0.999)	100.0 (0.815–1.000)
Negative	1	18					

Two-tailed Fisher's exact test. Pos. pred. value = positive predictive value, Neg. pred. value = negative predictive value.

doi:10.1371/journal.pntd.0003637.t003

Table 4. Comparison of the serological tests with the consensus of infection status (Table 2).

Assay and Result	Serodia TP-PA	Espline TP	Syphilitop Optima	Mastafluor FTA-ABS IgG	Mastablot TP IgG	VDRL	RPR-100
n Tpos	42	43	43	43	41	10	10
n Fpos	2	0	3	0	0	0	0
n Tneg	23	25	22	25	19	27	27
n Fneg	1	0	0	0	0	2	2
n Total	68	68	68	68	60	39	39
% Sensitivity (95% CI)	97.7 (0.877–0.999)	100.0 (0.918–1.000)	100.0 (0.918–1.000)	100.0 (0.918–1.000)	100.0 (0.914–1.000)	83.3 (0.516–0.979)	83.3 (0.516–0.979)
% Specificity (95% CI)	92.0 (0.740–0.990)	100.0 (0.863–1.000)	88.0 (0.688–0.975)	100.0 (0.863–1.000)	100.0 (0.824–1.000)	100.0 (0.872–1.000)	100.0 (0.872–1.000)
% Pos. pred. value (95% CI)	95.5 (0.845–0.994)	100.0 (0.918–1.000)	93.5 (0.821–0.986)	100.0 (0.918–1.000)	100.0 (0.914–1.000)	100.0 (0.692–1.000)	100.0 (0.692–1.000)
% Neg. pred. value (95% CI)	95.8 (0.789–0.999)	100.0 (0.863–1.000)	100.0 (0.846–1.000)	100.0 (0.863–1.000)	100.0 (0.824–1.000)	93.1 (0.772–0.992)	93.1 (0.772–0.992)
% Efficiency	95.6	100.0	95.6	100.0	100.0	94.9	94.9

Two-tailed Fisher's exact test in all serology tests, $p < 0.0001$. Pos. pred. value = positive predictive value, Neg. pred. value = negative predictive value. Sample size differs because in some cases sample material was on short supply for further tests. For the NTTs only serum was tested and included here. Tests in all other animals were preformed with plasma and can be found in S2 Table.

doi:10.1371/journal.pntd.0003637.t004

sensitivity in serum samples was lower (83.3%) than the average of the TTs (99.54%). The specificity of VDRL and RPR in serum samples was 100% and therefore higher than the performance of the standard TP-PA (92.0%) and Syphilitop-Optima rapid test (88.0%). The performance data of NTTs for plasma samples are listed in S2 Table.

Anti-*T. pallidum* antibodies were found in 97.3% of baboons with genital ulceration and in 6 of 20 animals that were clinically healthy (30.0%, Table 5, Fig. 1). For comparison, a remarkable proportion of genital-ulcerated baboons (13.5%; n = 5/37, Table 5) had a negative PCR outcome of their respective skin tissue biopsy. Vice versa, we found 30% (n = 6/20) genital non-genital-ulcerated baboons with positive *T. pallidum* PCR of their corresponding skin biopsy. Yet, genital-ulcerated baboons were significantly more often reactive for *T. pallidum* in serology (TTs, $p < 0.0001$) and skin tissue PCR ($p < 0.0001$) compared to clinically healthy and thus non-ulcerated baboons. Both treponemal rapid tests, Espline TP and Syphilitop Optima, were more sensitive than the Serodia TP-PA (Two-tailed Fisher's exact test, $p < 0.0001$, Table 4), although specificity in the Syphilitop Optima was much lower than Serodia TP-PA. The same applied, when Mastafluor FTA-ABS IgG and Mastablot TP IgG were compared to the Serodia TP-PA particle agglutination assay. In both tests the proportion of positive results matched the results of the Serodia TP-PA (Two-tailed Fisher's exact test, $p < 0.0001$, Table 4).

Anti-*T. pallidum* IgM antibodies were not detected

No correlation was found when anti-*T. pallidum* IgG positive serum samples were tested in the immunoblot assay Mastablot TP for the presence of IgM antibodies against *T. pallidum*. Only a limited number of animals, 6 out of 59 (10.2%), tested positive for both anti-*T. pallidum* IgG and IgM antibodies. We did not find any samples that were positive for anti-*T. pallidum* IgM only.

Severe genital ulceration is associated with high antibody titers

Even after $\log_{10}$-transformation of antibody titers, the clinically healthy and the initial stage genital-ulcerated group (Fig. 1; it is not known whether NHPs develop three stages similar to humans; initial stage refers to the severity of genital ulceration as describe elsewhere [18]) were not normally distributed. The Kruskal-Wallis test using Dunn's correction for multiple

Table 5. Crosstab of the results obtained from 57 baboon and 11 control samples.

Genital Ulceration	CNA		Initial		Moderate		Severe		Negative Control	
n	20		14		7		16		11	
	+	-	+	-	+	-	+	-	+	-
PCR	6	14	10	4	7	0	15	1	n/t	n/t
IHC	2	18	3	11	4	3	7	9	n/t	n/t
Serodia TP-PA	8	12	13	1	7	0	16	0	0	11
Espline TP	6	14	14	0	7	0	16	0	0	11
Syphilitop Optima	9	11	14	0	7	0	16	0	0	11
Mastafluor FTA-ABS IgG	6	14	14	0	7	0	16	0	0	11
Mastablot TP IgG	6	8	12	0	7	0	16	0	0	11
VDRL (Serum)	2	7	1	7	2	3	5	1	0	11
RPR (Serum)	2	7	2	6	2	3	4	2	0	11

CNA = clinically non-affected.

doi:10.1371/journal.pntd.0003637.t005

comparisons showed that antibody titers in the severe genital-ulcerated group of baboons were significantly higher when compared to clinically healthy animals (mean rank diff. = -30.04, $p \leq 0.0001$) and baboons with an initial stage of genital ulceration (mean rank diff. = -17.56, $p \leq 0.05$). The same was found for moderate genital-ulcerated baboons, which had significantly higher antibody titers against *T. pallidum* than animals without genital ulceration (mean rank diff. = -19.95, $p \leq 0.05$). Fig. 1 provides an overview.

Discussion

All serological tests used in this study, TTs as well as NTTs, were able to detect anti-*T. pallidum* antibodies in serum of infected baboons. While the presence of anti-*T. pallidum* IgG antibodies in all infected animals correlates to lifelong antibody titers in human infection [28], the absence of IgM type antibodies against the spirochete in infected baboons (only 6 animals were positive for anti-*T. pallidum* IgM antibodies, Mastablot IgM, S1 Table) could most likely be linked to the timing of sampling from two weeks to months post infection. This may not be fully consistent with human cases, in which IgM antibody titers are low to moderately high in primary syphilis, peak in secondary and sometimes tertiary syphilis, and are low in latent syphilis [29]; the majority of baboons, which were tested in this study were unlikely in a latent stage of infection. Although animals are known to be chronically infected and untreated over many months to years, genital ulceration in infected baboons rarely heals up. However, the expected lower amplitude of IgM titers [30] and the expected reduced half-life of IgM [31] may have contributed to the result. Results were excluded from multiple comparisons analysis and evaluation of performance characteristics, because of the uncertainty and low number of IgM positive samples. The finding of few chronically infected individuals positive for IgM antibodies against *T. pallidum* could be explained by persisting IgM titers, as described for other human spirochete infections (e.g., borreliosis) [32,33].

All TTs and NTTs used in this study are commercially available and licensed for use with human serum or plasma samples. Using TT as a screening test in NHPs (Fig. 2) is in accordance with international standards and EU Guidelines for the Management of Syphilis in humans [22,34,35] and is often referred to as "reverse testing algorithm" [36]. Although there were no data about efficiency, sensitivity, and specificity available from testing samples in

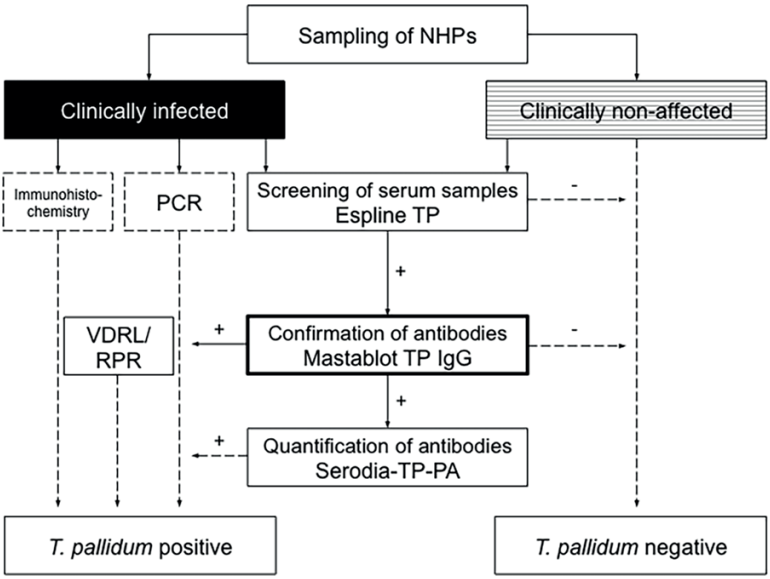


Fig 2. *T. pallidum* test algorithm for the screening of wild non-treated baboons. Based on the test performances Espline TP is recommended as the initial screening test followed by a confirmatory test e.g. Mastablot TP IgG that has been identified as most reliable standard. Dashed lines indicate reported results, while continuous lines represent the workflow.

doi:10.1371/journal.pntd.0003637.g002

baboons or other NHPs, it was reasonable to assume that a TP-PA would be a reliable standard in testing baboon serum samples both qualitatively and quantitatively (Table 3). A number of alternative TTs (Espline TP, Syphilitop Optima, Mastafluor FTA-ABS IgG, and Mastablot TP IgG) were included in the analysis to rule out uncertainties of TP-PA test performance in baboon sera. Mastafluor FTA-ABS IgG and Mastablot TP IgG were specifically added to cover different *T. pallidum* antigens than the rapid TTs Espline TP and Syphilitop Optima (TmpA; see Material and Methods). Interestingly, sensitivity, specificity, and the corresponding positive and negative predictive values were lower in Serodia TP-PA when compared to all other TTs, excluding the Syphilitop Optima, which had a higher sensitivity, but weaker specificity (Table 4). This was in contrary to the TP-PA manufacturer's information about performance characteristics in human test sera (Table 1) and also when compared to FTA-ABS and immunoblot IgG results of human samples [21, 37]. In our experience the interpretation of the gelatin TP-PA requires a certain level of training and serum pre-absorption in baboon samples. This was achieved by incubating the test serum with non-sensitized particles so that unspecific binding factors were pre-absorbed. Another difficulty of the test was related to the endpoints, which can differ over time. Despite those difficulties, Serodia TP-PA has an advantage in that the readability was made by the naked eye and may be operated in resource poor laboratory settings. Also, it was the only TT that can be used for semi quantitative titer quantification. Validation of dried blood spots with a Serodia TP-PA assay for external quality assurance of *T.*

pallidum serology as published elsewhere [38], provides an interesting outlook for the confirmation of screening test results, e.g., from the use of Espline TP in remote areas at the wildlife-human interface.

Both rapid TTs that could be used for screening, the Espline TP and Syphilitop Optima, were easy to perform. Results are reported within 15 min and especially the Espline TP comes in a handy cassette format. Both tests require only 25 μl of sample material. According to the manufacture's description serum and plasma samples can be used for Espline TP, while Syphilitop Optima must not use plasma samples. Comparing Espline TP to the Serodia TP-PA in serum samples of baboons, Espline TP had higher sensitivity (97.7 vs. 91.3%) and nearly the same specificity (96.0 vs 95.5%; Table 3). When the test performances are compared to the consensus of infection status (Table 2), sensitivity is 100% in both screening tests, but Syphilitop Optima achieved only 88.0% specificity compared to the Espline TP with 100% (Table 4). The reason for these differences remains uncertain but might be promoted by protein variations of the coated filter membrane. Espline TP uses at least one additional antigen (Tp 15) that is not included into Syphilitop Optima. Also, proteins (Tp 47 and Tp15-17) are used to coat two different areas of the membrane in the Espline TP, but they are combined in one field on the reaction membrane of the Syphilitop Optima. Unfortunately, there was no information available about the quantity of protein coated to the membrane in Syphilitop Optima. The Espline TP uses a combination of alkaline and non-phosphatase labeled TP recombinant antigens. Due to the lack of information, it was not possible to compare the tests in that aspect, but differences may have an influence on the binding affinity of simian *T. pallidum* antibodies.

When compared to the consensus of infection status (Table 4) both Mastafluor FTA-ABS IgG and Mastablot TP IgG have higher sensitivity and specificity and thus can be recommended as a confirmatory test in baboons. However, it may be noted, that in human infection FTA tests are no longer recommended for the diagnosis of syphilis [22,39], which should make Mastablot TP IgG the preferred option.

In the context of the performance characteristics that are reported by the manufacturers (Table 1), Serodia TP-PA had a slightly weaker sensitivity (97.7 vs. 100%) and reduced specificity (92.0 vs. 100%, Table 4). The Espline TP rapid test had nearly identical sensitivity and specificity as indicated by the manufacturer, whereas Syphilitop Optima had an equal sensitivity (100%) with a reduced specificity value, 88% vs. 95% as communicated by the manufacturer. Future studies may benefit from heat pre-treatment of serum samples to reduce interference caused by complement proteins and unspecific binding of antibodies. Heat pre-treatment was not part of any manufacturer's protocol of the tests that were used in this study.

Although the quantification of antibodies may not be of interest for disease prevalence studies in wild baboons, it may be an important tool for characterizing simian infection. The finding that severe genital-ulcerated baboons had significant higher anti-*T. pallidum* antibody titers than clinically non-affected animals (CNA vs. SEV, $p < 0.0001$; Fig. 1) or those with less severe genital ulceration (CNA vs. MOD and INI vs. SEV, both $p \leq 0.05$; Fig. 1) is consistent with what can be expected from the course of infection. The rating of chronicity of infection in baboons at Lake Manyara National Park was based on gross-pathology and histological examination of skin tissue samples [18].

Treponemal tests as a screening tool for non-human primate infection

The use of a NTT for the initial screening in the traditional algorithm in human infection is to avoid the detection of previously treated and non-active cases [40]. NTTs are known to produce a higher percentage of false positives [41] and test performance data of the NTTs as they are reported in Table 4 need to be interpreted with caution since it is neither known when an

individual was infected nor how long anti-cardiolipid antibodies can be found in the due course of infection in wild baboons.

The decision to use and recommend a TT as a screening test for *T. pallidum* infection in NHPs was based on the following three reasons and is in accordance with the current European Guidelines on the Management of Syphilis [23]. First, wild baboons are rarely treated and once infected, treponemal clearance may be an exception rather than the norm. Second, there is a paucity of data on cross-reactivity of proteins derived from human *T. pallidum* strains with antibodies against the simian strain in baboons. Lastly since the majority of baboons were chronically infected, we had reason to belief that a number of these chronically infected baboons were non-reactive in NTTs, as it was described in untreated human syphilis infection [25–27]. However, while a lifelong anti-*T. pallidum* antibody titer in baboons provides a most useful readout for the identification of a disease hot spot that offers the possibility for simian-human infection, therapeutic interventions in wild NHPs, as it is already conducted in baboons at Gombe Stream National Park in Tanzania (Collins et al. pers. communication) may benefit from the use of the traditional algorithm since NTTs may allow the differentiation of active and inactive infection.

It is generally believed that yaws has no animal reservoir. Until identical *T. pallidum* strains are found circulating in nonhuman primates and humans in their natural environment this understanding cannot change. Yet, to this end, more research is needed before nonhuman primates can definitely be ruled out to serve as a natural reservoir for human infection. We have only recently begun to explore the range of nonhuman primate infection in Africa. Because the human-livestock-wildlife interface is constantly growing, the potential for inter-species transmission increase significantly. It is also possible that simian strains do naturally infect humans but do not cause clinical manifestations, as it is the case in Guinea baboons (*Papio papio*) in Senegal; or it may be that at least the East African simian strains cause genital ulceration in humans and may therefore not be diagnosed as yaws based on their genetics. Clearly, further research is needed before any answers can be given and serological surveys are an important tool to support these investigations and to complete our picture of *T. pallidum* infection in humans and nonhuman primates.

Based on the outcome of this study we propose an algorithm for the screening of wild **non-treated** NHP populations (Fig. 2). The algorithm aims to identify *T. pallidum* infection in wild baboons and other NHPs and may complement the current yaws eradication campaign [42].

Conclusion

All tests used in this study provided reliable results to detect anti-*T. pallidum* antibodies in serum of baboons. We therefore favor hypothesis A, which suggests that commercially available serological tests are able to detect simian anti-*T. pallidum* IgM and IgG in serum samples of baboons, with the exception of IgM class anti-*T. pallidum* antibodies. It would be necessary to examine more animals in the initial stage of infection in order to test this part of the hypothesis, something that is difficult to achieve since the time of infection in wild baboons in general is not known. While NTTs may help to plan treatment and control of infections in baboons, TTs are most useful to screening non-treated baboon population for the presence of *T. pallidum*. Hypothesis B was partly rejected because some serological tests were not equally reliable in their sensitivity and specificity in baboon samples compared to human serum samples.

For screening purposes, the immunochromatography based Espline TP test provided the highest sensitivity and specificity values and in addition had the most suitable format for use in the field. For confirmation, the treponemal test Mastablot TP IgG had the best performance characteristics and is therefore recommended as a gold standard. Serodia TP-PA was able to quantify antibodies against *T. pallidum* in baboons and results were consistent with the

chronicity of infection. Based on this study a testing algorithm for the screening of NHP populations for *T. pallidum* infection is proposed, which may help future yaws eradication campaigns or wildlife management to identify baboons as a potential reservoir for human yaws infection.

Supporting Information

S1 Table. Individual-based dataset on clinical manifestations, molecular and serological test results.
(XLSX)

S2 Table. Comparison of NTT results from serum and plasma samples of baboons.
(XLSX)

S1 Checklist. STARD checklist for reporting of studies of diagnostic accuracy *(Version January 2003)*.
(DOC)

Acknowledgments

We thank Tanzania Wildlife Research Institute, especially J. Keyyu; Tanzania National Parks, especially I. A. V. Lejora; Lake Manyara National Park (LMNP) headquarter staff as well as the Tanzania Commission for Science and Technology for their support to sample baboons at LMNP. For assistance, we thank E. Kiwele, K. Mollel, D. M. Kambarage, R. A. Muhala, and P. Wambura. A. Wehrend supervised the PhD study of SK from which sample material originates. S. Wienstroth is thanked for technical support in the laboratory. K. Mätz-Rensing provided negative control samples from the German Primate Center. ALL. DIAG S.A.S., T. Lamy is thanked to support the study with one free kit of Syphilitop Optima tests.

The reviewers are thanked for their comments and suggestions, which improved the manuscript.

Author Contributions

Conceived and designed the experiments: SK HL. Performed the experiments: SK FD SF. Analyzed the data: SK SF HL. Contributed reagents/materials/analysis tools: SK SF. Wrote the paper: SK FD EKB SF HL. Field immobilization of baboons: SK EKB.

References

1. Kapembwa M. Endemic treponematoses. In: Cook GC, Zumla AI, editors. Manson's Tropical Diseases. Philadelphia: Saunders-Elsevier; 2009. pp. 1139–1148.
2. Ho EL, Lukehart SA. Syphilis: using modern approaches to understand an old disease. J Clin Invest. 2011; 121: 4584–4592. doi: 10.1172/JCI57173 PMID: 22133883
3. Brinkman MB, McGill MA, Pettersson J, Rogers A, Matejkova P, Smajs D, et al. A novel *Treponema pallidum* antigen, TP0136, is an outer membrane protein that binds human fibronectin. Infect Immun. 2008; 76: 1848–1857. doi: 10.1128/IAI.01424-07 PMID: 18332212
4. McGill MA, Edmondson DG, Carroll JA, Cook RG, Orkiszewski RS, Norris SJ. Characterization and serologic analysis of the Treponema pallidum proteome. Infect Immun. 2010; 78: 2631–2643. doi: 10.1128/IAI.00173-10 PMID: 20385758
5. Carlson JA, Dabiri G, Cribier B, Sell S. The immunopathobiology of syphilis: the manifestations and course of syphilis are determined by the level of delayed-type hypersensitivity. Am J Dermatopathol. 2011; 33: 433–460. doi: 10.1097/DAD.0b013e3181e8b587 PMID: 21694502
6. Lafond R, Lukehart S. Biological Basis for Syphilis. Clin Microbiol Rev. 2006; 19: 29–49. PMID: 16418521

7. Sun R, Lai DH, Ren RX, Lian S, Zhang HP. *Treponema pallidum*-specific antibody expression for the diagnosis of different stages of syphilis. Chin Med J. 2013; 126: 206–210. PMID: 23324264
8. Harris A, Rosenberg AA, Riedel LM. A microflocculation test for syphilis using cardiolipin antigen. J Vener Dis Inf. 1946; 27: 169–174. PMID: 21003298
9. Portnoy J, Brewer JH, Harris A. Rapid plasma reagin card test for syphilis and other treponematoses. Public Health Rep. 1962; 77: 645–652. PMID: 14488129
10. Harper K, Knauf S.*Treponema pallidum* infection in primates: clinical manifestations, epidemiology, and evolution of a stealthy pathogen. In: Brinkworth JF, Pechenkina K, editors. Primates, pathogens, and evolution. New York: Springer; 2013. pp. 189–219.
11. Harper KN, Fyumagwa RD, Hoare R, Wambura PN, Coppenhaver DH, Sapolsky RM, et al. *Treponema pallidum* Infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS ONE. 2012; 7: e50882. doi: 10.1371/journal.pone.0050882 PMID: 23284649
12. Zobaníková M, Strouhal M, Mikalova L, Cejkova D, Ambrozova L, Pospisilova P, et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl Trop Dis. 2013; 7: e2172. doi: 10.1371/journal.pntd.0002172 PMID: 23638193
13. Fribourg-Blanc A, Mollaret HH, Niel G. Serologic and microscopic confirmation of treponemosis in Guinea baboons. Bull Soc Pathol Exot Filiales. 1966; 59: 54–59. PMID: 5333741
14. Smith JL, David NJ, Indgin S, Israel CW, Levine BM, Justice J, et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br J Vener Dis. 1971; 47: 226–251. PMID: 4936861
15. Knauf S, Liu H, Harper KN. *Treponemal* infection in nonhuman primates may serve as a potential reservoir for human yaws. Emerg Infect Dis. 2013; 19: 2058–2060. doi: 10.3201/eid1912.130863 PMID: 24274094
16. Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. Primates Med. 1969; 3: 113–121. PMID: 5006024
17. Levrero F, Gatti S, Gautier-Hion A, Menard N. Yaws disease in a wild gorilla population and its impact on the reproductive status of males. Am J Phys Anthropol. 2007; 132: 568–575. PMID: 17274014
18. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IA, Nordhoff M, et al. *Treponema* infection associated with genital ulceration in wild baboons. Vet Pathol. 2012; 49: 292–303. doi: 10.1177/0300985811402839 PMID: 21411621
19. Leslie DE, Azzato F, Karapanagiotidis T, Leydon J, Fyfe J. Development of a real-time PCR assay to detect *Treponema pallidum* in clinical specimens and assessment of the assay's performance by comparison with serological testing. J Clin Microbiol. 2007; 45: 93–96. PMID: 17065262
20. Liu H, Rodes B, Chen CY, Steiner B. New tests for syphilis: rational design of a PCR method for detection of *Treponema pallidum* in clinical specimens using unique regions of the DNA polymerase I gene. J Clin Microbiol. 2001; 39: 1941–1946. PMID: 11326018
21. Binnicker MJ, Jespersen DJ, Rollins LO. *Treponema*-specific tests for serodiagnosis of syphilis: comparative evaluation of seven assays. J Clin Microbiol. 2011; 49: 1313–1317. doi: 10.1128/JCM.02555-10 PMID: 21346050
22. French P, Gomberg M, Janier M, Schmidt B, van Voorst Vader P, Young H. IUSTI: 2008 European guidelines on the management of syphilis. Int J STD AIDS. 2009; 20: 300–309. doi: 10.1258/ijsa.2008.008510 PMID: 19386965
23. International Union Against Sexually Transmitted Infections (IUSTI). European Guidelines on the Management of Syphilis. IUSTI. 2014; Available: http://www.iusti.org/
24. Sidana R, Mangala HC, Murugesh SB, Ravindra K. Prozone phenomenon in secondary syphilis. Indian J Sex Transm Dis. 2011; 32: 47–49. doi: 10.4103/0253-7184.81256 PMID: 21799578
25. Harner RE, Smith JL, Israel CW. The FTA-ABS test in late syphilis. A serological study in 1,985 cases. JAMA. 1968; 203: 545–548. PMID: 4865707
26. Hart G. Syphilis tests in diagnostic and therapeutic decision making. Ann Intern Med. 1986; 104: 368–376. PMID: 3511822
27. Reeves RR. Unreliability of syphilis screening tests. OTO-HNS. 2000; 122: 156.
28. Peeling RW, Ye H. Diagnostic tools for preventing and managing maternal and congenital syphilis: an overview. Bull World Health Organ. 2004; 82: 439–446. PMID: 15356937
29. Bosshard PP. Usefulness of IgM-specific enzyme immunoassays for serodiagnosis of syphilis: comparative evaluation of three different assays. J Infect. 2013; 67: 35–42. doi: 10.1016/j.jinf.2013.03.011 PMID: 23542782
30. Salazar JC, Hazlett KR, Radolf JD. The immune response to infection with *Treponema pallidum*, the stealth pathogen. Microbes Infect. 2002; 4: 1133–1140. PMID: 12361913

31. Teunis PF, van Eijkeren JC, Ang CW, van Duynhoven YT, Simonsen JB, et al. Biomarker dynamics: estimating infection rates from serological data. Stat Med. 2012; 31: 2240–2248. doi: 10.1002/sim.5322 PMID: 22419564

32. Hofmann H. Lyme borreliosis—Problems of serological diagnosis. Infection. 1996; 24: 470–472. PMID: 9007597

33. O'Neill P, Nicol CS. IgM class antitreponemal antibody in treated and untreated syphilis. Br J Vener Dis. 1972; 48: 460–463. PMID: 4568465

34. Egglestone SI, Turner AJ. Serological diagnosis of syphilis. PHLS Syphilis Serology Working Group. Commun Dis Public Health. 2000; 3: 158–162. PMID: 11014025

35. Young H. Guidelines for serological testing for syphilis. Sex Transm Infect. 2000; 76: 403–405. PMID: 11141863

36. Loeffelholz MJ, Binnicker MJ. It is time to use treponema-specific antibody screening tests for diagnosis of syphilis. J Clin Microbiol. 2012; 50: 2–6. doi: 10.1128/JCM.06347-11 PMID: 22090405

37. Sena AC, White BL, Sparling PF. Novel *Treponema pallidum* serologic tests: a paradigm shift in syphilis screening for the 21st century. Clin Infect Dis. 2010; 51: 700–708. doi: 10.1086/655832 PMID: 20687840

38. Smit PW, van der Vlis T, Mabey D, Changalucha J, Mngara J, Clark BD, et al. The development and validation of dried blood spots for external quality assurance of syphilis serology. BMC Infect Dis. 2013; 13: 102. doi: 10.1186/1471-2334-13-102 PMID: 23442198

39. Workowski KA, Berman S. Sexually transmitted diseases treatment guidelines. MMWR. 2010; 59: 1–110.

40. Pastuszczak M, Wojas-Pelc A. Current standards for diagnosis and treatment of syphilis: selection of some practical issues, based on the European (IUSTI) and U.S. (CDC) guidelines. Postepy Dermatol Alergol. 2013; 30: 203–210. doi: 10.5114/pdia.2013.37029 PMID: 24278076

41. Nandwani R, Evans DT. Are you sure it's syphilis? A review of false positive serology. Int J STD AIDS. 1995; 6: 241–248. PMID: 7548285

42. Maurice J. WHO plans new yaws eradication campaign. Lancet. 2012; 379: 1377–1378. PMID: 22509526

5.3.2 Gene Target Selection for Loop-Mediated Isothermal Amplification for Rapid Discrimination of *Treponema pallidum* Subspecies

The loop-mediated isothermal amplification (LAMP) assay for the yaws bacterium *Treponema pallidum* was developed to generate proof of concept for gene targets that can be used to distinguish between all three subspecies of the bacterium; subsp. *pallidum* causing syphilis, subsp. *pertenue* causing yaws and subsp. *endemicum* causing bejel.

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Gene target selection for loop-mediated isothermal amplification for rapid discrimination of *Treponema pallidum* subspecies

Sascha Knauf[1]*, Simone Lüert[1], David Šmajs[2], Michal Strouhal[2], Idrissa S. Chuma[1,3], Sieghard Frischmann[4], Mohammed Bakheit[4]

1 Work Group Neglected Tropical Diseases, Infection Biology Unit, German Primate Center, Leibniz- Institute for Primate Research, Göttingen, Germany, **2** Department of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic, **3** Sokoine University of Agriculture, Faculty of Veterinary Medicine, Morogoro, Tanzania, **4** Mast Diagnostica GmbH, Reinfeld, Germany

* sknauf@dpz.eu

OPEN ACCESS

Citation: Knauf S, Lüert S, Šmajs D, Strouhal M, Chuma IS, Frischmann S, et al. (2018) Gene target selection for loop-mediated isothermal amplification for rapid discrimination of *Treponema pallidum* subspecies. PLoS Negl Trop Dis 12(4): e0006396. https://doi.org/10.1371/journal.pntd.0006396

Editor: Sitara SR Ajjampur, Christian Medical College, Vellore, INDIA

Received: December 31, 2017

Accepted: March 21, 2018

Published: April 12, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. Sequence data of the TP_0619 locus obtained from the baboon sample were uploaded to NCBI GenBank (accession number: MG573304).

Funding: The nonhuman primate sample (6RUM2090716) originates from a study that was funded by the German Research Foundation (DFG; http://www.dfg.de/) KN1097/3-1 (SK). No author has been paid to write this article by a

Abstract

We show proof of concept for gene targets (*polA*, *tprL*, and TP_0619) that can be used in loop-mediated isothermal amplification (LAMP) assays to rapidly differentiate infection with any of the three *Treponema pallidum* subspecies (*pallidum* (*TPA*), *pertenue* (*TPE*), and *endemicum* (*TEN*)) and which are known to infect humans and nonhuman primates (NHPs). Four *TPA*, six human, and two NHP *TPE* strains, as well as two human *TEN* strains were used to establish and validate the LAMP assays. All three LAMP assays were highly specific for the target DNA. Amplification was rapid (5–15 min) and within a range of 10E+6 to 10E+2 of target DNA molecules. Performance in NHP clinical samples was similar to the one seen in human *TPE* strains. The newly designed LAMP assays provide proof of concept for a diagnostic tool that enhances yaws clinical diagnosis. It is highly specific for the target DNA and does not require expensive laboratory equipment. Test results can potentially be interpreted with the naked eye, which makes it suitable for the use in remote clinical settings.

Author summary

Sustainable eradication of human yaws benefits from applicable and reliable assays to detect possible reemerging yaws cases. Our study provides proof of concept for LAMP assays that are capable of rapid diagnoses and discrimination of active *Treponema pallidum* infection. While current clinical diagnosis is based on the clinical manifestations in combination with serology, the selected targets and LAMP assays allow for DNA based differentiation between skin ulcers caused by the subsp. *pallidum* (syphilis), subsp. *pertenue* (yaws), and the subsp. *endemicum* (bejel). The presented LAMP assays require limited expensive technical equipment and can be run in virtually any clinical setting. The method is thus capable of enhancing yaws diagnosis in particular in a low capacity environment.

pharmaceutical company or other agency. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: MB and SF are employed by Mast Diagnostica GmbH, a company that sells the Mast Isoplex DNA Kit that was used for LAMP reactions in this study. This however did not influence the design or the outcome of the presented study. All other authors declare no potential conflict of interest.

Introduction

Human yaws is a tropical skin disease of children caused by the bacterium *Treponema pallidum* subsp. *pertenue* (*TPE*) [1]. Skin ulcers are the most characteristic clinical manifestations associated with infection in all three active disease stages (primary, secondary, and tertiary yaws) [2]. The disease is currently subject to global eradication efforts [3], which face challenges that arise from the biology and distribution of the yaws bacterium as well as its diagnosis and treatment [4]. It is largely believed that the first yaws eradication campaigns conducted in the mid-1950s to late 1960s were successful in terms of reducing the prevalence by 95% but failed to eradicate the disease when local efforts to prevent new cases proved insufficient [5]. The majority of affected populations belong to poor and marginalized societies, with only rudimentary access to health care systems ('Where the road ends, yaws begins') [6]. Until today, standard diagnosis of yaws in clinical settings is based on clinical manifestations in combination with serology [1]. *T. pallidum* (*TP*) elicits a strong antibody response [7, 8]. Although it is possible to distinguish current infection (active or latent) from past infection when non-treponemal and treponemal tests are used in combination [9], it remains impossible based on serology and in some instances clinical manifestations, to differentiate yaws infection (*TPE*) from syphilis (caused by subsp. *pallidum* (*TPA*)) or bejel (caused by the subsp. *endemicum* (*TEN*)). Moreover, it has been shown that other diseases are capable of mimicking yaws infection. In particular, *Haemophilus ducreyi* has been reported to cause yaws-like skin ulcers [10]. Lastly, a larger number of skin ulcers in rural Africa remains etiologically undiagnosed [11], which increases the chance of overlooked infection with *TPE*. Other diseases which are capable of mimicking yaws infection are cutaneous leishmaniasis, scabies, or fungal infections [1]. Eradication of yaws is further challenged by the finding that nonhuman primates (NHPs) are infected with *TP* [12, 13]. Notably, all whole genome sequenced simian strains must be considered *TPE* strains [14, 15]. NHPs therefore must be considered as a possible natural reservoir for human infection [13]. The West African simian *TPE* strain Fribourg-Blanc, which was isolated from a Guinea baboon (*Papio papio*) in the 1960s [16], caused sustainable infection when inoculated into humans [17].

Post-eradication surveillance following the currently ongoing mass-azithromycin treatment phase [4] would benefit from rapid and cost-effective molecular tests that are able to distinguish *TPE* infection [18] from infections with all other *TP* subspecies (*TPA* and *TEN*) and bacteria that are involved in tropical skin ulcers and which may fall together with *TP* seropositivity. Potentially a single overlooked yaws case would result in a failure of global yaws eradication. Loop-mediated isothermal amplification (LAMP) was first described by Notomi et al. in 2000 [19] and since then has been extensively used to improve infectious disease diagnostics [20]. The highly specific method recognizes the DNA target using six distinct sequences initially and four distinct sequences subsequently [19]. LAMP uses a DNA polymerase with high strand displacement activity to perform a fast running auto-cycling strand displacement synthesis. Reactions run at constant temperature (isothermal) and therefore do not require expensive technical equipment such as PCR cycling machines. Our objective was to identify suitable gene targets that can be used for LAMP assay design to rapidly distinguish between yaws infection, including simian strains, and syphilis or bejel.

Materials and methods

Ethical statement and *TP* strain selection

DNA samples of human *TPA* laboratory strain Mexico A, Nichols, Seattle 81–4, SS14, *TPE* strain Gauthier, CDC-1, CDC-2, Samoa D, Sei Geringging K403, Kampung Dalan K363, as

well as the simian *TPE* strain Fribourg-Blanc were obtained from rabbit-in vivo inoculation experiments (S.A. Lukehart and DS). These experiments were not directly associated with this study. DNA extracts from human *TEN* strain Bosnia A and Iraq B originate from whole genome amplified clinical samples (DS) that were not directly associated with this study. DNA from a *TP*-infected olive baboon (*Papio anubis*; 6RUM2090716) originates from a clinical sample collected for a different study at Ruaha National Park (RNP) in Tanzania in 2015 (DFG KN1097/3-1 (SK)). Details and further reference for each strain included into the study can be found in the Supplementary S1 Table. 'Good Veterinary Practice' rules were applied to all procedures where animals were handled.

Study design and data collection

Three different LAMP assays were designed. First, we generated a LAMP assay that is able to detect DNA of all three *TP* subspecies (*TPA*, *TPE*, and *TEN*). This assay served as an initial control and was designed for the use in NHPs where little is known about the *TP* subspecies that circulate in wild NHP populations. Second, a LAMP assay was designed to distinguish *TPE* strain infection from infection with *TPA* or *TEN* strains. Third, a LAMP assay that differentiates between infection with *TPE* or *TEN* and infection with *TPA* strains has been established. All LAMP reactions were run with four human *TPA*, six human *TPE* and two simian *TPE* strains, as well as two human *TEN* strains of known copy number (S1 Table). All tests were run as triplicates and included a DNA-free negative control.

Dilution series of target DNA were used to identify the analytic limits of detection for each of the specific LAMP reactions using appropriate strain material. 10-fold serial dilutions of the target DNA were applied to cover a range of at least five decimal powers, from the maximum of *TP* copy numbers (strain Nichols 10E+5, all other strains 10E+6) until 10E+0. Negative controls that contained no DNA and dilution steps that contained $\leq$10E+2 *TP* copies were run as at least six replicates. A StepOnePlus Real-Time PCR System (ThermoFisher Scientific) was used to run the reactions and to collect the data. Due to software restrictions, it was necessary to introduce a (neglectable) thermal cycling step into the protocol. Each LAMP run therefore encompassed continuous 40 cycling steps each consistent of 63˚C for five seconds followed by 64˚C for one minute and data collection.

LAMP reactions

LAMP reactions were performed in a volume of 25.0 μl using the Mast Isoplex DNA Kit (#REF67dnamp, Mast Diagnostica GmbH). According to the manufacture's guidance, each reaction consisted of 12.5 μl of the kit's 2x reaction mix, 1.0 μl Bst polymerase enzyme, 1.0 μl fluorochrome dye, and 2.0 μl of the primer mix. One microliter target DNA was included and distilled water (molecular grade) was used to top up the reaction volume until 25.0 μl were reached. All primers were heat pre-treated at 95˚C for 5 min and immediately cooled on ice prior to adding them to the master mix. The primer mix contained 1.6 μM each FIP and BIP, 0.2 μM each F3 and B3, as well as 0.8 μM each LF and LB primer. All reactions were run on a MircoAmp Fast Optical 96-well reaction plate (#4346907, ThermoFisher).

Oligonucleotide primer design

Oligonucleotide primers were designed using the PrimerExplorer v5 Software (http://primerexplorer.jp/e/). Each LAMP primer set consisted of six oligonucleotide primers (Table 1). The design followed the description given by Yoshida et al. 2005 [21]. Briefly, a set of four primers (F3, B3, the forward inner primer [FIP], and backward inner primer [BIP]), which bind six loci of the target gene (F1, F2, F3, B1, B2, and B3) are necessary. The two inner primers (FIP and

Table 1. Oligonucleotide primers used in this study.

LAMP	Type	Sequence	Gene target
TP	F3	5'-ATTGGTCCTAAGACGGCT	*polA* (TP_0105)
	B3	5'-GCGGAATACAACAGGAATC	
	FIP	5'-CAGCGCTTCTTTTAAGGAATAGGTAGCACATCTTCTCCACTGT	
	BIP	5'-CGCACGAAGATAGTGTGTGGACATGGTACATCGTCACG	
	LF	5'-CGATAAATACCATCAAGTGTGCCAAA	
	LB	5'-GAAGAAAGATGCATTTTTTTCTCGTTC	
TPA/*TEN*	F3	5'-TGCAGTCTTTTTTTGTGCGG	*tprL* (TP_1031)
	B3	5'-TCAAATCATTGGTGGTGCGA	
	FIP	5'-CTGCGGGGCAATGCCTAGCTGCTCCCGGGGTTTGG	
	BIP	5'-GTAACTGGTCACGCCCAGCTGCCCGTGCGTGTACTCGTTC	
	LF	5'-AGAAAAAACGGGCAGTGCG	
	LB	5'-GGGGCATTAAGTTTAAAAAGAACCC	
TPE/*TEN*	F3	5'-TCCCTCGAGACTCGATTCC	TP_0619
	B3	5'-GTACGCATCTGTCGGTAGC	
	FIP	5'-GTTGCACAGAAGGGGGGTGGCATCCTTCTGatGCGTCAGC	
	BIP	5'-GTGCTTTTCGGGAGGATTCCCGAGTGCACTGCGCTCATCT	
	LF	5'-TTCCGCGTTCGGTGCTC	
	LB	5'-CTTTCGGGACGATGAGATAATGC	

https://doi.org/10.1371/journal.pntd.0006396.t001

BIP) are a sequence combination of sense and antisense sequences of the DNA. This is essential for the priming in the first stage and the self-priming in the later stages. Therefore, FIP primers consist of the combination of sequences defined as F1c (c = complementary) and F2. Likewise, BIP primers are composed of primer sequences B1c and B2. To enhance amplification efficacy, two loop primers LF and LB were added to each of the LAMP primer sets. To confirm the specificity of the newly designed primers, we performed a search for orthologous sequences using BLASTn at the NCBI homepage (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Gene locus selection

The LAMP primer set '*TP*' targets the polymerase I (*polA*) gene (TP_0105) of *TP*. The locus is highly specific for all *TP* subspecies [22] and has only one orthologue in the lagomorph infecting *Treponema paraluisleporidarum* ecovar *Cuniculus*. The latter is not capable of infecting humans [23, 24]. This locus therefore allows the reaction to become positive for DNA of any known *TPA*, *TPE*, or *TEN* strain (Fig 1A).

LAMP primer set '*TPA*/*TEN*' targets the *tprL* gene (TP_1031) of *TP*. At this locus, a 278-bp long deletion exists that distinguishes known human *TPA* and *TEN* strains from human and simian *TPE* strains [25, S1 Fig]. This primer set was specifically designed to bind within the deletion part, which creates the specificity for *TPA* and *TEN* strains (Fig 1B).

LAMP primer set '*TPE*/*TEN*' targets the *T. pallidum* TP_0619 gene, which has recently been investigated in a multilocus-typing study on *TPE* strains [26]. This locus has a 179-bp long sequence part that distinguishes known human and NHP *TPE* as well as *TEN* strains from *TPA* human strains (Fig 1C, S2 Fig). Primer sequence data of all three LAMP primer sets are listed in Table 1.

Quantification of strain material

All *TP* strains used in this study were quantified using an established [27] but slightly modified TaqMan PCR (qPCR) targeting the *polA* gene. A dilution series of a plasmid containing the

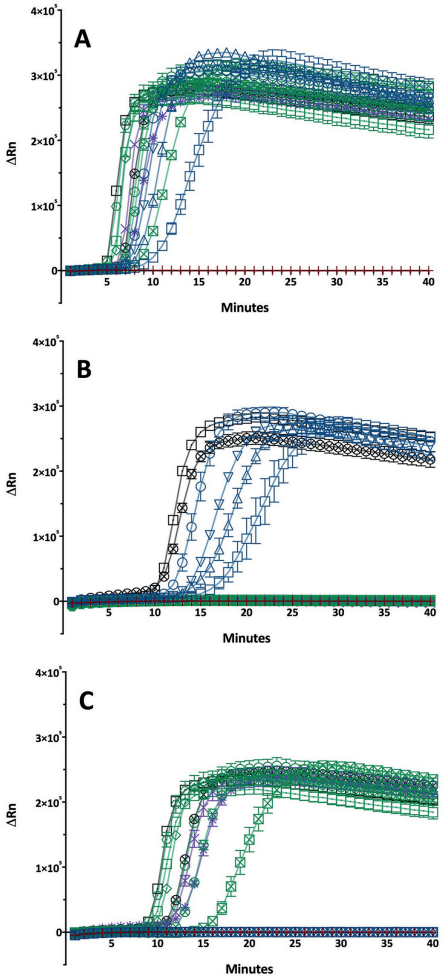


Fig 1. Performance characteristics of the three LAMP assays. (A) LAMP targeting the *polA* gene. All *TP* strains become positive. (B) LAMP assay targeting the *tprL* locus, which results in amplification of *TPA* and *TEN* strains. (C) Only *TPE* and *TEN* strains generate positive results in the LAMP assay targeting the TP_0619 locus. Note that the strain material in the three assays had different *TP* copy numbers as indicated in S1 Table. Data points are presented as mean±SEM values. Red cross = Negative control; *TPA* strains (blue) up-pointing triangle = Mexico A, down-pointing triangle = SS14, circle = Nichols, square = Seattle 81–4; *TPE* strains (green) square = Gauthier, square turned = Sei Geringing K403, hexagon = Kampung Dalan K363, circle = Samoa D, circle with cross = CDC-1, square with cross CDC-2; *TPE* simian strains (purple) cross = Fribourg-Blanc, star = RNP; *TEN* strains (black) square = Bosnia A, circle with cross = Iraq B.

https://doi.org/10.1371/journal.pntd.0006396.g001

target amplicon was used as a standard curve from 10E+7 to 10E+0 copy numbers. Briefly, each reaction volumes contained 10.0 μl TaqMan Universal Mastermix II (no Uracil N-glycosylase, Applied Biosystems), 1.8 μl each 10 μM primer and probe, 3.6 μl molecular grade water (RNase-free; Qiagen), and one microliter of the target DNA. Samples were quantified using a StepOne Plus Real-time system with the following temperature steps: 50˚C for two minutes, 95˚C for ten minutes, followed by 50 cycles of 95˚C for 15 seconds and 60˚C for one minute. At the end of each cycle, fluorescence was measured. All samples and standards were run as triplicates.

Data analyses and statistics

LAMP performance as well as qPCR data were retrieved from the StepOnePlus Real-Time PCR System and extracted as RAW data into Excel sheets utilizing the StepOne Software v2.3 (Life Technologies). Statistical analyses were performed with Prism 7.0 (GraphPad Software). In LAMP dilution series with low copy numbers (≤10E+2) and in qPCR data, single replicate outliers were excluded.

Results

LAMP assays are highly specific for *T. pallidum* target DNA

The LAMP assay targeting the *polA* gene was positive for all tested *TP* strain samples including the four *TPA*, six human *TPE*, two simian *TPE*, and the two human *TEN* strains (Fig 1A, S1 Table). The *tprL* targeting LAMP was positive for all tested *TPA* and *TEN* strains, while human and NHP *TPE* strains did not amplify (Fig 1B). The LAMP assay that uses a part of the TP_0619 gene generated positive results for all *TPE* strains including simian *TPE* strains as well as the two human *TEN* strains (Fig 1C). The onset of exponential fluorescence increase (ΔRn) started reproducibly between 5 min and 15 min incubation time (Fig 1A–1C). Melting curves for each LAMP assay are shown in S3 Fig. All curves were of appropriate shape and without any additional peaks indicative for unwanted side products of primer dimers.

Limits of detection are suitable for clinical samples

Analytic limits of detection were assessed as demonstrated in several published studies [28–30]. The LAMP assay that targets the *polA* locus amplifies all *TP* strains but differed slightly in its detection limit across the different *TP* subspecies. While the *TPA* strain Nichols failed to amplify between 10E+3 and 10E+2 copies (Fig 2A), the *TPE* strain Gauthier showed a non-exponential increase in fluorescence at 10E+2 copies (Fig 2B). *TEN* strain Bosnia A failed to exponentially amplify at 10E+1 copies (Fig 2C; Table 2). The LAMP targeting the *tprL* locus had a detection limit of 10E+2 copies for Nichols (Fig 2D) and 10E+3 for *TEN* strain Bosnia A (Fig 2E; Table 2).

The LAMP assay that utilizes the TP_0619 locus amplified *TPE* (strain Gauthier) and *TEN* (strain Bosnia A) DNA until a total copy number of 10E+2 copies was reached (Fig 2F and 2G; Table 2).

Discussion

Suitability of the selected gene targets for *T. pallidum* subsp. discriminating LAMP assays

In many areas where endemic treponematoses occur, syphilis can also be found at meaningful prevalence rates (e.g., Ghana 3.7% [31], Papua New Guinea 7.9% (men)-12.9% (women) [32]).

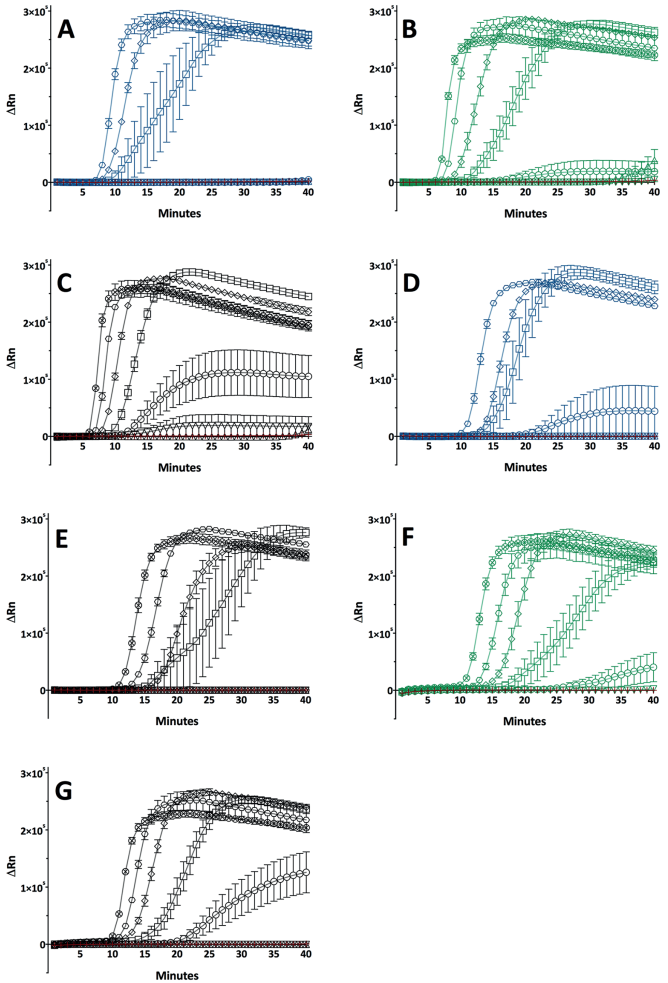


Fig 2. Detection limits of LAMP assays. (A-C) *TP* LAMP assay involving (A) *TPA* strain Nichols with copy numbers 10E+5 to 10E+0, (B) *TPE* strain Gauthier tested in a range of 10E+6 to 10E+0 copies, and (C) *TEN* strain Bosnia A in 10E+6 to 10E+0 copy numbers. (D-E) *TPA* LAMP assay run with *TPA* strain Nichols from 10E+5 to 10E+0 copy numbers, and the same LAMP assay run with (E) *TEN* strain Bosnia A in copy numbers ranging from 10E+6 to 10E+0. (F-G) *TPE*/*TEN* LAMP assay involving (E) *TPE* strain Gauthier tested in 10E+6 to 10E+0 copy numbers, and (F) *TEN* strain Bosnia A tested in 10E+6 to 10E+0 copies. Red cross = Negative control, *TPA* strain Nichols (blue), *TPE* strain Gauthier (green), *TEN* strain Bosnia A (black); symbols represent copy numbers: circle with cross = 10E+6, hexagon = 10E+5, square turned = 10E+4, square = 10E+3, circle = 10E+2, down-pointing triangle = 10E+1, up-pointing triangle = 10E+0.

https://doi.org/10.1371/journal.pntd.0006396.g002

Table 2. Detection limits for the three LAMP assays. Corresponding graphs can be found in Fig 2. + = exponential amplification, (+) = no exponential amplification, — = no amplification.

	Gene	*TP*	Dilution step (total *TP* copy number per reaction)						
LAMP	**Locus**	**Strain**	**10E+6**	**10E+5**	**10E+4**	**10E+3**	**10E+2**	**10E+1**	**10E+0**
TP	*polA* (TP_0105)	Nichols	n/a	+	+	+	-	-	-
		Gauthier	+	+	+	+	(+)	-	-
		Bosnia A	+	+	+	+	+	(+)	-
TPA/TEN	*tprL* (TP_1031)	Nichols	n/a	+	+	+	+	-	-
		Bosnia A	+	+	+	+	-	-	-
TPE/TEN	TP_0619	Gauthier	+	+	+	+	(+)	-	-
		Bosnia A	+	+	+	+	+	-	-

https://doi.org/10.1371/journal.pntd.0006396.t002

While this is a problem for the serological based diagnosis of yaws in the presence of etiologically unrelated skin ulcers, it is not an issue for LAMP assays, which specifically target the DNA of the pathogen. The *TPE/TEN* LAMP was able to reliably discriminate yaws and simian *TPE* infection from infection with syphilis causing strains. It will, however, not discriminate yaws-causing strains from those known to cause bejel (*TEN* strains). While in theory this could be a problem, bejel is a disease found in the dry areas of Sahelian Africa and Saudia Arabia and thus its spatial distribution does not overlap with yaws reporting countries in Western and Central Africa, Southeast Asia, and the Pacific Islands [33]. In cases where a clear differentiation between yaws and bejel infection is important, the combination of the LAMP targeting *tprL1* and TP_0619 will enable the distinction of both subspecies since only *TEN* strains will amplify in both assays. In the future, either *TEN* specific assays or a LAMP multiplex assay can be designed [34]. The latter, however, would require a fluorescence measuring device and thus may restricts the use in remote tropical health care facilities. Our study used the StepOnePlus Real-Time PCR System, but all LAMP assays described in this study can be run equally well on a portable system (e.g., ESEQuant TS2, Qiagen) that allows easy transportation and use under field conditions. In low-income clinical settings, it would even be possible to detect amplification by the naked eye through the detection of turbidity generated by the precipitation of magnesium pyrophosphate or through the addition of calcein, a fluorescent metal indicator [35]. Lyophilization allows for ambient storage of formulated LAMP reagents [36].

As indicated in the methods, all three gene targets that were selected for the LAMP assays are highly specific for the human and NHPs specific pathogenic *TP*, but also the lagomorph infecting *T. paraluisleporidarum* ecovar *Cuniculus* and *Lepus*, respectively. However, lagomorph infecting treponemes are not capable of infecting humans [23, 24] and most probably also NHPs. False positive test results due to infection with non-*TP* bacteria are therefore unlikely. In light of a recently published report on failure of qPCR due to variations in primer binding sites [37], it should be noted that the number of published genomes, in particular non-draft genomes, in any of the *TP* subspecies is low. At this stage, a general statement on genome variability at the selected gene target sites is therefore not possible. However, based on our research, which included representatives of the full range of published *TP* genomes (Table 1, S1 and S2 Figs), the relevant primer binding sites are conserved across the different subspecies and strains.

LAMP as a rapid and reliable discrimination tool

It has been proposed that yaws eradication in humans is possible through total community treatment in combination with subsequent total target treatment [38]. Rapid and reliable identification of yaws infection is important because successful global eradication requires an

infinite zero-case scenario. In the first years after eradication has been declared in humans, it might well happen that few cases reemerge from yet unidentified relapsing latent yaws cases as well as there is a theoretical change that sporadic spillover from infected wild NHPs in Africa occurs. Either way, an available molecular test such as a LAMP assay could effectively and timely identify new cases from etiologically unrelated skin ulcers at the very beginning and could help to prevent yaws from re-emerging in areas where PCR machines and expensive laboratory equipment are not available. The analytic limits of detection for all three LAMP assays were around 10E+2 copies per reaction (Table 2), which is sufficient for clinical samples from human primary and secondary syphilis infection [39]. The same numbers can be expected for human yaws samples. Furthermore, the amount of *TP* in chronically infected monkeys also falls within the detection range of the *TPE*/*TEN* LAMP [27].

Application for epidemiological studies in nonhuman primates and one health

NHP *TPE* strains have been discussed as a possible source for human yaws infection in Africa [13]. The identification of NHP populations that harbor the pathogen, not only in Africa but also Asia [12], must be considered an important research priority [4]. Post-treatment surveillance needs to focus in particular on areas where NHPs and humans are in close contact. The *TPE*/*TEN* LAMP performance of the NHP samples (strain Fribourg-Blanc and DNA extracted from a clinical sample of a baboon at Ruaha National Park in Tanzania (RNP)) that were included into this study were similar to the results obtained for the human yaws-causing strains (Fig 1C). This is not surprising, given the fact that NHP *TPE* strains are genetically and functionally highly similar to human yaws causing strains [14, 15]. However, the full diversity of NHP infecting *TP* is unknown and it is possible that monkeys from Sahelian Africa and Saudia Arabia may carry *TEN* strains. In this case, the *TPE*/*TEN* LAMP assay would become positive. Due to the fact that currently all naturally occurring NHP infections with *TP* should be whole genome sequenced to fuel our understanding on yaws epidemiology and evolution, the *TPE*/*TEN* LAMP assay result would be more of academic than practical interest. The whole genome data derived from simian isolates would reveal the subspecies status of the isolate.

In humans, infections with all *TP* subspecies have reported potential to cause atypical clinical manifestations. A striking example is the frequent syphilis-like manifestations associated with *TEN* strains [40,41]. A rapid, highly sensitive and specific LAMP assay would therefore contribute to the identification of atypical clinical manifestations caused by *TP*. It would further help to identify possible NHP-to-human infection in countries like Tanzania, where human yaws has not been reported since decades. Syphilis screening programs in Tanzania would currently not detect possible NHP-to-human transmission events, since serological tests cannot discriminate between the *TPA* and *TPE* infection. Our target selection for LAMP assays that discriminate infection with *TP* from other causes of skin ulcers, represents a basis for the implementation of a One Health approach in yaws eradication and its post-eradication surveillance. Fig 3 illustrates the proposed new way of diagnosing *TPE* infection in humans. The new LAMP assays would simplify and accelerate yaws diagnosis. We note here that we have reached proof of concept for the suitability of the described gene targets, but further validation in a statistically adequate number of clinical samples is necessary to achieve confidence of the LAMP assays to be used in a non-research environment.

Conclusion

The selected gene targets are suitable for the diagnosis and discrimination of all three *TP* subspecies, which is currently not possible using clinical signs of infection in combination with

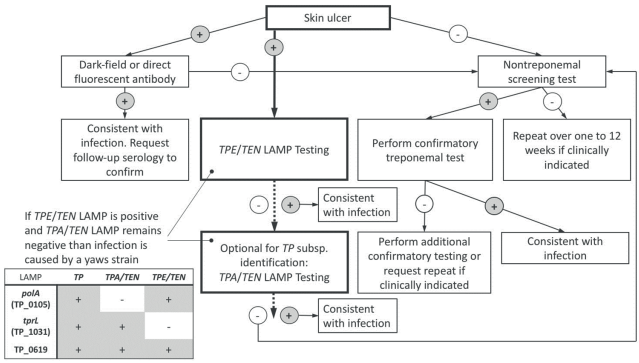


Fig 3. LAMP assays will simplify the existing test algorithm for human treponematoses. The test algorithm proposed by Ratnam 2005 [42] is shown in thin lines. The new LAMP assays for diagnosis of yaws is shown in bold lines and will simplify and accelerate yaws diagnosis. Furthermore, it enables the clinician to rapidly discriminate *TPE* infection from infection with other *TP* subspecies or pathogens.

https://doi.org/10.1371/journal.pntd.0006396.g003

serology. The next step must be to conduct tests to evaluate the sensitivity and specificity of the newly created assays in various clinical samples that originate from humans and NHPs. The designed LAMP assays do not require expensive laboratory equipment and can be run in virtually any clinical setting. Results are available within a few minutes and thus outrun the existing methods, which require reasonable laboratory infrastructure [4].

Supporting information

S1 Fig. *tprL* alignment showing the sequence differences between TPA/TEN and TPE strains. Nucleotide position 1 in the alignment refers to nucleotide position 1,125587 in the TPA str. Nichols genome (CP004010) and nucleotide position 410 to 1,125,993, respectively. Green annotations on the consensus sequence indicate the outer LAMP primer binding sites F3 and B3 (Table 1). GenBank accession numbers for TPA: CP010561 str. Nichols Houston clone J, CP010560 str. Nichols Houston clone E, CP004010 str. Nichols, AE000520 str. Nichols, HM 585246 str. DAL-1, CP003115 str. DAL-1, CP010558 str. Chicago, CP001752 str. Chicago, CP010422 str. Seattle Nichols, CP003679 str. Sea 81–4, CP010559 str. CDC-A, CP004011 str. SS14, CP000805 str. SS14, CP016066 str. PT_SIF1278, CP016057 str. PT_SIF1142, CP016056 str. PT_SIF121140, CP01050 str. PT_SIF0954, CP016049 str. PT_SIF0908, CP016045 str. PT_SIF0697, CP016055 str. PT_SIF1135, CP016053 str. PT_SIF1063, CP016069 str. PT_SIF1348, CP016068 str. PT_SIF1299, CP016067 str. PT_SIF1280, CP016065 str. PT_SIF1261, CP016064 str. PT_SIF1252, CP016063 str. PT_SIF1242, CP016062 str. PT_SIF1200, CP0116061 str. PT_SIF1196, CP016060 str. PT_SIF1183, CP016058 str. PT_SIF1156, CP016051 str. PT_SIF1002, CP016048 str. PT_SIF0877, CP016047 str. PT_SIF0857, CP016046 str. PT_SIF0751, CP016059 str. PT_SIF1167, CP016052 str. PT_SIF1020, CP016054 str. PT_SIF1127, CP010562 str. UW074B, CP010563 str. UW189B, CP010565 str. UW254B, CP010564 str. UW228B, CP010566 str. UW391B, CP015162 str. Amoy, NC_018722 str. MexA, CP003064 str. MexA, HM585253 str. MexA, TEN: CP007548 str. BosA, TPE: JX079830 str. Gauthier, HM585235.1 str. Gauthier, CP002375.1 str. CDC-2,

JX079832 str. CDC-2, CP020365 str. Ghana-051, CP020366 str. CDC-2575, CP002374.1 str. SamoaD, JX079831 str. SamoaD, ERR1470330, ERR1470331, ERR1470334, ERR1470338, ERR1470343, ERR1470344, JX079833 str. Fribourg-Blanc, CP003902.1 str. Fribourg-Blanc.
(DOCX)

S2 Fig. TP_0619 alignment showing the sequence differences between TPA and TPE/TEN strains. Nucleotide position 1 in the alignment refers to nucleotide position 671,979 in the TPA str. Nichols genome (CP004010) and nucleotide position 239 to 672,216, respectively. Green annotations on the consensus sequence indicate the outer LAMP primer binding sites F3 and B3 (Table 1). GenBank accession numbers for TPA: CP003115 str. DAL-1, CP000805 str. SS14, CP004011 str. SS14, CP001752 str. Chicago, CP003064 str. MexA, CP004010 str. Nichols, AE000520 str. Nichols, CP003679 str. Sea81-4, CP010559 str. CDC A, CP010558 str. Chicago, CP010560 str. Nichols Houston, CP010561 str. Nichols Houston, CP016045 str. PT_SIF0697, CP016046 str. PT_SIF0751, CP016047 str. PT_SIF0857, CP016048 str. PT_SIF0877, CP016049 str. PT_SIF0908, CP016050 str. PT_SIF0954, CP016051 str. PT_SIF1002, CP016052 str. PT_SIF1020, CP016053 CP016054 str. PT_SIF1127, CP016055 str. PT_SIF1135, CP016056 str. PT_SIF1140, CP016057 str. PT_SIF1142, CP016058 str. PT_SIF1156, CP016059 str. PT_SIF1167, CP016060 str. PT_SIF1183, CP016061 str. PT_SIF1196, CP016062 str. PT_SIF1200, CP016063 str. PT_SIF1242, CP016064 str. PT_SIF1252, CP016065 str. PT_SIF1261, CP016066 str. PT_SIF1278, CP016067 str. PT_SIF1280, CP016068 str. PT_SIF1299, CP016069 str. PT_SIF1348, CP010422 str. Seattle Nichols, CP10562 str. UW074B, CP10563 str. UW189B, CP10564 str. UW228B, CP10565 str. UW254B, CP10566 str. UW391B, TEN: CP007548 str. BosA, KY120834.1 str. IraqB, TPE: CP002375 str. CDC2, CP002376 str. Gauthier, CP002374 str. SamoaD, CP020366 str. CDC2575, CP020365 str. Ghana-051, CP003902 str. Fribourg-Blanc, MG573304 str. RuahaNP-1 (6RUM2090716).
(DOCX)

S3 Fig. Melting curves of the different TP LAMP assays. Melting curves were of appropriate shape and without any additional peaks indicative for unwanted side products or primer dimers. LAMP targeting (A) the *polA* gene, (B) the *tprL* locus, and (C) the TP_0619 locus.
(DOCX)

S1 Table. Details and further reference on *T. pallidum* strains included into the study. RNP = Ruaha National Park.
(DOCX)

S1 References.
(DOCX)

Acknowledgments

Barbara Molini, Charmie B. Gordornes, and Sheila A. Lukehart (University of Washington) are thanked for the provision of *TPA* and *TPE* strain materials. Alexander Hahn, Anna K. Großkopf, and Sarah S. Schlagowski (Junior Research Group Herpes viruses, German Primate Center) are thanked for the co-use of the StepOnePlus Real-Time PCR System. Luisa Hallmaier-Wacker (Work Group Neglected Tropical Diseases, German Primate Center) is thanked for proofreading the manuscript.

Author Contributions

Conceptualization: Sascha Knauf, Mohammed Bakheit.

Data curation: Sascha Knauf, Mohammed Bakheit.

Formal analysis: Sascha Knauf, Sieghard Frischmann, Mohammed Bakheit.

Investigation: Sascha Knauf, Idrissa S. Chuma.

Methodology: Sascha Knauf, Simone Lüert, David Šmajs, Michal Strouhal, Idrissa S. Chuma, Sieghard Frischmann, Mohammed Bakheit.

Project administration: Sascha Knauf.

Resources: Sascha Knauf, Mohammed Bakheit.

Supervision: Sascha Knauf.

Validation: Sascha Knauf, Simone Lüert, Mohammed Bakheit.

Visualization: Sascha Knauf.

Writing – original draft: Sascha Knauf, Mohammed Bakheit.

Writing – review & editing: Sascha Knauf, Simone Lüert, David Šmajs, Michal Strouhal, Idrissa S. Chuma, Sieghard Frischmann, Mohammed Bakheit.

References

1. Mitjà O, Asiedu K, Mabey D. Yaws. Lancet. 2013; 381(9868):763–73. Epub 2013/02/19.
2. Marks M, Mitjà O, Solomon AW, Asiedu KB, Mabey DC. Yaws. Br Med Bull. 2015; 113(1):91–100. Epub 2014/12/20. https://doi.org/10.1093/bmb/ldu037 PMID: 25525120
3. Asiedu K, Fitzpatrick C, Jannin J. Eradication of yaws: historical efforts and achieving WHO's 2020 target. PLoS Negl Trop Dis. 2014; 8(9):e3016. https://doi.org/10.1371/journal.pntd.0003016 PMID: 25254372
4. Marks M, Mitjà O, Vestergaard LS, Pillay A, Knauf S, Chen CY, et al. Challenges and key research questions for yaws eradication. Lancet Infect Dis. 2015; 15(10):1220–5. https://doi.org/10.1016/S1473-3099(15)00136-X PMID: 26362174
5. Asiedu K, Amouzou B, Dhariwal A, Karam M, Lobo D, Patnaik S, et al. Yaws eradication: past efforts and future perspectives. Bull World Health Organ. 2008; 86(7):499–A. Epub 2008/08/02. https://doi.org/10.2471/BLT.08.055608 PMID: 18670655
6. Fitzpatrick C, Asiedu K, Jannin J. Where the Road Ends, Yaws Begins? The Cost-effectiveness of Eradication versus More Roads. PLoS Negl Trop Dis. 2014; 8(9):e3165. https://doi.org/10.1371/journal.pntd.0003165 PMID: 25255131
7. McGill MA, Edmondson DG, Carroll JA, Cook RG, Orkiszewski RS, Norris SJ. Characterization and serologic analysis of the *Treponema pallidum* proteome. Infect Immun. 2010; 78(6):2631–43. Epub 2010/04/14. https://doi.org/10.1128/IAI.00173-10 PMID: 20385758
8. Brinkman MB, McGill MA, Pettersson J, Rogers A, Matejková P, Šmajs D, et al. A novel *Treponema pallidum* antigen, TP0136, is an outer membrane protein that binds human fibronectin. Infect Immun. 2008; 76(5):1848–57. Epub 2008/03/12. https://doi.org/10.1128/IAI.01424-07 PMID: 18332212
9. Marks M, Yin YP, Chen XS, Castro A, Causer L, Guy R, et al. Metaanalysis of the performance of a combined treponemal and nontreponemal rapid diagnostic test for syphilis and yaws. Clin Infect Dis. 2016; 63(5):627–33. https://doi.org/10.1093/cid/ciw348 PMID: 27217216
10. Mitjà O, Lukehart SA, Pokowas G, Moses P, Kapa A, Godornes C, et al. *Haemophilus ducreyi* as a cause of skin ulcers in children from a yaws-endemic area of Papua New Guinea: a prospective cohort study. Lancet Glob Health. 2014; 2(4):e235–41. https://doi.org/10.1016/S2214-109X(14)70019-1 PMID: 25103064.
11. Gibbs S. Skin disease and socioeconomic conditions in rural Africa: Tanzania. Int J Dermatol. 1996; 35 (9):633–9. Epub 1996/09/01. https://doi.org/10.1111/j.1365-4362.1996.tb03687.x PMID: 8876289.
12. Klegarth AR, Ezeonwu CA, Rompis A, Lee BPY, Aggimarangsee N, Chalise M, et al. Survey of treponemal infections in free-ranging and captive macaques, 1999–2012. Emerg Infect Dis. 2017; 23(5):816–9. https://doi.org/10.3201/eid2305.161838 PMID: 28418297

13. Knauf S, Liu H, Harper KN. Treponemal infection in nonhuman primates as possible reservoir for human yaws. Emerg Infect Dis. 2013; 19(12):2058–60. Epub 2013/11/28. https://doi.org/10.3201/eid1912.130863 PMID: 24274094

14. Knauf S, Gogarten J, Schuenemann VJ, De Nys HM, Duex A, Strouhal M, et al. African nonhuman primates are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. bioRxiv. 2017. https://doi.org/10.1101/135491

15. Zobaníková M, Strouhal M, Mikalová L, Čejková D, Ambrozova L, Pospíšilová P, et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl Trop Dis. 2013; 7(4):e2172. Epub 2013/05/03. https://doi.org/10.1371/journal.pntd.0002172 PMID: 23638193

16. Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. Primates Med. 1969; 3 (0):113–21. Epub 1969/01/01. PMID: 5006024.

17. Smith JL, David NJ, Indgin S, Israel CW, Levine BM, Justice J Jr., et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br J Vener Dis. 1971; 47(4):226–51. Epub 1971/08/01. PMID: 4936861

18. Pahil S, Goyal K. Rapid molecular diagnosis of chronic skin ulcers. Lancet Glob Health. 2917; 2: e385.

19. Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, et al. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. 2000; 28(12):E63. PMID: 10871386

20. Mori Y, Notomi T. Loop-mediated isothermal amplification (LAMP): a rapid, accurate, and cost-effective diagnostic method for infectious diseases. J Infect Chemother. 2009; 15(2):62–9. https://doi.org/10.1007/s10156-009-0669-9 PMID: 19396514.

21. Yoshida A, Nagashima S, Ansai T, Tachibana M, Kato H, Watari H, et al. Loop-mediated isothermal amplification method for rapid detection of the periodontopathic bacteria *Porphyromonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola*. J Clin Microbiol. 2005; 43(5):2418–24. https://doi.org/10.1128/JCM.43.5.2418-2424.2005 PMID: 15872275

22. Liu H, Rodes B, Chen CY, Steiner B. New tests for syphilis: rational design of a PCR method for detection of *Treponema pallidum* in clinical specimens using unique regions of the DNA polymerase I gene. J Clin Microbiol. 2001; 39(5):1941–6. Epub 2001/04/28. https://doi.org/10.1128/JCM.39.5.1941-1946.2001 PMID: 11326018

23. Šmajs D, Zobaníková M, Strouhal M, Čejková D, Dugan-Rocha S, Pospíšilová P, et al. Complete genome sequence of *Treponema paraluiscuniculi*, strain Cuniculi A: the loss of infectivity to humans is associated with genome decay. PLoS One. 2011; 6(5):e20415. https://doi.org/10.1371/journal.pone.0020415 PMID: 21655244

24. Graves S, Downes J. Experimental infection of man with rabbit-virulent *Treponema paraluis-cuniculi*. Br J Vener Dis. 1981; 57(1):7–10. PMID: 7470837

25. Centurion-Lara A, Giacani L, Godornes C, Molini BJ, Brinck Reid T, Lukehart SA. Fine analysis of genetic diversity of the *tpr* gene family among treponemal species, subspecies and strains. PLoS Negl Trop Dis. 2013; 7(5):e2222. https://doi.org/10.1371/journal.pntd.0002222 PMID: 23696912

26. Godornes C, Giacani L, Barry A, Mitjà O, Lukehart S (2017) Development of a multilocus sequence typing (MLST) scheme for *Treponema pallidum* subsp. *pertenue*: Application to yaws in Lihir Island, Papua New Guinea. PLoS Negl Trop Dis. 11: e0006113. https://doi.org/10.1371/journal.pntd.0006113 PMID: 29281641

27. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IA, Nordhoff M, et al. *Treponema* infection associated with genital ulceration in wild baboons. Vet Pathol. 2012; 49(2):292–303. Epub 2011/03/18. https://doi.org/10.1177/0300985811402839 PMID: 21411621.

28. Iwamoto T, Sonobe T, Hayashi K. Loop-mediated isothermal amplification for direct detection of *Mycobacterium tuberculosis* complex, *M. avium*, and *M. intracellulare* in sputum samples. J Clin Microbiol. 2003; 41(6):2616–22. Epub 2003/06/07. https://doi.org/10.1128/JCM.41.6.2616-2622.2003 PMID: 12791888

29. Aryan E, Makvandi M, Farajzadeh A, Huygen K, Bifani P, Mousavi SL, et al. A novel and more sensitive loop-mediated isothermal amplification assay targeting IS6110 for detection of *Mycobacterium tuberculosis* complex. Microbiol Res. 2010; 165(3):211–20. https://doi.org/10.1016/j.micres.2009.05.001 PMID: 19515543.

30. Das A, Babiuk S, McIntosh MT. Development of a loop-mediated isothermal amplification assay for rapid detection of capripoxviruses. J Clin Microbiol. 2012; 50(5):1613–20. https://doi.org/10.1128/JCM.06796-11 PMID: 22357504

31. Sarkodie F, Hassall O, Owusu-Dabo E, Owusu-Ofori S, Bates I, Bygbjerg IC, et al. Syphilis screening practices in blood transfusion facilities in Ghana. Int J Infect Dis. 2016; 43:90–4. https://doi.org/10.1016/j.ijid.2015.12.020 PMID: 26751240.

32. Vallely A, Page A, Dias S, Siba P, Lupiwa T, Law G, et al. The prevalence of sexually transmitted infections in Papua New Guinea: a systematic review and meta-analysis. PLoS One. 2010; 5(12):e15586. https://doi.org/10.1371/journal.pone.0015586 PMID: 21203468

33. Giacani L, Lukehart SA. The endemic treponematoses. Clin Microbiol Rev. 2014; 27(1):89–115. Epub 2014/01/08. https://doi.org/10.1128/CMR.00070-13 PMID: 24396138

34. Liu N, Zou D, Dong D, Yang Z, Ao D, Liu W, et al. Development of a multiplex loop-mediated isothermal amplification method for the simultaneous detection of *Salmonella* spp. and *Vibrio parahaemolyticus*. Sci Rep. 2017; 7:45601. https://doi.org/10.1038/srep45601 PMID: 28349967

35. Tomita N, Mori Y, Kanda H, Notomi T. Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. Nat Protoc. 2008; 3(5):877–82. https://doi.org/10.1038/nprot.2008.57 PMID: 18451795.

36. Chander Y, Koelbl J, Puckett J, Moser MJ, Klingele AJ, Liles MR, et al. A novel thermostable polymerase for RNA and DNA loop-mediated isothermal amplification (LAMP). Front Microbiol. 2014; 5:395. https://doi.org/10.3389/fmicb.2014.00395 PMID: 25136338

37. Marks M, Fookes M, Wagner J, Butcher R, Ghinai R, Sokana et al. (2017) Diagnostics for yaws eradication: insights from direct next generation sequencing of cutaneous strains of *Treponema pallidum*. Clin Infect Dis. 2017; https://doi.org/10.1093/cid/cix892 PMID: 29045605.

38. Marks M, Mitjà O, Fitzpatrick C, Asiedu K, Solomon AW, Mabey DC, et al. Mathematical modeling of programmatic requirements for yaws eradication. Emerg Infect Dis. 2017; 23(1):22–8. Epub 2016/12/17. https://doi.org/10.3201/eid2301.160487 PMID: 27983500

39. Tipple C, Hanna MO, Hill S, Daniel J, Goldmeier D, McClure MO, et al. Getting the measure of syphilis: qPCR to better understand early infection. Sex Transm Infect. 2011; 87(6):479–85. https://doi.org/10.1136/sti.2011.049494 PMID: 21752804

40. Grange P, Mikalová L, Gaudin C, Strouhal M, Janier M, et al. *Treponema pallidum* 11qj subtype may correspond to a *Treponema pallidum* subsp. *endemicum* strain. Sex Transm Dis. 2016; 43: 517–8.

41. Noda AA, Grillová L, Lienhard R, Blanco O, Rodríguez I, et al. Bejel in Cuba: molecular identification of *Treponema pallidum* subsp. *endemicum* in patients diagnosed with venereal syphilis. Clin Microbiol Infect. 2018; https://doi.org/10.1016/j.cmi.2018.02.006 PMID: 29454847.

42. Ratnam S. The laboratory diagnosis of syphilis. Can J Infect Dis Med Microbiol. 2005; 16(1):45–51. Epub 2007/12/27. PMID: 18159528

5.3.3 Multiplex Mediator Displacement Loop-Mediated Isothermal Amplification for Detection of *Treponema pallidum* and *Haemophilus ducreyi*

Here we further validated the use of a gene target to detect *T. pallidum* for the application of a multiplex system to detect multiple pathogens in a single reaction.

RESEARCH

Multiplex Mediator Displacement Loop-Mediated Isothermal Amplification for Detection of *Treponema pallidum* and *Haemophilus ducreyi*

Lisa Becherer, Sascha Knauf, Michael Marks, Simone Lueert, Sieghard Frischmann, Nadine Borst, Felix von Stetten, Sibauk Bieb, Yaw Adu-Sarkodie, Kingsley Asiedu, Oriol Mitjà,[1] Mohammed Bakheit[1]

Yaws, a neglected tropical disease caused by the bacterium *Treponema pallidum* subspecies *pertenue*, manifests as ulcerative skin lesions. Nucleic acid amplification tests, like loop-mediated isothermal amplification (LAMP), are versatile tools to distinguish yaws from infections that cause similar skin lesions, primarily *Haemophilus ducreyi*. We developed a novel molecular test to simultaneously detect *T. pallidum* and *H. ducreyi* based on mediator displacement LAMP. We validated the *T. pallidum* and *H. ducreyi* LAMP (TPHD-LAMP) by testing 293 clinical samples from patients with yaws-like lesions. Compared with quantitative PCR, the TPHD-LAMP demonstrated high sensitivity and specificity for *T. pallidum* (84.7% sensitivity, 95.7% specificity) and *H. ducreyi* (91.6% sensitivity, 84.8% specificity). This novel assay provided rapid molecular confirmation of *T. pallidum* and *H. ducreyi* DNA and might be suitable for use at the point of care. TPHD-LAMP could support yaws eradication by improving access to molecular diagnostic tests at the district hospital level.

Author affiliations: University of Freiburg, Freiburg, Germany (L. Becherer, N. Borst, F. von Stetten); Georg-August-University of Goettingen, Goettingen, Germany (S. Knauf); German Primate Centre, Goettingen (S. Knauf, S. Lueert); London School of Hygiene & Tropical Medicine, London, UK (M. Marks); Hospital for Tropical Diseases, London (M. Marks); Mast Diagnostica GmbH, Reinfeld, Germany (S. Frischmann, M. Bakheit); Hahn-Schickard, Freiburg, Germany (N. Borst, F. von Stetten); Papua New Guinea Department of Health, Port Moresby, Papua New Guinea (S. Bieb); Kwame Nkrumah University of Science and Technology, Kumasi, Ghana (Y. Adu-Sarkodie); World Health Organization, Geneva, Switzerland (K. Asiedu); Lihir Medical Centre-International SOS, Newcrest Mining, Lihir Island, Papua New Guinea (O. Mitjà); Fundacio Lluita contra la Sida-Hospital Universitari Germans Trias i Pujol, Badalona, Spain (O. Mitjà); Barcelona Institute for Global Health-Hospital Clinic-University of Barcelona, Barcelona, Spain (O. Mitjà)

DOI: https://doi.org/10.3201/eid2602.190505

Yaws, a neglected tropical disease caused by the bacterium *Treponema pallidum* subspecies *pertenue*, predominantly affects children living in low-income, rural communities of warm and humid regions (*1*). Clinical manifestations include lesions of the skin, bone, and cartilage, progressing to severe destructive lesions if left untreated (*2*). Manifestations of primary yaws include papillomas or ulcerative lesions; manifestations of secondary yaws include a wide range of rashes, often accompanied by bone and joint involvement (*2*). Currently, 15 countries in West and Central Africa, Southeast Asia, and the Pacific region are known to be yaws-endemic. The World Health Organization (WHO) released a yaws eradication strategy (the Morges strategy) in 2012 (*3*). The mainstay of the strategy is mass drug administration (MDA) with single-dose azithromycin in yaws-endemic communities, followed by routine surveillance and retreatment for 3–6 months until no cases remain (*3*).

Serologic tests, including the *T. pallidum* particle agglutination and rapid plasma reagin tests, remain the primary diagnostic tools for yaws (*2*). Newer point-of-care serologic tests have replaced traditional laboratory-based serologic assays in many settings (*4*–*7*). Despite their central role in yaws diagnosis, serologic assays have several limitations. First, treponemal serologic assays usually remain positive over a patient's lifetime, and these tests cannot distinguish previous from current infection. Second, studies in Africa and in countries in the Pacific region have demonstrated that *Haemophilus ducreyi* causes cutaneous lesions similar to those observed in yaws (*8*–*11*). Persons with clinically suspicious lesions caused by

[1]These senior authors contributed equally to this article.

H. ducreyi can have a reactive serologic test for yaws because of latent *T. pallidum* infection. Nucleic acid amplification tests (NAATs) can distinguish active yaws, involving a lesion with detectable *T. pallidum* DNA, from latent yaws, in which patients will have reactive serology without detectectable *T. pallidum* DNA from lesions. In addition, before seroconversion, a small proportion of patients with early active yaws will have a positive NAAT but negative serologic results.

NAATs could play a central role in yaws eradication efforts, particularly for diagnosis and surveillance after MDA in yaws-endemic areas (*12*). PCR has been standard for molecular diagnosis and has a high specificity and sensitivity for *T. pallidum* and *H. ducreyi*, but the process is time-consuming and requires expensive laboratory equipment. Most yaws-endemic countries have limited access to PCR to aid national yaws eradication programs. A point-of-care NAAT could provide reliable post-MDA molecular surveillance, as well as help in monitoring for azithromycin resistance. Loop-mediated isothermal amplification (LAMP) is an alternative for molecular diagnosis that might be more suitable than PCR as a point-of-care NAAT in resource-limited environments. LAMP has fast processing times and high specificity and can be performed on less expensive devices than those needed for PCR.

Multiplex technologies, such as mediator displacement (MD) LAMP (*13*), have extended the usability of LAMP for simultaneous detection of >1 target and could be an efficient and cost-effective solution. MD detection uses an MD probe composed of a generic mediator attached to a generic overhang of a DNA target-specific sequence and a universal reporter molecule with a fluorophore and quencher for detection. We developed and validated a biplex MD LAMP assay to simultaneously identify *T. pallidum* and *H. ducreyi*.

Methods

Participants

We obtained samples from larger trials conducted on Lihir Island (n = 57) and Karkar Island (n = 184), Papua New Guinea; and in Ghana (n = 52). Details of the studies in which the samples were collected are provided elsewhere (*14,15*). In brief, samples were collected as part of a randomized control trial comparing azithromycin doses of 30 mg/kg against doses of 20 mg/kg to treat patients in a pilot study for yaws elimination (*14,15*). Swabs were collected from persons with yaws-like ulcers and placed in AssayAssure Multilock (Sierra Molecular, https://sierramolecular.com) transport medium, then frozen at –20°C until transported to Mast Diagnostica GmbH laboratory in Reinfeld, Germany. DNA was extracted from the samples by using innuPREP MP Basic Kit A (Analytik Jena, https://www.analytik-jena.com) according to manufacturer's instructions. Isolated DNA was kept frozen at –20°C until it was used for biplex *T. pallidum* and *H. ducreyi* LAMP (TPHD-LAMP), singleplex *T. pallidum* and *H. ducreyi* LAMP assays, and quantitative PCR (qPCR) testing.

Ethics Approval

Participants, or parents or guardians of persons <18 years of age, provided written consent for inclusion in clinical surveys and etiologic studies. Children also provided assent when appropriate. The studies were approved by the National Medical Research Advisory Committee of the Papua New Guinea Ministry of Health (MRAC nos. 12.36 and 14.31), the Ghana Health Service (approval no. GHS 13/11/14), the London School of Hygiene & Tropical Medicine (approval no. 8832), and WHO (approval no. RPC720).

TPHD-LAMP Assay

We devised the TPHD-LAMP assay on the basis of 2 previously published assays: a singleplex LAMP assay (*16*), which we modified by adding an MD probe; and a biplex LAMP assay of *T. pallidum* and *H. ducreyi* (*13*). TPHD-LAMP primers target the polymerase I (polA) gene of *T. pallidum* and the 16S ribosomal RNA (16S rRNA) of *H. ducreyi*. We further optimized the assays for improved functionality by redesigning primers and probes and modifying reagent concentrations (Appendix Tables 1–3, https://wwwnc.cdc.gov/EID/article/26/2/19-0505-App1.pdf).

We performed a 2-step validation of the TPHD-LAMP assay. In the first step, we assessed the analytical sensitivity and specificity of the assay. In the second step, we used clinical samples collected in Ghana and Papua New Guinea to compare the performance of TPHD-LAMP against qPCR for individual targets. In a secondary analysis, we compared the performance of singleplex LAMP assays for each individual target against qPCR assays.

Assessment of Analytical Performance

We determined the analytical limit of detection (LOD) for the TPHD-LAMP assay by using target sequences cloned into plasmids. We determined the LOD of each of the 2 components separately, as well as the LOD of the biplex TPHD-LAMP assay (Appendix). We varied the plasmid DNA concentrations between 3×10^1 copies/reaction and 3×10^5 copies/reaction in 8 replicates to reproduce the *Treponema*

bacterial load in skin infections, which ranges from 10^2–10^4 copies/reaction (*17*). In addition, we tested the TPHD-LAMP in the presence of a high number of copies, 3×10^5 copies/reaction, of *H. ducreyi* or *T. pallidum* in the presenece of a low number of copies of the second target to optimize each component and to simulate clinical samples that might contain both targets. We conducted primer titration experiments to minimize the preferential amplification of *H. ducreyi* DNA targets in persons with both infections. We estimated the LOD by counting the fraction of positive amplifications and performed probit regression analysis by using SPSS Statistics 25 (IBM, https://www.ibm.com).

We assessed the analytical specificity of the primer sets in silico by using ortholog target gene sequences from GenBank (Appendix Table 4) and found all primer sets were highly specific for *T. pallidum* and *H. ducreyi*. Based on these results, we tested the specificity of TPHD-LAMP in vitro against endemic pathogens associated with cutaneous ulcerative syndromes by using a panel of 13 organisms: *Escherichia coli, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter cloacae, Salmonella enterica (*Paratyphi and Typhi*), Streptococcus pneumoniae, Streptococcus pyogenes, Staphylococcus aureus, Corynebacterium diphtheria, Corynebacterium ulcerans, Proteus mirabilis,* and *Enterococcus faecalis* (Appendix). We calculated interassay and intraassay variability of the TPHD-LAMP assay by using 3 batches of the TPHD-LAMP mix, prepared individually on 3 separate days and processed in different runs of 3 replicates per batch (Appendix).

Clinical Performance of the TPHD-LAMP

We performed clinical validation by comparing the performance of the TPHD-LAMP and qPCR assays to identify *T. pallidum* and *H. ducreyi* in patient samples collected in Ghana and Papua New Guinea. TPHD-LAMP reactions (10 μL per assay) were composed of 1× RM MPM buffer (MAST Diagnostica GmbH, https://mast-group.com), 8 U Bst 2.0 WarmStart DNA Polymerase (New England Biolabs, https://www.neb.com), 0.05 μmol/L universal reporter, and MD primer mix (Appendix). We incubated primer mixes for 5 m at 70°C before LAMP to prevent nonspecific amplification initiated by primer dimerization. We performed real-time TPHD-LAMP reactions at 64°C in a Rotor-Gene Q (QIAGEN, https://www.qiagen.com) and acquired fluorescence signals every minute by using the Cy5-readout gain for *T. pallidum* and the FAM-readout gain for *H. ducreyi*. The singleplex LAMP reactions (10 μL per assay) using intercalating dye were composed of 1× RM MPM buffer, 8 U Bst 2.0 WarmStart DNA Polymerase, and 1 μL of 10× SYBR Green staining reagent, DNA free (AppliChem, https://www.applichem.com) and primer mix (Appendix Table 1). We also performed singleplex LAMP reactions in a Rotor-Gene Q at 63°C with the FAM-readout gain. We used a cutoff of 60 m for biplex TPHD-LAMP and singleplex LAMP assays and considered samples with amplification beyond 60 m negative.

For performance analyses, we compared the TPHD-LAMP assay against TaqMan qPCR assays targeting *polA* of *T. pallidum* (*18*) and an optimized TaqMan qPCR assay targeting the 16S rRNA gene of *H. ducreyi* on the same DNA extract (Appendix Table 4, Figure 1). The 16S rRNA gene has been previously used in qPCR assays to detect *H. ducreyi* (*19*). We ran all tests in duplicate and included positive controls and DNA-free negative controls in each run. We used an identical sample volume, 2.5 μL/reaction, for TPHD-LAMP, singleplex LAMP, and qPCR. For samples that tested negative by qPCR but positive by TPHD-LAMP, we repeated qPCR in a single reaction with higher sample volumes (3 μL) to identify true negative test results.

Statistical Analysis

For clinical validation, we compared the sensitivity and specificity of the TPHD-LAMP assay against TaqMan qPCR assays. In a secondary analysis, we compared the performance of singleplex LAMP assays to qPCR. We performed all analysis by using R version 3.4.3 (https://www.R-project.org).

Results

Analytical Sensitivity and Specificity

The LOD for the TPHD-LAMP assay was 357 copies/reaction (95% CI 265–535 copies/reaction) for *T. pallidum* and 293 copies/reaction (95% CI 199–490 copies/reaction) for *H. ducreyi*. When we added the second target at the higher concentration of 3×10^5 copies/reaction to simulate clinical samples from persons infected with both bacteria, the LOD increased to 808 copies/reaction (95% CI 550–2,128 copies/reaction) for *T. pallidum* and 622 copies/reaction (95% CI 415–1,687 copies/reaction) for *H. ducreyi* (Appendix Figure 2). The TPHD-LAMP assay was negative for all other pathogens tested within 60 m, demonstrating high analytical specificity (Appendix Figure 3). We observed a minimal interassay or intraassay variation (Appendix Figure 4).

Validation of TPHD-LAMP in Clinical Samples

For clinical validation, we used a sample set consisting of 293 lesion swabs collected from patients with suspected *T. pallidum* infection. Samples were collected in Lihir Island (n = 57; 19.5%) and Karkar Island (n = 184; 62.8%), Papua New Guinea; and in Ghana (n = 52; 17.7%). A total of 184 (62.8%) cases were in male patients and 109 (37.2%) in female patients; the median age of case-patients was 10 years (interquartile range [IQR] 8–12 years).

Using qPCR, we detected *T. pallidum* in 59 (20.1%) samples, *H. ducreyi* in 155 (52.9%) samples, and *T. pallidum* and *H. ducreyi* co-infection in 19 (6.5%) samples. When tested by TPHD-LAMP, we detected *T. pallidum* in 60 (20.5%) samples and *H. ducreyi* in 163 (55.6%) samples. We detected both targets in 12 (4.1%) samples. Taking qPCR as the reference standard, the diagnostic sensitivity of the TPHD-LAMP assay for *T. pallidum* was 84.7% and the specificity was 95.7%. For *H. ducreyi*, the sensitivity of the TPHD-LAMP assay was 91.6% and the specificity was 84.8% (Table 1). Kappa coefficients (κ), ranging from 0.7 to 0.9 for the detection of *T. pallidum* and from 0.7 to 0.8 for *H. ducreyi*, show substantial to excellent agreement between qPCR and TPHD-LAMP. Moderate agreement between qPCR and TPHD-LAMP (κ = 0.5) also was demonstrated for the simultaneous detection of both targets. The median time to amplification of *T. pallidum* was 11 min (IQR 9–15 min) and the median time to amplification of *H. ducreyi* was 10 min (IQR 8–24 min).

For samples in which only 1 organism was detected by qPCR, the sensitivity of the TPHD-LAMP assay was higher for both *T. pallidum* (92.5%) and *H. ducreyi* (94.1%) than for samples with both organisms confirmed by qPCR. For samples confirmed to contain both bacteria by qPCR, sensitivity for *T. pallidum* was 68.4% (p = 0.048) and sensitivity for *H. ducreyi* was 73.7% (p = 0.01) (Table 1).

Using qPCR as the reference standard, the singleplex *T. pallidum* LAMP assay had a sensitivity of 78.0% and specificity of 97.9%; for the singleplex *H. ducreyi* LAMP assay the sensitivity was 91.0% and specificity was 75.3% (Table 2). We did not see a noticeable variation in the performance of the biplex TPHD-LAMP and singleplex LAMP assays between locations from which samples were collected (Tables 1 and 2).

Discussion

We provide data demonstrating a high analytical performance of a multiplex LAMP assay for *T. pallidum* and *H. ducreyi* and a high sensitivity and specificity comparable to qPCR. The TPHD-LAMP assay also performed better than singleplex LAMP assays, likely reflecting better performance of the MD technology used in the biplex LAMP compared with standard intercalating dyes used in singleplex LAMP assays.

The LOD of the TPHD-LAMP assay was 300 copies/reaction for both targets, which is comparable to qPCR, which has standard reproducibility in a range

Table 1. Comparison of clinical performance of biplex loop-mediated isothermal amplification for detection of *Treponema pallidum* and *Haemophilus ducreyi* (TPHD-LAMP) against singleplex TaqMan quantitative PCR*

Characteristics	Sample size	*Treponema pallidum*	*Haemophilus ducreyi*
Total samples, no.	293		
No. positive		60	163
Sensitivity, % (95% CI)		84.7 (72.5–92.4)	91.6 (85.8–95.3)
Specificity, % (95% CI)		95.7 (92.0–97.8)	84.8 (77.4–90.1)
Lesions containing a single pathogen†	195		
No. positive		48	151
Sensitivity, % (95% CI)		92.5 (78.5–98.0)	94.1 (88.4–97.2)
Specificity, % (95% CI)		95.7 (92.0–97.8)	84.8 (77.4–90.1)
Lesions containing both pathogens†	19		
No. positive		12	12
Sensitivity, % (95% CI)		68.4 (43.5–86.4)	73.7 (48.6–89.9)
Specificity, % (95% CI)		NA	NA
Samples from Lihir Island, no.	57		
No. positive		21	13
Sensitivity, % (95% CI)		90.5 (68.2–98.3)	76.5 (50.0–92.2)
Specificity, % (95% CI)		94.4 (80.0–99.0)	100.0 (89.1–100)
Samples from Karkar Island, no.	184		
No. positive		33	119
Sensitivity, % (95% CI)		78.1 (59.6–90.1)	94.2 (87.5–97.7)
Specificity, % (95% CI)		94.7 (89.5–97.5)	74.7 (63.4–83.5)
Samples from Ghana, no.	52		
No. positive		6	31
Sensitivity, % (95% CI)		100.0 (51.7–100)	90.9 (75.5–97.6)
Specificity, % (95% CI)		100.0 (90.4–100)	94.7 (71.9–99.7)

*NA, not applicable.
†Determined by quantitative PCR.

Table 2. Comparison of clinical performance of singleplex loop-mediated isothermal amplification for detection of *Treponema pallidum* and *Haemophilus ducreyi* against singleplex TaqMan quantitative PCR*

Characteristics	Sample size	*Treponema pallidum*	*Haemophilus ducreyi*
Total samples, no.	293		
No. positive		51	175
Sensitivity, % (95% CI)		78.0 (64.9–87.3)	91.0 (85.0–94.8)
Specificity, % (95% CI)		97.9 (94.8–99.2)	75.3 (67.2–82.1)
Lesions containing a single pathogen†	195		
No. positive		34	158
Sensitivity, % (95% CI)		82.5 (66.6–92.1)	92.6 (86.5–96.2)
Specificity, % (95% CI)		97.9 (94.8–99.2)	75.4 (67.2–82.1)
Lesions containing both pathogens†	19		
No. positive samples		17	17
Sensitivity, % (95% CI)		68.4 (43.5–86.4)	78.9 (53.9–93.0)
Specificity, % (95% CI)		NA	NA

*NA, not applicable.
†Determined by quantitative PCR.

of 10^1–10^6 copies/reaction. The LOD increased to ≈600 copies/reaction in samples that contained both targets, which is consistent with our clinical validation of the TPHD-LAMP; sensitivity for both bacteria was slightly higher when samples contained only a single target. Kappa coefficients confirmed substantial agreement (κ>0.7) for the individual targets and moderate agreement (κ = 0.5) for simultaneous detection of both targets in a sample.

Detection of *T. pallidum* is the programmatic priority, but detection of *H. ducreyi* is beneficial for clinical management of patients with suspected yaws. The median time to amplification was <15 m for both *T. pallidum* and *H. ducreyi*, indicating the TPHD-LAMP assay could provide rapid, molecular confirmation of the presence of *T. pallidum* or *H. ducreyi*. Further optimization of the assay to enhance the performance of the *T. pallidum* component, particularly in the context of co-infection, will be required to ensure cases of yaws are not missed.

Implementing qPCR at the point of care is operationally challenging because it requires relatively expensive equipment, in particular thermocyclers, which can cost up to 10 times as much as a tubescanner capable of performing the TPHD-LAMP assay. Because qPCR is available only in a limited number of national and international reference laboratories, TPHD-LAMP might be an alternative molecular test to support expansion of yaws eradication activities. We did not conduct a cost-effectiveness analysis of the TPHD-LAMP assay, but such an assessment should consider equipment costs, cost per assay, and the relative performance of each assay to assess the cost per case diagnosed. However, our data suggest that the TPHD-LAMP assay might be a cost-saving alternative to qPCR, especially at the point of care.

Our study had some limitations. We tested samples from only 2 geographic regions for clinical validation of the TPHD-LAMP. Primer binding site mutations have affected the performance of other diagnostic assays for *T. pallidum* strains. Although we selected conserved genomic regions when designing the TPHD-LAMP primers, further experimental validation of the TPHD-LAMP assay with samples from a broader range of settings is needed. We conducted clinical validation of the assay in a controlled laboratory setting, but conditions at the point of care, including temperature, humidity, and a range of other environmental factors, might affect reagents in storage and in performing assays. Further optimization, including freeze-dried reagents in combination with dried oligonucleotides, might improve robustness and facilitate rollout of the assay in yaws-endemic countries.

In yaws-endemic countries, clinical manifestations combined with serologic tests are still the standard tool for the clinical management of yaws, but serologic tests have limitations and molecular assays are needed to support WHO yaws eradication efforts (*12*). Molecular assays also can detect mutations in the 23S RNA gene associated with azithromycin resistance (*15*,*20*,*21*), which is essential to monitor for drug resistance as yaws eradication efforts expand. qPCR is the most common NAAT currently available but remains restricted to a small number of laboratories in yaws-endemic countries. MD LAMP could facilitate surveillance for resistance and we plan further studies to evaluate a modified TPHD-LAMP assay for this purpose. Further, multicountry evaluations are warranted to assess performance of the assay when deployed in yaws-endemic countries and to assess the role the test could play in support of national yaws eradication programs. Nonetheless, the performance characteristics of the TPHD-LAMP suggest it has the potential to increase access to molecular diagnosis of yaws, especially at the point of care.

Acknowledgments

We thank members of the study teams and communities who participated in the field studies from which we obtained samples.

Financial support: The trial conducted in Ghana and Karkar Island, Papua New Guinea, was funded by a grant from the Neglected Tropical Diseases Support Center to WHO (no. NTD-SC/NCT 053). M.M. was supported by the Wellcome Trust under grant no. 102807. The study was partially funded by a grant from the German Research Foundation (no. KN 1097/3-2) to S.K. This work was partially funded by a grant from the German Federal Ministry of Education and Research (EuroTransBio no. 031B0132B) to L.B. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions, or policies of the institutions with which they are affiliated.

Author contributions: L.B. conducted laboratory work, analyzed the data, and wrote the first draft of the manuscript. M.M. and O.M. designed the field studies and analyzed the data. S.K., S.L., S.F., and N.B. contributed to laboratory work or analysis. S.B., Y.A.-S., and K.A. led the field studies. M.B. contributed to laboratory work and analyzed the data. All authors revised the manuscript.

Potential conflicts of interest: S.F. and M.B. are employees of Mast Diagnostica GmbH, which produces and sells LAMP kits and products. A patent covering the technique described in the paper has been applied for by the University of Freiburg, Freiburg, Germany, and Hahn-Schickard, Villingen-Schwenningen, Germany.

About the Author

Ms. Becherer is a scientific researcher in the nucleic acid analysis group at the Department of Microsystems Engineering, University of Freiburg, Freiburg, Germany. Her work focuses on the development of novel methods for nucleic acid amplification, with a special focus on isothermal amplification.

References

1. Marks M, Solomon AW, Mabey DC. Endemic treponemal diseases. Trans R Soc Trop Med Hyg. 2014;108:601–7. https://doi.org/10.1093/trstmh/tru128
2. Mitjà O, Asiedu K, Mabey D. Yaws. Lancet. 2013;381:763–73. https://doi.org/10.1016/S0140-6736(12)62130-8
3. The World Health Organization. Eradication of yaws—the Morges strategy. Wkly Epidemiol Rec. 2012;87:189–94.
4. Ayove T, Houniei W, Wangnapi R, Bieb SV, Kazadi W, Luke L-N, et al. Sensitivity and specificity of a rapid point-of-care test for active yaws: a comparative study. Lancet Glob Health. 2014;2:e415–21. https://doi.org/10.1016/S2214-109X(14)70231-1
5. Marks M, Goncalves A, Vahi V, Sokana O, Puiahi E, Zhang Z, et al. Evaluation of a rapid diagnostic test for yaws infection in a community surveillance setting. PLoS Negl Trop Dis. 2014;8:e3156. https://doi.org/10.1371/journal.pntd.0003156
6. Fitzpatrick C, Asiedu K, Sands A, Gonzalez Pena T, Marks M, Mitja O, et al. The cost and cost-effectiveness of rapid testing strategies for yaws diagnosis and surveillance. PLoS Negl Trop Dis. 2017;11:e0005985. https://doi.org/10.1371/journal.pntd.0005985
7. Marks M, Yin Y-P, Chen X-S, Castro A, Causer L, Guy R, et al. Metaanalysis of the performance of a combined treponemal and nontreponemal rapid diagnostic test for syphilis and yaws. Clin Infect Dis. 2016;63:627–33. https://doi.org/10.1093/cid/ciw348
8. González-Beiras C, Marks M, Chen CY, Roberts S, Mitjà O. Epidemiology of *Haemophilus ducreyi* infections. Emerg Infect Dis. 2016;22:1–8. https://doi.org/10.3201/eid2201.150425
9. Marks M, Chi K-H, Vahi V, Pillay A, Sokana O, Pavluck A, et al. *Haemophilus ducreyi* associated with skin ulcers among children, Solomon Islands. Emerg Infect Dis. 2014;20:1705–7. https://doi.org/10.3201/eid2010.140573
10. Mitjà O, Lukehart SA, Pokowas G, Moses P, Kapa A, Godornes C, et al. *Haemophilus ducreyi* as a cause of skin ulcers in children from a yaws-endemic area of Papua New Guinea: a prospective cohort study. Lancet Glob Health. 2014;2:e235–41. https://doi.org/10.1016/S2214-109X(14)70019-1
11. Ghinai R, El-Duah P, Chi K-H, Pillay A, Solomon AW, Bailey RL, et al. A cross-sectional study of 'yaws' in districts of Ghana which have previously undertaken azithromycin mass drug administration for trachoma control. PLoS Negl Trop Dis. 2015;9:e0003496. https://doi.org/10.1371/journal.pntd.0003496
12. Marks M, Mitjà O, Vestergaard LS, Pillay A, Knauf S, Chen C-Y, et al. Challenges and key research questions for yaws eradication. Lancet Infect Dis. 2015;15:1220–5. https://doi.org/10.1016/S1473-3099(15)00136-X
13. Becherer L, Bakheit M, Frischmann S, Stinco S, Borst N, Zengerle R, et al. Simplified real-time multiplex detection of loop-mediated isothermal amplification using novel mediator displacement probes with universal reporters. Anal Chem. 2018;90:4741–8. https://doi.org/10.1021/acs.analchem.7b05371
14. Marks M, Mitjà O, Bottomley C, Kwakye C, Houinei W, Bauri M, et al.; study team. Comparative efficacy of low-dose versus standard-dose azithromycin for patients with yaws: a randomised non-inferiority trial in Ghana and Papua New Guinea. Lancet Glob Health. 2018;6:e401–10. https://doi.org/10.1016/S2214-109X(18)30023-8
15. Mitjà O, Godornes C, Houinei W, Kapa A, Paru R, Abel H, et al. Re-emergence of yaws after single mass azithromycin treatment followed by targeted treatment: a longitudinal study. Lancet. 2018;391:1599–607. https://doi.org/10.1016/S0140-6736(18)30204-6
16. Knauf S, Lüert S, Šmajs D, Strouhal M, Chuma IS, Frischmann S, et al. Gene target selection for loop-mediated isothermal amplification for rapid discrimination of *Treponema pallidum* subspecies. PLoS Negl Trop Dis. 2018;12:e0006396. https://doi.org/10.1371/journal.pntd.0006396
17. Tipple C, Hanna MOF, Hill S, Daniel J, Goldmeier D, McClure MO, et al. Getting the measure of syphilis: qPCR to better understand early infection. Sex Transm Infect. 2011;87:479–85. https://doi.org/10.1136/sti.2011.049494
18. Chen C-Y, Chi K-H, George RW, Cox DL, Srivastava A, Rui Silva M, et al. Diagnosis of gastric syphilis by direct

immunofluorescence staining and real-time PCR testing. J Clin Microbiol. 2006;44:3452–6. https://doi.org/10.1128/JCM.00721-06
19. Orle KA, Gates CA, Martin DH, Body BA, Weiss JB. Simultaneous PCR detection of *Haemophilus ducreyi, Treponema pallidum*, and herpes simplex virus types 1 and 2 from genital ulcers. J Clin Microbiol. 1996;34:49–54.
20. Chen C-Y, Chi K-H, Pillay A, Nachamkin E, Su JR, Ballard RC. Detection of the A2058G and A2059G 23S rRNA gene point mutations associated with azithromycin resistance in *Treponema pallidum* by use of a TaqMan real-time multiplex PCR assay. J Clin Microbiol. 2013;51:908–13. https://doi.org/10.1128/JCM.02770-12
21. Lukehart SA, Godornes C, Molini BJ, Sonnett P, Hopkins S, Mulcahy F, et al. Macrolide resistance in *Treponema pallidum* in the United States and Ireland. N Engl J Med. 2004;351:154–8. https://doi.org/10.1056/NEJMoa040216

Address for correspondence: Mohammed Bakheit, Mast Diagnostica GmbH, Feldstraße 20, 23858 Reinfeld, Germany; email: bakheit@mast-diagnostica.de

5.3.4 The Impact of Storage Buffer, DNA Extraction Method, and Polymerase on Microbial Analysis

HTS approaches for the characterisation of microbial communities are associated with technical caveats that can significantly impact the research outcome. The following publication describes the optimal sample handling protocol to minimize the bias for different sample types. The use of a standardized microbial mock-community is particularly noteworthy as it introduced a high level of standardization. The outcome of this study contributed significantly to the optimisation of sampling and methodology protocols to investigate disease reservoirs in wildlife in my laboratory.

SCIENTIFIC REPORTS

OPEN

The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis

Received: 30 November 2017
Accepted: 6 April 2018
Published online: 19 April 2018

Luisa K. Hallmaier-Wacker[1,2], Simone Lueert[1,2], Christian Roos[2] & Sascha Knauf[1]

Next-generation sequencing approaches used to characterize microbial communities are subject to technical caveats that can lead to major distortion of acquired data. Determining the optimal sample handling protocol is essential to minimize the bias for different sample types. Using a mock community composed of 22 bacterial strains of even concentration, we studied a combination of handling conditions to determine the optimal conditions for swab material. Examining a combination of effects simulates the reality of handling environmental samples and may thus provide a better foundation for the standardization of protocols. We found that the choice of storage buffer and extraction kit affects the detected bacterial composition, while different 16S rRNA amplification methods only had a minor effect. All bacterial genera present in the mock community were identified with minimal levels of contamination independent of the choice of sample processing. Despite this, the observed bacterial profile for all tested conditions were significantly different from the expected abundance. This highlights the need for proper validation and standardization for each sample type using a mock community and blank control samples, to assess the bias in the protocol and reduce variation across the datasets.

Microorganisms colonize various anatomical sites and play a crucial role in the balance of health and disease. The vaginal microbiome is known to maintain the health of women and thereby prevents urogenital diseases[1]. The advent of cultivation-independent molecular approaches, such as 16S rRNA amplicon sequencing, has allowed for a better understanding of the microbes that inhabit different biological niches. However, these powerful tools are not without important technical caveats that can lead to a distortion in the acquired data[2]. Such limitations have been well documented, and include sample collection, storage buffer, DNA extraction, amplification primers and methods, sequencing technology, and analysis techniques[3,4]. While it is impossible to negate all of these influences, it is important to understand the bias inherent in the analysis. Studies focusing on one or two technical limitations have made recommendations for improving the bias such as reducing the number of PCR cycles[5] or adding additional lysis pre-treatment[6].

DNA extraction, a critical step in culture-independent bacterial profiling, has been identified as a key driver of technical variation[3]. Most common studies on the microbiome of swab material use commercially available DNA extraction kits that vary in their lysis approach from mechanical to enzymatic treatment. Various studies have focused on technical variations in extraction kits, yet a field-wide consensus on sample extraction has not been reached[3,6–9]. Due to the large variety of microbiota and sample types, a single standard for all sample types is unlikely to be achieved. Despite the knowledge that the choice of extraction kit can have a significant effect on the results, there is often a lack of proper validation across sample types[3].

Similar to DNA extraction kits, the choice of sample storage buffer has been shown to influence the detected bacterial community[10–12]. The ideal storage choice largely depends on the available resources during sampling such as the availability of freezing conditions[11]. Selecting the optimal storage buffers is dependent upon its compatibility with all downstream analyses including the extraction method. Many studies, however, only focus on the effect of a single technical variation instead of examining the effect of different combination of storage buffer, DNA extraction kit, and amplification methods[2]. Studying a combination of effects mirrors the reality of sample handling more closely and may thus provide a better foundation for the standardization of sampling handling protocols prior to microbial analysis.

[1]Work Group Neglected Tropical Diseases, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077, Göttingen, Germany. [2]Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077, Göttingen, Germany. Correspondence and requests for materials should be addressed to S.K. (email: sknauf@dpz.eu)

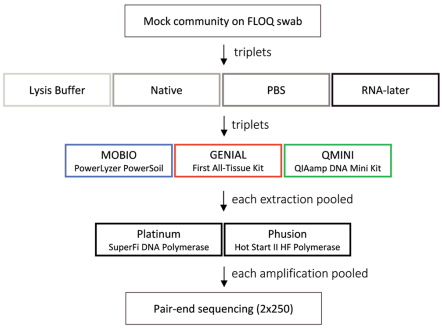


Figure 1. Outline of experimental design. A schematic showing the different treatment variables.

Extraction Method	Abbreviation	Lot #	Lysis type	Elution Volume
MOBIO PowerLyzer PowerSoil Kit	MOBIO	PL16C30	Mechanical, Column-based	100
GEN-IAL First All-Tissue Kit	GENIAL	0091.01	Enzymatic, Phenol-Chloroform	20
QIAamp DNA Mini Kit	QMINI	154035749	Enzymatic, Column-based	70

Table 1. Commercial extraction kits used in this study.

In this study, we used a mock community, composed of an even concentration of cells from 22 bacterial strains (19 genera), to assess the effect of storage buffers, extraction kits, and amplification methods (Fig. 1). Using a mock community to examine the effect of different sample handling conditions rather than environmental samples of unknown microbe composition is essential to be able to systematically compare the effects[3]. In addition to the use of a mock community, a blank control was included in all sample procedures to monitor any buffer, kit, or reagent specific contamination[13]. The aim of this study was to evaluate the performance of combinations of handling conditions commonly used in microbiome studies and to contribute to the ongoing debate on standardization in microbiome research.

Methods

Preparation of swab mock community samples. A cell mixture of 22 different bacterial strains at a concentration of 1×10^8 cells/mL of each organism (Microbial mock community, HM-280) in phosphate buffer saline (PBS) was obtained through Biodefense and Emerging Infectious Research (BEI) Resources, NIAID, NIH as part of the Human Microbiome Project (Manassas, USA; Supplementary Table S1). To simulate physiological conditions, 10 µl of mock community containing 1×10^6 cells/mL of each organism was added to a flocked swab (FLOQSwabs, Copan Improve Diagnostics, Brescia, Italy) and then placed in 500 µl of the respective storage buffer (Fig. 1). Four different storage buffers were used; PBS (PAN-Biotech GmbH, Aidenbach, Germany), a custom-made lysis buffer (10 mM Tris, pH 8.0, 0.1 M EDTA, pH 8.0 and 0.5% SDS), RNA-later (Thermo Fisher Scientific Inc., Waltham, MA, USA), and no buffer (native). A blank control swab sample was placed in each storage buffer without additive. All swab samples were frozen at −80 °C for one week prior to DNA extraction. Suitable precautions were taken during sample handling and processing to insure sterility during all procedures.

DNA Extraction methods. Three commercially available DNA extraction kits were used in this study to extract bacterial DNA from swab material stored in four different storage buffers (Table 1). Extraction was performed in triplets and the extracted DNA from each buffer was subsequently pooled prior to 16 S rRNA gene amplification. Processing of swab samples prior to DNA extraction is illustrated in Supplementary Fig. S1.

QIAamp DNA Mini Kit (QMINI). Samples were extracted using the QIAamp Mini Kit (Qiagen GmbH, Hilden, Germany) according to the standard protocol with minor modifications. Briefly, proteinase K (20 mg/µl) was added and the samples were incubated for 50 minutes at 56 °C. Then, AL buffer (Qiagen GmbH) and ethanol were added in the appropriate amount. The DNA from the lysate was subsequently purified using the spin columns provided by the manufacturer and eluted in 70 µl AVE buffer (Qiagen GmbH).

MOBIO PowerLyzer PowerSoil Kit (MOBIO). A maximum of 750 µl of swab lysate was added to the 0.1 mm PowerLyzer® Glass Bead Tube (Qiagen GmbH). DNA extraction was continued from step 2 as described in the MOBIO PowerLyzer PowerSoil Kit protocol (Qiagen GmbH). The DNA was eluted in a final volume of 100 µl of Solution C6 provided in the kit.

GEN-IAL First All-Tissue Kit (GENIAL). The first All-Tissue Kit (GEN-IAL, Troisdorf, Germany) was applied according to the manufacturer's protocol with minor modifications. Briefly, 5 µl proteinase K and 5 µl dithiothreitol (DTT) was added to the lysate and incubated at 65 °C for 60 min at 600 rpm in a thermomixer (Eppendorf, Hamburg, Germany). The lysate was purified according to the standard protocol and the DNA pellet was resuspended in 20 µl of C6 buffer (Qiagen GmbH).

16S rRNA gene amplification. For each pooled extraction, the V4 region of the 16 S ribosomal RNA (16 S rRNA) gene was amplified in triplets using the universal primers 515 F and 806 R adapted with linker regions and barcoded sequences used for dual-indexing[14]. Platinum SuperFi DNA Polymerase (Thermo Fisher Scientific) and the Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific) were both tested for amplification. Each PCR reaction consisted of 12.5 µl of 2x PCR master mix, 6 µl of Microbial DNA-Free water (Qiagen GmbH), 1.25 µl of each primer (0.5 mM each, Metabion, Martinsried, Germany) and 4 µl of template in a total reaction volume of 25 µl. PCR cycling conditions comprised of a pre-denaturation step of 30 s at 98 °C, followed by 30 cycles of 98 °C for 10 s, 55 °C for 15 s and 72 °C for 60 s, and a final 10 min extension step at 72 °C. For a selection of four samples, five additional cycles were added to the amplification procedure to examine if additional cycles may be favorable for samples with low concentrations. The amplicon triplets were pooled, purified using 0.7x AMPure XP beads (Beckman Coulter, Brea, USA) and quantified using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific). Amplicon integrity was verified for a representative number of 11 samples using a BioAnalyzer 2000 (Agilent, Palo Alto, USA) prior to pooling equimolar amounts (10 nM) of each amplicon for sequencing. For the blank samples, the maximum volume (5 µl) of sample was added to the library, as the concentrations prior to sequencing were below 10 nM. Illumina MiSeq. 2 × 250 bp paired-end sequencing (Illumina V2 chemistry) was performed in the Transcriptome and Genome Analysis Laboratory at the University of Göttingen[14]. All generated read files analyzed in this study were uploaded to the NCBI Sequence Read Archive (SRA) (SRP125723).

Mock community data processing and analysis. The sequencing reads were processed using the mothur software package (v.1.36.1)[15]. According to the MiSeq SOP[14], contigs were assembled, sequences trimmed, identical sequences merged, and chimeras removed (UCHIME[16]). Subsequently, sequences were aligned to the SILVA bacterial reference database[17]. Non-bacterial sequences, cross-sample singletons, and poorly aligned sequences were removed. The seq.error command was run for each mock sample in mothur and subsequently averaged to determine the error rate of the run. Due to low read numbers, blank control sample reads (control swabs containing no mock community) were removed from the dataset and analyzed separately. As subsampling is currently still an accepted method of normalization in microbial ecology[18], the reads of the remaining mock community samples were rarefied to 95,870 sequences/sample. A separate file with the theoretical sequence composition (actual) of the 22 bacterial strains of mock community was created and adjusted for the 16 S rRNA copy number (Supplementary Table S1) and normalized to the sequence count of the run (95,870 reads)[19]. After merging the actual (theoretical) mock community composition with the practically obtained sequences, the merged file was classified using the Bayesian classifier implemented in mothur[20]. Operational taxonomic units (OTUs) were assigned based on 97% sequence similarity and subsequently the alpha and beta diversity was analyzed. For alpha diversity, the richness (OTUs observed and Chao1) and community diversity (Inverse Simpson Metrix) was analyzed using the summary.single command in mothur. Additionally, the percentage of contaminant OTUs (OTUs that do not cluster to the theoretical mock community) was examined. Beta diversity was analyzed using Bray-Curtis dissimilarity index[21]. The dissimilarity matrix was visualized using nonmetric multidimensional scaling (NMDS) plots and Newick formatted dendrograms (visualized in FigTree v.1.4.2, http://tree.bio.ed.ac.uk/software/figtree/).

Statistical comparison of sequence data. To evaluate and compare the type of extraction and amplification method used, the values of the alpha or beta diversity measurement were pooled for each variable (e.g. the buffer type). The statistical significance of the pooled data was analyzed in GraphPad Prism 6 (GraphPad software, La Jolla, CA, USA). In case of normal distribution (Kolmogorov-Smirnov normality test), the parametric paired two-tailed students t-test was used for comparison. In all other cases the non-parametric Wilcoxon matched-pairs signed rank test was used. For multiple comparisons, a one-way ANOVA with Bonferroni's multiple comparisons test was applied. Differences in community structure between storage buffers and extraction methods were tested using analysis of molecular variance (AMOVA) in mothur[22]. Non-metric multidimensional scaling (NMDS) plot of Bray-Curtis dissimilarities and UPGMA-clustered dendrograms (Bray-Curtis) were used to visualize data points. Parsimony (mothur) hypothesis testing was performed to test whether the differential clustering of the PBS samples in the dendrograms was significant[23]. Differences in the 30 most abundant OTUs were assessed using the metastats command in mothur[24] and p-values for differences in individual OTUs were corrected for multiple comparisons using Bonferroni correction. Values of $p < 0.05$ were considered statistically significant.

Results

The pooled library (n = 28 mock samples, n = 36 blank/control samples) produced 12,968,125 16 S rRNA sequence reads, of which 9,920,805 reads were retained after quality control (77%). A total of 8,974,393 sequences, with a mean read count of 249,288 reads per sample, were retained after the sequences corresponding to the blank control samples were removed. After rarefying to 95,870 sequences per sample, *de novo* OTU picking returned 228 OTUs, of which 19 OTUs corresponding to the mock community make up more than 99% of the pooled community. The average error rate of the run was found to be 0.040% (±0.004).

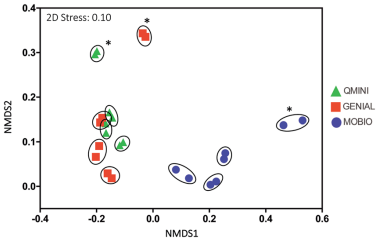


Figure 2. Clustering of samples amplified with two different polymerases on a non-metric multidimensional scaling (NMDS) plot of the Bray-Curtis dissimilarities. Points are colored by applied extraction kit. The encircled pairs correspond to a single sample where each data point represents one 16 S rRNA amplification with Phusion Hot Start II High-Fidelity DNA Polymerase and the another with Platinum SuperFi DNA Polymerase. Sample pairs labeled with * were stored in PBS.

Effect of different amplification method. The choice of polymerase (Platinum SuperFi DNA polymerase vs. Phusion Hot Start II High-Fidelity DNA polymerase) was not found to significantly change the number of observed OTUs ($p = 0.08$ [paired t-test] or Inverse Simpson index, $p = 0.48$, [paired t-test]). Furthermore, pairwise comparison of the Bray-Curtis dissimilarity between the two polymerases yielded only small variations (maximum difference 0.076, Supplementary Table S2) indicating near identical bacterial community profile for a single sample (Fig. 2). Since the results indicate that these two applied high-fidelity polymerases do not significantly impact the observed microbial diversity, we pooled the data from the two polymerases for identical sample for the analyses of buffer and extraction kit choice. The addition of five cycles in 16 S rRNA gene amplification shows only a minor impact on the detected bacterial composition when tested on MOBIO extractions (Supplementary Fig. S2a). There was, however, a significant increase of the number of OTUs detected with additional cycles ($p = 0.029$, Supplementary Fig. S2b), indicating that lower cycle numbers are favorable.

Effect of storage buffer. The effect of the four storage buffer (lysis buffer, native, PBS or RNA-later) on the alpha diversity was assessed based on OTU richness (identified absolute number of taxa) and evenness (Inverse Simpson index). The choice of storage buffer had no significant influence on the OTU richness of the swab samples ($p = 0.158$ [ANOVA], Fig. 3a), nor the overall evenness. However, PBS treated samples that were extracted with MOBIO, detected a lower evenness compared to all other treatment conditions (Wilcoxon test, Fig. 3b).

Pairwise AMOVA of Bray-Curtis dissimilarity showed that the storage buffer choice had a significant impact on the community structure ($p = 0.004$, AMOVA). A dendrogram of the Bray-Curtis dissimilarity shows that the PBS stored samples clustered separately from the other buffer types which was confirmed by parsimony analysis ($p = 0.001$, Fig. 3c). To examine which OTUs drive the differential clustering, we examined the read count for each OTU. Four bacterial OTUs corresponding to *Neisseria*, *Pseudomonas*, *Porphyromonas* and *Helicobacter* are significantly different in the PBS stored samples for all extraction kits (Fig. 3d–g). These results indicated that PBS buffer significantly alters single OTUs as well as the overall bacterial composition compared to all other storage buffers, independent of extraction kit choice. The bacterial profile of the blank control samples indicated that this effect is not caused by a buffer specific contamination as there appears to be no obvious buffer or kit specific profile (Supplementary Fig. S3).

Effect of extraction method. Richness, both the observed number of OTUs and Choa1, were analyzed to see the effect of the extraction kit choice on the alpha diversity. Pairwise comparison showed no significant effect on OTU richness between the different extraction kits ($p = 0.893$ [ANOVA], Table 2). In general, all extraction kits detect a higher OTU richness compared to the expected richness of the mock community (Table 2). In addition to assessing richness, evenness was analyzed using the Inverse Simpson index. The evenness of the samples extracted using MOBIO was significantly lower compared to the QMINI and GEN-IAL extractions ($p = 0.008$, $p = 0.023$, Wilcoxon test, Table 2). The evenness did not significantly vary between QMINI and GEN-IAL. Yet, the mean ($\pm$SEM) observed evenness (5.21 ± 0.08) was significantly lower than the expected evenness of the mock community (18.3). The same five OTUs, *Enterococcus, Neisseria, Escherichia, Pseudomonas*, and *Bacillus* dominate the bacterial profile independent of extraction kit choice (Fig. 3c).

Pairwise AMOVA of Bray-Curtis dissimilarity indicated that the extraction kit choice significantly impacted the community structure ($p = 0.001$, AMOVA). To assess which extraction kit more accurately represents the bacterial community structure, a theoretical ideal mock community (actual) composition was created for comparison (see methods for details). In the ideal scenario, the experimental data would be identical to the actual composition and there would be no Bray-Curtis dissimilarity. To assess the extraction kits, Bray-Curtis dissimilarity was calculated between the observed and actual mock community for each sample (Fig. 4). The samples extracted with the same commercial kit were grouped in a boxplot and pairwise comparison was performed. The QMINI kit produced a significantly better representation of the bacterial community compared to all other kits tested (paired t-test, all $p < 0.01$, Fig. 4). On the contrary, the MOBIO kit performed significantly poorer than all

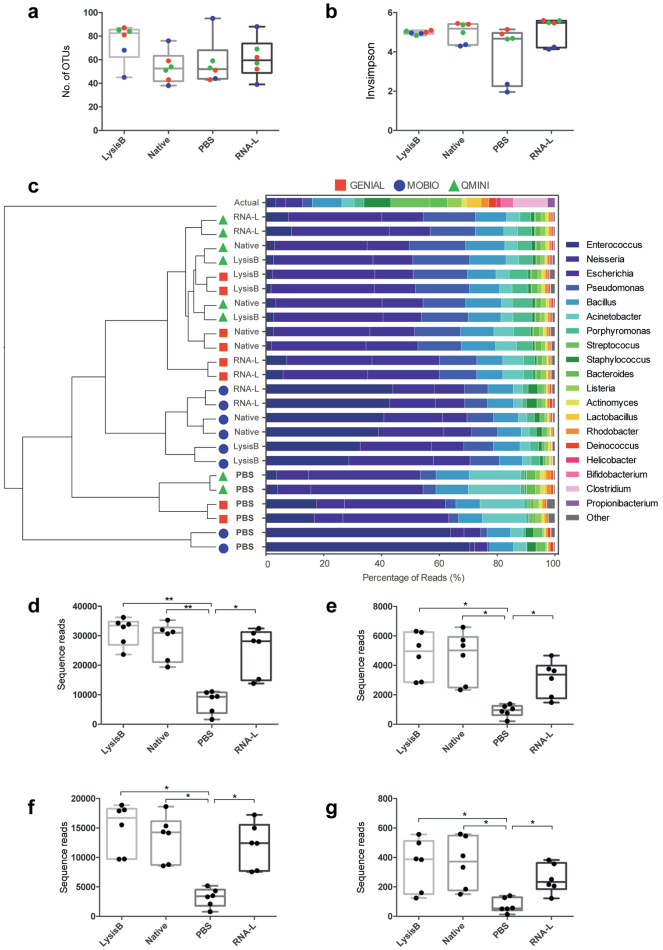


Figure 3. PBS stored samples significantly distort individual OTUs and cluster separately from other buffer types. Boxplots (median ± range) of (**a**) the number of OTUs and (**b**) the Inverse of the Simpson index for each buffer type. (**c**) UPGMA clustering on Bray-Curtis dissimilarities including taxa plots showing the relative abundance of OTUs in percentage of reads. Differential clustering of PBS to all other buffers was found to be significant (parsimony test, $p = 0.001$) (**d**–**g**) Individual bacterial OTUs are significantly underrepresented for PBS-stored samples. Number of sequence reads for OTUs corresponding to (**d**) *Neisseria*, (**e**) *Pseudomonas*, (**f**) *Porphyromonas*, and (**g**) *Helicobacter*. (Wilcoxon test, $*p < 0.05$, $**p < 0.01$).

other tested kits (all $p < 0.01$, Fig. 4). Overall, all the extraction kits distort the bacterial profile compared to the expected bacterial composition of the mock community (Fig. 4).

Discussion

We compared a variety of storage buffers, extraction kits, and amplification methods to examine which combination of handling conditions best represents the microbial diversity of an even mock community (Fig. 1). Different combinations of factors that most closely resemble the reality of sample handling were analyzed to facilitates the

Extraction Method	Observed OTUs	Chao1	InvSimpson
MOBIO PowerLyzer PowerSoil Kit	62.88 ± 8.38	69.79 ± 10.31	3.9 ± 0.40
GEN-IAL First All-Tissue Kit	59.75 ± 5.82	66.02 ± 7.51	5.3 ± 0.09
QIAamp DNA Mini Kit	64.00 ± 4.87	78.04 ± 5.94	5.1 ± 0.13
Actual/Expected Mock Community	22	22	18.3

Table 2. Alpha diversity measurements (mean ± SEM) for each of the DNA extraction kits (n = 8).

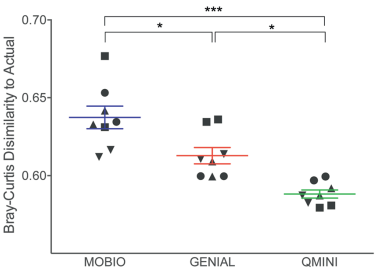


Figure 4. Bray-Curtis dissimilarity between observed and expected strain proportion for each of the tested extraction methods. The expected strain proportion (actual) was generated for comparison and represents the theoretically composition of the mock community (see methods for detail). The pair-wise proportions (expected to observed) from samples extracted with the same commercial kit were grouped in a single boxplot (mean ± SEM). Symbols illustrate different buffer types (■ PBS, ▼RNA-later, ▲native, ●lysis buffer) (Paired t-test, *$p < 0.05$, ***$p < 0.001$).

establishment of standards for the analyses of microbial compositions in swab samples. We show that the choice of storage buffer and extraction kit affects the detected bacterial composition, while different amplification methods had only a minor effect.

Using a mock community, four storage buffers were tested that have been previously used in various studies[25–28]. All samples in this study were frozen at −80 °C rapidly after collection. The samples stored in RNA-later, lysis buffer and native performed similarly to each other and revealed a similar detected bacterial diversity (Fig. 3). Samples stored in RNA-later have been previously reported to decrease DNA purity, lower DNA extraction yields, and to significantly alter the microbial diversity compared to native frozen samples[10,29]. This, however, was not observed in our study. It is likely, that this reflects differences in the sample material (microbes on swab vs. fecal samples) as it has been observed that fecal samples are harder to disperse evenly in RNA-later which may affect the storage and extraction efficiency[10]. Interestingly, compared to the other buffer types, swabs stored in PBS show an altered bacterial composition. There is no indication of a PBS buffer specific contamination profile in the blank samples that could explain this differential clustering. Moreover, PBS buffer in combination with the MOBIO extraction kit detected a lower evenness, which indicates that PBS seems to be particularly incompatible with certain extraction kits. PBS is a balanced salt solution that maintains pH, osmotic balance and is therefore frequently used as a wash buffer in cell and tissue culture. PBS storage has been recommended by manufacturers protocols and has been previously used when examining various extraction kits[12,30]. Other studies examining the effect of different storage conditions have not tested PBS despite its use in DNA extraction from swab material[6,10–12]. It is not clear what properties of PBS effect the mock community differently from other storage buffers. Due to its properties, the buffer may stabilize certain cell types and therefore create a different bacterial profile. Interestingly, despite the different bacterial profile, the PBS samples perform similarly to the other buffer types when comparing them to the mock community. This indicates that the choice of buffer can affect the bacterial profile and specific OTUs, but does not lead to a significantly worse representation of the bacterial community. Our findings support the notion that standardization in sample collection and handling is essential to allow comparison of data within a study[31]. Additionally, field-wide standardization across handling protocols is vital for each sample type, so that cross-study comparisons become possible.

All extraction methods used in this study identify all 19 OTUs present in the mock microbial community (22 bacterial strains of 19 genera, Supplement Table S1). However, all kits detected a higher richness compared to the actual richness of the mock control. A low concentration of mock community (approximately 1×10^7 cells/mL of each organism) was used in this study to simulate the expected bacterial amount in vaginal or oral swab samples[32]. Therefore, it was not surprising that additional OTUs were detected[13]. However, 99% of the pooled library clusters into 19 OTUs which correspond to the bacteria in the mock community. This indicates that the additionally detected OTUs correspond to a small fraction of sequence reads and may therefore be a result of contamination. This study in combination with previous work suggests that the expected biomass of

vaginal and oral swab samples is sufficient for amplicon-based microbial detection without the need of additional target enrichment[13]. The use of a mock microbial community in this study allowed for direct assessment of the extraction kit performance. This comparison indicated that QMINI provides the best representation of the bacterial community when compared to MOBIO and GENIAL. Using a mock community, Yuan *et al.* also found that an altered version of QMINI provided the best bacterial profile[6]. A study using oral swabs confirmed that QMINI extracts DNA with significantly greater yield and good quality compared to other extraction kits[2]. This is in contrast to previous studies on fecal and soil samples, which found that MOBIO most effectively extracts microbial DNA of various bacterial strains[33]. These reported differences in optimal extraction kit may be due to the differences in sample type. The overall bacterial DNA and exogenous material (e.g. fiber) differs substantially between fecal and swab material[34]. Standardization of the extraction kit may thus only be appropriate within each sample type.

In this study, we find that the choice between the two polymerases and the addition of five cycles in amplification of the 16S rRNA gene did not have a significant effect on the bacterial community structure (Fig. 2). Contrary to our findings, Wu *et al.* report that the choice of polymerase had an effect on the microbial community structure, however, the two polymerases that were tested had considerable differences in the fidelity (20 times and 4 times higher than *Taq*)[35]. The two hot-start polymerases used in our study, had significantly higher fidelity (100 times and 52 times higher compared to *Taq*) and are both recommended for NGS applications by the manufacturers. This may likely explain the lack of observable differences. Unlike polymerase choice, which had no effect on the detected evenness or richness, the addition of five PCR cycles to the amplification method led to an overestimation of the bacterial richness. Previous studies have already suggested that this increase is due to an upsurge of chimeric structures with increased cycle numbers[3,5,35]. This supports the notion that lower cycles numbers are favorable for amplicon sequencing[5].

All tested conditions in this study lead to a distortion of the bacterial community structure compared to the expected bacterial mock composition (Fig. 4). *Enterococcus, Neisseria, Escherichia*, and *Pseudomonas* dominated the detected profile in our study, while other bacteria genera such as *Lactobacillus* were underrepresented. Knowledge of which genera are underestimated in the detected bacterial profile (e.g. *Lactobacillus*) is essential to properly estimate the bias when studying certain bacterial communities (e.g., the vaginal microbiome). In a recent study using the same mock community, the bacterial profile resembled the one detected in our study, indicating that the observed distortion is most likely not due to laboratory or kit specific contamination[3,13]. Instead, the bias could be attributed to a variety of factors that were not examined in this study, such as differential susceptibility of bacteria to lysis[6]. To increase lysis efficiency of a broader spectrum of bacteria, enzymatic pre-treatment has been studied as a potential solution, with mixed results[6,36,37]. Another potential cause for the observed bias is the use of primers for 16 S rRNA gene amplification. Although these are universal, amplification may favour certain bacterial strains thus creating bias in the analysis[38,39]. Shotgun metagenomics has been proposed as a solution as it negates some of the bias caused by the amplification, however, this technique does not negate all of technical caveats as storage and extraction kit choice can still have a major impact on the results[3,40]. Continual improvement to the sample handling conditions for both amplicon sequencing and shotgun metagenomics using mock communities is therefore essential.

Conclusion

For now, investigators should standardize the sample handling methods for each sample type as consistency among sample collection, sample storage and sample processing is able to significantly reduce variation. Preliminary tests on specific sample types should be used to ensure that the comparative analysis is as accurate as possible. Caution is, however, warranted when drawing conclusions about the relative abundance of bacterial populations in a single sample and when combining data for meta-analyses.

References

1. Sobel, J. D. Is there a protective role for vaginal flora? *Curr. Infect. Dis. Rep.* **1**, 379–383 (1999).
2. Vesty, A., Biswas, K., Taylor, M. W., Gear, K. & Douglas, R. G. Evaluating the Impact of DNA Extraction Method on the Representation of Human Oral Bacterial and Fungal Communities. *PLoS One* **12**, e0169877, https://doi.org/10.1371/journal.pone.0169877 (2017).
3. Brooks, J. P. *et al.* The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol.* **15**, 66, https://doi.org/10.1186/s12866-015-0351-6 (2015).
4. V Wintzingerode, F., Göbel, U. B. & Stackebrandt, E. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* **21**, 213–229 (1997).
5. Ahn, J.-H., Kim, B.-Y., Song, J. & Weon, H.-Y. Effects of PCR cycle number and DNA polymerase type on the 16S rRNA gene pyrosequencing analysis of bacterial communities. *J. Microbiol.* **50**, 1071–1074 (2012).
6. Yuan, S., Cohen, D. B., Ravel, J., Abdo, Z. & Forney, L. J. Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS One* **7**, e33865, https://doi.org/10.1371/journal.pone.0033865 (2012).
7. Mackenzie, B. W., Waite, D. W. & Taylor, M. W. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Front. Microbiol.* **6**, 130, https://doi.org/10.3389/fmicb.2015.00130 (2015).
8. Kennedy, N. A. *et al.* The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. *PLoS One* **9**, e88982, https://doi.org/10.1371/journal.pone.0088982 (2014).
9. Willner, D. *et al.* Comparison of DNA extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. *PLoS One* **7**, e34605, https://doi.org/10.1371/journal.pone.0034605 (2012).
10. Dominianni, C., Wu, J., Hayes, R. B. & Ahn, J. Comparison of methods for fecal microbiome biospecimen collection. *BMC Microbiol.* **14**, 103, https://doi.org/10.1186/1471-2180-14-103 (2014).
11. Choo, J. M., Leong, L. E. & Rogers, G. B. Sample storage conditions significantly influence faecal microbiome profiles. *Sci. Rep.* **5**, 16350, https://doi.org/10.1038/srep16350 (2015).
12. Bai, G. *et al.* Comparison of storage conditions for human vaginal microbiome studies. *PLoS One* **7**, e36934, https://doi.org/10.1371/journal.pone.0036934 (2012).

13. Salter, S. J. *et al*. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87, https://doi.org/10.1186/s12915-014-0087-z (2014).
14. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
15. Schloss, P. D. *et al*. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).
16. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
17. Quast, C. *et al*. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
18. Weiss, S. *et al*. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**, 27, https://doi.org/10.1186/s40168-017-0237-y (2017).
19. Větrovský, T. & Baldrian, P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One* **8**, e57923, https://doi.org/10.1371/journal.pone.0057923 (2013).
20. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
21. Bray, J. R. & Curtis, J. T. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27**, 325–349 (1957).
22. Excoffier, L., Smouse, P. E. & Quattro, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491 (1992).
23. Schloss, P. D. & Handelsman, J. Introducing TreeClimber, a test to compare microbial community structures. *Appl. Environ. Microbiol.* **72**, 2379–2384 (2006).
24. White, J. R., Nagarajan, N. & Pop, M. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Comput. Biol.* **5**, e1000352, https://doi.org/10.1371/journal.pcbi.1000352 (2009).
25. Perez, G. I. P. *et al*. Body site is a more determinant factor than human population diversity in the healthy skin microbiome. *PLoS One* **11**, e0151990, https://doi.org/10.1371/journal.pone.0151990 (2016).
26. Yildirim, S. *et al*. Primate vaginal microbiomes exhibit species specificity without universal Lactobacillus dominance. *The ISME journal* **8**, 2431–2444 (2014).
27. Pyles, R. B. *et al*. Cultivated vaginal microbiomes alter HIV-1 infection and antiretroviral efficacy in colonized epithelial multilayer cultures. *PLoS One* **9**, e93419, https://doi.org/10.1371/journal.pone.0093419 (2014).
28. Chehoud, C. *et al*. Complement modulates the cutaneous microbiome and inflammatory milieu. *Proceedings of the National Academy of Sciences* **110**, 15061–15066 (2013).
29. Hale, V. L., Tan, C. L., Knight, R. & Amato, K. R. Effect of preservation method on spider monkey (Ateles geoffroyi) fecal microbiota over 8 weeks. *J. Microbiol. Methods* **113**, 16–26 (2015).
30. Abusleme, L., Hong, B.-Y., Dupuy, A. K., Strausbaugh, L. D. & Diaz, P. I. Influence of DNA extraction on oral microbial profiles obtained via 16S rRNA gene sequencing. *J. Oral Microbiol.* **6**, https://doi.org/10.3402/jom.v6.23990 (2014).
31. Stulberg, E. *et al*. An assessment of US microbiome research. *Nature microbiology* **1**, 15015, https://doi.org/10.1038/nmicrobiol.2015.15 (2016).
32. Virtanen, S., Kalliala, I., Nieminen, P. & Salonen, A. Comparative analysis of vaginal microbiota sampling using 16S rRNA gene analysis. *PLoS One* **12**, e0181477 (2017).
33. İnceoğlu, Ö., Hoogwout, E. F., Hill, P. & van Elsas, J. D. Effect of DNA extraction method on the apparent microbial diversity of soil. *Appl. Environ. Microbiol.* **76**, 3378–3382 (2010).
34. Sender, R., Fuchs, S. & Milo, R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* **14**, e1002533, https://doi.org/10.1371/journal.pbio.1002533 (2016).
35. Wu, J.-Y. *et al*. Effects of polymerase, template dilution and cycle number on PCR based 16S rRNA diversity analysis using the deep sequencing method. *BMC Microbiol.* **10**, 255, https://doi.org/10.1186/1471-2180-10-255 (2010).
36. Gill, C., van de Wijgert, J. H., Blow, F. & Darby, A. C. Evaluation of Lysis Methods for the Extraction of Bacterial DNA for Analysis of the Vaginal Microbiota. *PLoS One* **11**, e0163148, https://doi.org/10.1371/journal.pone.0163148 (2016).
37. Rosenthal, R., Blundell, J. & Perkins, H. Strain-related differences in lysozyme sensitivity and extent of O-acetylation of gonococcal peptidoglycan. *Infect. Immun.* **37**, 826–829 (1982).
38. Soergel, D. A., Dey, N., Knight, R. & Brenner, S. E. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *The ISME journal* **6**, 1440–1444 (2012).
39. Hamady, M. & Knight, R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res.* **19**, 1141–1152 (2009).
40. Jovel, J. *et al*. Characterization of the gut microbiome using 16S or shotgun metagenomics. *Front. Microbiol.* **7**, https://doi.org/10.3389/fmicb.2016.00459 (2016).

Acknowledgements

The authors thank the Biodefense and Emerging Infectious Research (BEI) Resources, NIAID, NIH for providing the cells from Microbial Mock Community (Even, HM-280) as part of the Human Microbiome Project. We thank Dr. Dietmar Zinner, Uwe Schönmann and Dr. Angela Noll (German Primate Center) for their guidance and general support. We thank the staff of the Transcriptome and Genome Analysis Laboratory at the University of Göttingen for their assistance in optimizing the sequencing run.

Author Contributions

The study was designed by L.H.W., C.R., and S.K. Laboratory work was conducted at the German Primate Center and performed by L.H.W. and S.L. Data were analyzed by L.H.W. and S.K. All authors (L.H.W., S.L., C.R. and S.K.) contributed to the manuscript preparation.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-24573-y.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

5.4 The Importance of Baseline Data in Ecology

Independent of the directionality of pathogen transmission (wildlife-to-human or human-to-wildlife), our perspective on microorganisms is mostly driven by the clinical manifestations they cause as well as their economic impact. This creates a bias towards virulent microorganisms that are associated with high morbidity and mortality and neglects the fact that global preparedness requires baseline data on microbial communities within the context of natural ecosystems. The comparison of related non-pathogenic and pathogenic microorganisms can help to identify the potential for microbes to become pathogenic before the emergence of the next pandemic. Until today, less than 0.1% of the estimated species of bacteria (Locey and Lennon, 2016; Overmann et al., 2013) and less than 2.0% of estimated fungal species (Overmann and Scholz, 2017) are known. Moreover, there are an estimated 1.67 million viruses from key zoonotic viral families that infect mammals and birds which are yet undiscovered (Carroll et al., 2018). Although wildlife is generally not seen as a source of pathogen threat, it plays a critical role as an indirect source for candidate pathogens that can impact human and livestock health (Karesh et al., 2012; Kock, 2014). A good example of this was the spread of the severe acute respiratory syndrome (SARS) in November 2002 (Drosten et al., 2003). The syndrome was caused by a newly emerged coronavirus of rhinolophid bat origin (as reviewed in (Drexler et al., 2014)). The virus infected more than 8,000 individuals with near to ten percent fatality and caused an approximately 40 billion USD economic loss in a single year (Lee and McKibbin, 2004). These incidences emphasize the need for comprehensive epidemiological investigations and disease surveillance in wildlife as a key requirement to identify and monitor disease reservoirs (Artois et al., 2009). Since a substantial number of the recent pandemics had a viral aetiology of animal origin, much of the current work is directed on the discovery of zoonotic viruses (Anthony et al., 2013; Carroll et al., 2018). An ‘all-pathogen’ inclusive wildlife epidemiosurveillance, in combination with long-term research monitoring, is the only way to understand infectious diseases beyond the scale of individual case reports. The ecological and evolutionary perspective on pathogens is a requirement to adequately manage diseases and to conduct a risk assessment of future transmission events (Karesh et al., 2012). The importance of baseline data is not restricted to the simple presence or absence of microbes that have the potential to cause diseases. It should further include temporal changes in prevalence and virulence (Brearley et al., 2013), which can be age-specific (Iritani et al., 2019), strain dependant (Alizon et al., 2013) or based on changes in the host community composition (Roche et al., 2012) or size (Swart et al., 2017).

6. The Control, Elimination and Eradication of Diseases associated with Wildlife Reservoirs

In its simplest form, the epidemiology of infectious diseases is made up of three components: the host (defined as infected), the agent (defined as the pathogen), and the environment (defined as the population of susceptible hosts) (Katz et al., 2014). Eradication programs that target diseases on the host and environmental level and which do not have a nonhuman reservoir have a great chance of success as demonstrated by the eradication of smallpox (Fenner, 1982). However, despite decades-long attempts to eradicate human-relevant dracunculiasis (guinea worm diseases), poliomyelitis, yaws, malaria, hookworm, and some other diseases (summarized by (Hopkins, 2013)), smallpox and rinderpest remain the only diseases that have

been eradicated (Barrett, 2003; Roeder et al., 2013). While in theory eradication appears feasible in all of the aforementioned human diseases, they share one thing in common that challenges eradication and that distinguished them from e.g., smallpox: their disease ecology does either involve a vector (e.g., in malaria (Benelli and Beier, 2017) or dracunculiasis (Biswas et al., 2013)) and/or is associated with a (potential) nonhuman disease reservoir (e.g., dracunculiasis (Callaway, 2016; Eberhard et al., 2014, 2016; Galán-Puchades, 2016; Galán-Puchades, 2017), yaws (Chapter 7) (Knauf et al., 2018, 2013)) or hookworm infection (George et al., 2016)). This underscores the need to study infectious diseases under the One Health concept (Chapter 3). The decision to control, eliminate or eradicate a disease depends on the technical and biological feasibility including the existence of a nonhuman reservoir (Barrett, 2004, 2003; Hallmaier-Wacker et al., 2017) as well as the benefit in excess of the costs (Barrett, 2004). While the control of diseases requires continuous management and disease surveillance, these measures become obsolete once eradication is achieved (Barrett, 2004). On the downside, disease eradication efforts are labour and cost intensive. Moreover, they require political commitment across all administrative national and international levels (Barrett, 2003) as well as intensified efforts to detect the last cases in the end stage of eradication (Buyon et al., 2018). Based on these difficulties, many diseases are subject to elimination, which describes the absence of a disease in a defined geographic area (e.g., a single country or region) and therefore does not require an infinite global zero-case scenario as it is the case for eradication (Dowdle, 1998). While elimination programs ease the collaboration and coordination across the different administrative levels, they require the installation of continuous regional disease control measures to ensure that the pathogen does not re-emerge from areas that are still endemic for the pathogen (Siembieda et al., 2011).

Wildlife reservoirs are in particular difficult to control as they are often not well studied and not directly accessible to surveillance and control measurements (Siembieda et al., 2011). Moreover, the majority of reservoirs in wildlife are multi-host systems of high complexity (Viana et al., 2014). In principle, there are three ways to control diseases in wildlife reservoirs: targeting (A) the pathogen, (B) the host or (C) both. In the first case, vaccination or treatment efforts can reduce the number of susceptible and infected individuals (Blancou et al., 2009; Porter et al., 2013). Efforts such as the collection and decontamination of infected carcasses can help to reduce the risk of environmental contamination with sit-and-wait pathogens (Chapter 2.1) (Wobeser, 2002). The oldest example for the successful control of a zoonotic wildlife disease in Europe was the vaccination of the red fox (*Canis vulpes*) against rabies (Blancou et al., 2009). Yet, targeting the pathogen requires feasible ways to distinguish infected from uninfected hosts, is cost and labour intensive and treatment is generally not an option in wildlife. Furthermore, vaccination and treatment can have unpredictable consequences such as the evolutionary selection of higher virulent strains (Boots, 2015) or the emergence of drug-resistance (Radhouani et al., 2014). Protecting hosts from the selection pressure of infection may, therefore, implement long-term ecological consequences for population health and resistance to diseases (Woodroffe, 2001).

Targeting the animal host is a double-edged sword as it is generally ethically more prejudiced (Godfroid, 2017). Also, it induced a perturbation effect that makes the management outcome harder to predict (Carter et al., 2007). Nevertheless, culling is currently still the most commonly applied approach to control wildlife diseases (Bolzoni and Leo, 2013). Yet, the

approach aims to reduce population size to control contact rate and prevalence with varying success (Prentice et al., 2019) basically because of its influence on the demography, behaviour and movement patterns as well as immunology (Prentice et al., 2019). Similar to targeting the pathogen, culling can impact the evolution of virulence if competition among pathogenic variants in the presence of superinfection exists (Bolzoni and Leo, 2013). Diseases, where host population control is frequently applied, are for example African swine fever (Vicente et al., 2019), brucellosis (Harrison et al., 2010) or bovine tuberculosis (Langton, 2019). Further difficulties are associated with multi-host communities, where it has been shown that the removal of a single host within a given ecosystem can have only a limited effect on long-term disease dynamics (Johnson et al., 2019). Due to public opposition to culling, fertility control has been proposed as an alternative method to regulate host populations (White et al., 1997). Apart from the classical control of diseases on the pathogen and host level, there are additional management options that can help to prevent the transmission of pathogens. Fencing (Mysterud and Rolandsen, 2019), feeding ban (Gortazar et al., 2015), vector control through burning of forest vegetation to reduce tick populations (Goodenough et al., 2017) or the use of wildlife corridors in combination with vaccination (Haydon et al., 2006) are just some of these examples.

Although eradication of infectious diseases seems to be a noble goal and therefore worthwhile the efforts, there is a justified scepticism associated with the diseases that are currently subject to eradication efforts (Caplan, 2009). Since the removal of microorganisms from an ecosystem is likely associated with fundamental ecological implications and long-term consequences, the control, elimination and eradication of pathogens is ecologically a highly invasive practice. The ecology of diseases is a dynamic and evolutionary process and ecological niches will likely get recolonized by other pathogens (Cohan and Kopac, 2017). These aspects must be considered when diseases are a target to control, elimination or eradication.

7. *Treponema pallidum*, an Ancient and Stealthy Pathogen

The following chapter will use *Treponema pallidum* infection in nonhuman primates as an example for the investigation of a potential disease reservoir in the context of natural ecosystems, one health, and wildlife health. The chapter is structured to reflect the different aspects that have been outlined in the previous introductory parts of the thesis.

7.1 Genetics of Human and Animal Uncultiviable Treponemal Pathogens

This review introduces the pathogen and highlights difficulties associated with its characterisation.

Infection, Genetics and Evolution 61 (2018) 92–107

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Review

Genetics of human and animal uncultivable treponemal pathogens

David Šmajs[a,*], Michal Strouhal[a], Sascha Knauf[b]

[a] Department of Biology, Masaryk University, Kamenice 5, Building A6, 625 00 Brno, Czech Republic
[b] Work Group Neglected Tropical Diseases, Pathology Unit, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany,

ARTICLE INFO

Keywords:
Treponema pallidum subsp. *pallidum*
T. pallidum subsp. *pertenue*
T. pallidum subsp. *endemicum*
T. paraluisleporidarum
Non-human primates
Molecular evolution

ABSTRACT

Treponema pallidum is an uncultivable bacterium and the causative agent of syphilis (subsp. *pallidum* [TPA]), human yaws (subsp. *pertenue* [TPE]), and bejel (subsp. *endemicum*). Several species of nonhuman primates in Africa are infected by treponemes genetically undistinguishable from known human TPE strains. Besides *Treponema pallidum*, the equally uncultivable *Treponema carateum* causes pinta in humans. In lagomorphs, *Treponema paraluisleporidarum* ecovar Cuniculus and ecovar Lepus are the causative agents of rabbit and hare syphilis, respectively.

All uncultivable pathogenic treponemes harbor a relatively small chromosome (1.1334–1.1405 Mbp) and show gene synteny with minimal genetic differences (> 98% identity at the DNA level) between subspecies and species. While uncultivable pathogenic treponemes contain a highly conserved core genome, there are a number of highly variable and/or recombinant chromosomal loci. This is also reflected in the occurrence of intrastrain heterogeneity (genetic diversity within an infecting bacterial population). Molecular differences at several different chromosomal loci identified among TPA strains or isolates have been used for molecular typing and the epidemiological characterization of syphilis isolates. This review summarizes genome structure of uncultivable pathogenic treponemes including genetically variable regions.

1. Introduction

Bacteria of the genus *Treponema* belong to a group of spirochetes that contain several pathogens of global importance for human and animal health. It includes nonpathogenic and pathogenic species, some of which are uncultivable (Krieg et al., 2010). While cultivable treponemes in humans can cause chronic infections in the oral cavity (Visser and Ellen, 2011), they can also cause a variety of economically important diseases in animals, e.g. lameness in sheep (Duncan et al., 2014) and goat populations (Crosby-Durrani et al., 2016) or digital dermatitis in cattle (Plummer and Krull, 2017). With regard to their global importance for human and animal health, we focus in this review on the uncultivable treponemes and the diseases they cause in humans and animals (Table 1).

2. Uncultivable pathogenic treponemes that cause human diseases

Since the end of the 15th century, the great pox now known as syphilis, became one of the most feared diseases. A cure was not available until the middle of 20th century when penicillin became available to treat syphilis infections. Subsequently, syphilis experienced the first and most dramatic decline worldwide. Physiologically, *Treponema pallidum* subsp. *pallidum* (TPA) is a flat wave bacterium with a number of unusual biological characteristics including few outer membrane proteins and a number of efficient immune evasion mechanisms that inspired the name “stealth” pathogen (for review see Radolf et al., 2016). Other uncultivable pathogenic treponemes that cause human diseases are *Treponema pallidum* subsp. *pertenue* (TPE), the causative agent of yaws, *Treponema pallidum* subsp. *endemicum* (TEN), the causative agent of bejel (also known as endemic syphilis), and *Treponema carateum*, the causative agent of pinta. Infections caused by TPE, TEN, and *T. carateum* are commonly considered to be endemic treponematoses (reviewed by Giacani and Lukehart, 2014).

2.1. Overall genome structure of human pathogenic uncultivable treponemes

Difficulties in stuying non-cultivable bacteria affected the availability of genetic data for these species (Weinstock et al., 2000). This changed recently, when modern molecular and genomics techniques resulted in dozens of completely or partially sequenced genomes from treponemes that were either propagated in rabbits or sequenced directly from human clinical samples (Table 2).

The genome structure of the currently analyzed human uncultivable

* Corresponding author.
E-mail addresses: dsmajs@med.muni.cz (D. Šmajs), mstrouhal@med.muni.cz (M. Strouhal), SKnauf@dpz.eu (S. Knauf).

https://doi.org/10.1016/j.meegid.2018.03.015
Received 24 November 2017; Received in revised form 16 March 2018; Accepted 20 March 2018
Available online 22 March 2018

Table 1
Human and animal uncultivable pathogenic treponemes and the diseases they cause.

Treponema sp.	Host	Disease	Invasiveness/ pathogenicity	Geographical occurrence
T. pallidum subsp. *pallidum*	Humans	Syphilis (lues)	Highly invasive, vertical transmission	Global
T. pallidum subsp. *pertenue*	Humans	Yaws (framboesie)	Moderately invasive no vertical transmission	Endemic; Africa, South America, Asia, Oceania
	Nonhuman primates	Facial and genital lesions	Unknown	Endemic; Africa
T. pallidum subsp. *endemicum*	Humans	Endemic syphilis (bejel)	Moderately invasive no vertical transmission	Endemic; south and Sahel region of Africa, Middle East
T. carateum	Humans	Pinta (carate)	Non-invasive no vertical transmission	Endemic; North and South America
T. paraluisleporidarum ecovar Cuniculus	Rabbits	Venereal spirochetosis and facial lesions	Not pathogenic for humans No vertical transmission	Europe
ecovar Lepus	Hares	Venereal spirochetosis and facial lesions	Unknown pathogenicity for humans	Europe

Table 2
An overview of complete and draft genome sequences of human uncultivable pathogenic treponemes available to date.

Strain/isolate	Place and year of isolation Reference	Genome sequence	No. of genomes	Clade classification[a]	Reference[b] (GenBank/ BioProject Acc. No.)	Genome size (bp)	No. of predicted genes[c]
TPA Nichols	Washington, DC, USA; 1912 Nichols and Hough (1913)	Complete	1	Nichols-like	Fraser et al. (1998), Pětrošová et al. (2013) (CP004010.2)	1,139,633	1,093
TPA Chicago	Chicago, USA; 1951 Turner and Hollander (1957)	Complete	1	Nichols-like	Giacani et al., (2010) (CP001752.1)	1,139,281	1,135
TPA DAL-1	Dallas, USA; 1991 Wendel Jr. et al. (1991)	Complete	1	Nichols-like	Zobaníková et al. (2012) (CP003115.1)	1,139,971	1,122
TPA Sea81-4	Seattle, USA; 1980 Giacani et al. (2014)	Complete	1	Nichols-like	Giacani et al. (2014) (CP003679.1)	1,139,203	1,116
TPA SS14	Atlanta, USA; 1977 Stamm et al. (1983)	Complete	1	SS14-like	Matějková et al. (2008), Pětrošová et al. (2013) (CP004011.1)	1,139,569	1,089
TPA Mexico A	Mexico; 1953 Turner and Hollander (1957)	Complete	1	SS14-like	Pětrošová et al. (2012) (CP003064.1)	1,140,038	1,089
TPA strains/ isolates[d]	Austria, Switzerland, Czech Republic, Argentina, The Netherlands, USA; 1973-2013 Arora et al. (2016)	Draft	24	Nichols-like: 5; SS14-like: 19	Arora et al. (2016) (PJRNA313497)	N/A[e]	N/A
TPA isolates	Lisbon, Portugal; 2009-2014 Pinto et al. (2016)	Draft	25	SS14-like	Pinto et al., (2016) (CP016045-69)	N/A	N/A
TPA isolates	Shanghai, China; 2014-2015 Sun et al. (2016)	SRA data	8	SS14-like[f]	Sun et al. (2016) (PJRNA305961)	N/A	N/A
TPA Amoy	Xiamen, China; 2011 Tong et al. 2017a	Draft	1	SS14-like	Tong et al. (2017a) (CP015162.1)	N/A	N/A
TPE Samoa D	Western Samoa; 1953 Turner and Hollander (1957)	Complete	1	N/A	Čejková et al. (2012) (CP002374.1)	1,139,330	1,125
TPE CDC-2	Akorabo, Ghana; 1980 Liska et al. (1982)	Complete	1	N/A	Čejková et al. (2012) (CP002375.1)	1,139,744	1,125
TPE Gauthier	Congo; 1960 Gastinel et al. (1963)	Complete	1	N/A	Čejková et al. (2012) (CP002376.1)	1,139,417	1,125
TPE CDC 2575	Akorabo, Ghana; 1980 Liska et al. (1982)	Complete	1	N/A	Strouhal et al. (2017) (CP020366.1)	1,139,577	1,124
TPE Ghana-051	Ghana; 1988 Engelkens et al., 1989	Complete	1	N/A	Strouhal et al. (2017) (CP020365.1)	1,139,577	1,124
TPE isolate IND1[g]	Indonesia; 1990 Noordhoek et al. (1991)	Draft	1	N/A	Arora et al. (2016) (PJRNA313497)	N/A	N/A
TEN Bosnia A	Bosnia; 1950 Turner and Hollander (1957)	Complete	1	N/A	Šťaudová et al. (2014) (CP007548.1)	1,137,653	1,125

[a] For clade classification, see end of Section 2.1.
[b] Only genome sequences published in the scientific journals are listed, for additional treponemal draft sequences search GenBank files.
[c] These genes include 54 genes coding for RNAmolecules not tranlated to proteins.
[d] Resequenced laboratory TPA, TPE and TEN strains with previously determined whole genome sequence are not shown.
[e] N/A, not applicable.
[f] Although classified as lineage 2 not related to either Nichols or SS14 clades (Sun et al., 2016), reanalysis of the original data revealed clustering with the SS14 clade (Strouhal et al., 2018).
[g] The IND1 genome represents genome of TPE Kampung Dalan K363 (Noordhoek et al., 1991).

pathogenic treponemes is surprisingly similar with only minor variability in the genome length (~1.14 Mbp; 1,137,653–1,140,038 bp; Šmajs et al., 2012). Furthermore, complete gene synteny has been observed in analyzed *Treponema pallidum* (TP) genomes. The minor differences in the annotation of treponemal genomes including differences in the number and length of annotated genes, resulted either from different annotation algorithms used during genome analyses or from single short indels (or substitutions) present in individual genomes. The

gene numbers used throughout this review correspond to the numbers used in the first author-annotated treponemal genome (AE000520; Fraser et al., 1998) and not to numbers in automatically annotated genomes. The number of annotated pseudogenes in treponemes of human origin is usually less than 1% of all annotated genes (the number of annotated genes range between 1,089 and 1,135; Šmajs et al., 2012). Moreover, the high genetic clonality in human infecting uncultivable treponemes corresponds with the lack of mobile genetic elements including transposons, pathogenicity islands, prophages and plasmids. Human pathogenic uncultivable treponemes are thus examples of a monomorphic bacterial pathogen adapted to a single or limited number of hosts. The limited metabolic capabilities of uncultivable treponemes is associated with reduced viability and inability to replicate outside of a human host (Šmajs et al., 2012; Radolf et al., 2016). The overall whole genome identity between TPA, TPE and TEN strains is greater than 99.7% (Štaudová et al., 2014).

While TPE appears to form a single phylogenetic group that includes isolates or strains derived from African non-human primates (NHP) (Knauf et al., 2017), human TPA isolates appear to belong to two distinct genetic groups, which are named by relatedness to reference strains, i.e. Nichols-like or SS14-like TPA strains (Nichols-like and SS14-like clades). Compared to the TPA Nichols strain, the TPA SS14 strain contains seven deletions, six insertions, and 405 single or multiple nucleotide substitutions (Pětrošová et al., 2013; Šmajs et al., 2016). In addition, there are differences between the Nichols and SS14 groups in the frequency of occurrence of the *tprD* and *tprD2* alleles (see below). Additionally, the two alleles (*tprD2* and *tprD*) differ at 328 nucleotide positions.

2.2. Variable genomic regions of human uncultivable pathogenic treponemes

Most of the chromosomal regions of the uncultivable human pathogenic treponemes are highly conserved. However, several chromosomal loci show considerable genetic variability. The most variable locus of the treponemal genomes is the *tprK* gene, especially the portions that correspond to variable regions V1-V7 of the encoded TprK lipoprotein. A population of treponemes infecting a single host differs in *tprK* sequences as a result of variation that is governed by a gene conversion mechanism (Centurion-Lara et al., 2004). It is believed that this variability produces novel antigenic variants, which are thought to promote chronic treponemal infections. In support of this assumption, *tprK* variability has been seen to increase during the course of human infections. However, during rabbit infections, *tprK* variability decreases (LaFond et al., 2003; LaFond et al., 2006; Heymans et al., 2009). The decreased *tprK* variability during rabbit infection is further supported by genomic data that indicated a deletion of a 1,204-bp region in the vicinity of TP0126, a locus that contains *tprK* donor sequences (Šmajs et al., 2002). A larger subpopulation of the TPA strain Nichols (which had been propagated over a long period in rabbit testes) showed this 1,204-bp long deletion while the remaining subpopulation contained the original sequence (Šmajs et al., 2002). Nonetheless, there is still controversy about the cellular localization of TprK (Centurion-Lara et al., 1999; Hazlett et al., 2001), though its variability suggests an important role in evasion of the host immune response and thus in development of the chronic nature of treponemal infections.

Besides *tprK*, an additional set of eleven paralogous *tpr* genes (or *tpr* pseudogenes) are present in the treponemal genome. These genes represent one of the most variable components of the treponemal chromosome. *tpr* genes are classified into three subfamilies including subfamily I, (*tprC, D, F, I*), II (*tprE, G, J*), and III (*tprA, B, H, K, L*) (Centurion-Lara et al., 2013). Some of the corresponding proteins have been predicted to reside in the outer membrane (TprB, C, D, E, F, I, J; Cox et al., 2010). Similar to TprK, some of the Tpr proteins are believed to contribute to antigenic variance, and thus promote the persistence of treponemal infections in the immunocompetent hosts (Leader et al., 2003).

As shown in earlier studies, some of the TPA strains or clinical isolates contain the *tprD* allele at the *tprD* locus, while others contain the *tprD2* allele at the same position (Centurion-Lara et al., 2013). The majority of the TPA strains or isolates within the Nichols-like clade contain the *tprD* allele (which is identical to the *tprC* allele), whereas most of the members of the SS14 clade contain the *tprD2* allele. However, in the SS14 genome, a minor *tprD* allele has been found in the *tprD* locus (Pětrošová et al., 2013); additionally, the *tprD2* allele has also been found in the Nichols-like TPA strain Sea81-4 (Giacani et al., 2014). A similar situation has been reported for TPE and TEN strains and isolates, where the *tprD* allele were denoted as the *tprD*-like allele (Centurion-Lara et al., 2013). While some TPE and TEN strains (human TPE Samoa D and CDC-2 strains, simian TPE strain Fribourg-Blanc, human TEN strain Bosnia A) contain *tprD2* allele in the *tprD* locus, other TPE strains (human TPE strains Gauthier, CDC 2575, and Ghana-051, simian TPE strain LMNP-1) contain the *tprD*-like allele in this locus. In TPA, TPE, and TEN strains or isolates the occurrence of *tprD* alleles does not correlate with the geographical origin, time of isolation, or with the host species from which the strain was isolated. This suggests the existence of a gene conversion mechanism operating between the *tprC* and *tprD* loci and/or the existence of interstrain recombination.

Pathogenic uncultivable treponemes contain two different rRNA-encoding *rrn* loci, which differ in regard to intergenic spacers. There are therefore two possible constitutions of these loci in the genome. A previous analysis of *rrn* operons in pathogenic uncultivable treponemes revealed that the genomic position of intergenic spacers did not correlate with the classification of the treponemal strains (Čejková et al., 2013). Both *rrn* spacer patterns (Ile/Ala and Ala/Ile) appear to combine independently in the individual treponemal genomes and intragenomic rearrangements mediated by a RecBCD-like system have been proposed as the mechanism of the observed differences in *rrn* spacer patterns (Čejková et al., 2013). Despite the variability found in the *rrn* spacer patterns (Ile/Ala and Ala/Ile), the sequences of 5S, 16S, and 23S rRNA genes have always been found to be identical in both operons, even when the A2058G or A2059G mutations in the 23S rRNA gene (causing resistance to macrolide antibiotics) were present (Stamm and Bergen, 2000a; Matějková et al., 2009; Šmajs et al., 2015).

Other variable genome positions comprise a 60-bp repeat region in the *arp* gene (TP0433) and a number of 24-bp repetitions in the TP0470 gene. The *arp* gene encodes an acidic repeat protein and the TP0470 gene encodes a tetratricopeptide repeat domain-containing protein. Both proteins are potentially associated with the cell envelope (Harper et al., 2008; Naqvi et al., 2015). While the number of repetitions among TPA strains with a fully sequenced genome (Fig. 1) in the *arp* gene varies from 7 to 16, the number among clinical TPA isolates varies from 4 to 19 (Peng et al., 2011) and among TPE from 2 to approximately 15. Similarly, the number of repetitions among TPA strains in the TP0470 gene varies from 10 to 29, and from 12 to 37 among TPE. As shown in Fig. 1, TPA strains appear to have a larger number of repetitions in the *arp* gene compared to TPE strains of human origin. Moreover, there appears to be an inverse correlation between the number of repetitions in the *arp* gene and TP0470 gene among TPE strains (Fig. 1). As shown by Harper et al. (2008), another interesting feature of *arp* gene repetitions is the fact that while the *arp* gene of TPA strains and isolates contains multiple, but distinct, repeat motifs, *arp* gene repetitions in TPE and TEN strains contain only identical repetitive sequences. Analysis of samples taken from the same patient revealed different number of *arp* repetitions in parallel samples suggesting relatively high *arp* gene variability and/or difficulties with its amplification (Mikalová et al., 2013).

The occurrence of pseudogenes is another striking feature of treponemal genomes; the occurrence of pseudogenes correlates with subspecies classification of pathogenic treponemes. The overall distribution of pseudogenes in the complete genomes of uncultivable treponemes is shown in Table 3.

The analysis of treponemal genomes with respect to the presence of

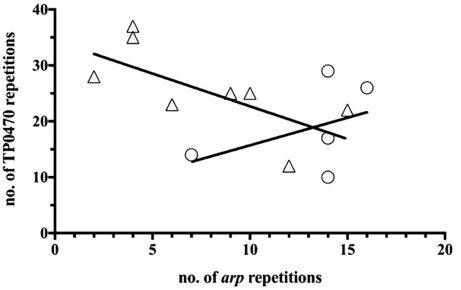


Fig. 1. Number of detected repetitions in the *arp* (TP0433) and TP0470 genes as found in complete treponemal genomes listed in Table 2. In addition, the number of repetitions for two TPE strains, TPE Kampung Dalan K363 and Sei Geringging K403 (Noordhoek et al., 1991), were added (M. Strouhal, unpublished data). The x-axis shows the number of repetitions in *arp* (TP0433) and the y-axis shows the number of repetitions in TP0470. Linear regression analysis showed a significant decrease in number of repetions at the TP0470 locus with the increasing number of repetitions in the *arp* gene in TPE strains (Δ; $F_{1,7} = 5.79$, $p = 0.0470$; $Y = -1.17 * X + 34.4$; $r^2 = 0.4528$). For the TPA strains (○) linear regression analysis showed no significant relationship between the number of repetitions in the TP0470 and the *arp* gene ($F_{1,4} = 0.88$, NS).

conserved regions that differed among TPA, TPE, TEN, and TPeC subspecies and species as well as between Nichols-like and SS14-like strains, but which were consistently found among strains or isolates of the same subspecies/group resulted in the identification of conserved signature indels (CSIs). CSIs were defined as aggregated nucleotide changes containing insertions and deletions that were consistently found in different treponemal subspecies/groups. An overview of such regions is shown in Fig. 2. Twelve out of the 21 regions presented in Fig. 2 were found in the intergenic regions (IGRs), which suggests that alteration of gene expression is a common event that differentiates treponemal subspecies/groups. However, no transcriptome studies supporting this hypothesis are available until now, mainly due to inherent difficulties in the preparation of samples (Šmajs et al., 2005).

2.3. Genetic recombination between T. pallidum strains and subspecies

In addition to several intrastrain recombination events mentioned in the previous section, there is an increasing body of evidence that recombination also occurs between different TP strains and subspecies. The whole genome sequence for TPA strain Mexico A, as an example, revealed a mosaic character of the TPAMA_0326 and TPAMA_0488 loci (Pětrošová et al., 2012). It is thought that these loci most likely contain traces of inter-strain recombination between TPA and TPE strains, which may have resulted from simultaneous infection of a single host via horizontal gene transfer. Moreover, the study showed that the TP0326 of SS14 strain was identical to the sequence found in the TEN Bosnia A genome. Subsequently, several loci identified in the genome of TEN strain Bosnia A (Štaudová et al., 2014) revealed surprising homology to the orthologous sequences of TPA strains in both SS14-like and Nichols-like clade. Recently, an analysis of the bejel clinical isolate 11q/j, obtained from a swab from an indurated penile ulceration, revealed two recombinant loci including TP0488 and TP0548 (Grange et al., 2016; Mikalová et al., 2017a). These findings suggested a possible role of repeated horizontal gene transfer in shaping the TPA, TPE, and TEN genomes. A set of putative recombinant genes identified in previous studies are shown in Table 4. It is likely that the number of putative recombinant genes will increase when more genome data from TP strains or isolates become available.

2.4. Genetic variability within infecting population of uncultivable treponemes

Genetic heterogeneity within an infecting population is frequently present among bacterial agents during infection (van der Woude and Bäumler, 2004; Palmer et al., 2009; Golubchik et al., 2013). The idea that a population of infecting treponemes is not identical but is composed of different subpopulations was presented in 1974 when Baseman et al., detected two different subpopulations of treponemes coming from the same infected rabbit tissue. This was further supported by the identification of a subpopulation of *TP, which was* resistant to phagocytosis (Lukehart et al., 1992). Over time, several examples of genetic heterogeneity in the infecting treponemal population accumulated in particular in the *tpr* gene family (Stamm and Bergen, 2000b; Centurion-Lara et al., 2000; LaFond et al., 2003; Šmajs et al., 2002). The first work mapping genetic heterogeneity on a larger scale was the work of Matějková et al. (2008), which was later extended by the work of Pětrošová et al. (2013). Both studies revealed that intrastrain heterogenous sites are not restricted to *tpr* genes, *tpr* donor sequences, or the vicinity of *tpr* genes since they are localized equally in other genomic regions encoding peptidoglycan synthesis enzyme MurC (TP0341), flagellum specific ATP synthase (TP0402), efflux pump OrpN ortholog (TP0967; Radolf and Kumar, 2017), *tp34* membrane lipoprotein (TP0971), and RNA binding protein (TP1029; encoding hypothetical protein). Additional work by Čejková et al. (2015) mapped genetic heterogeneity among ten treponemal whole genome TPA, TPE and TEN sequences; Strouhal et al. (2017) performed an analysis of two TPE genomes (CDC 2575 and Ghana-051) and Pinto et al. (2016) mapped this phenomenon among 25 clinical isolates that clustered with the TPA SS14 genome. The results of these studies are summarized in Table 5 with only the nucleotide changes leading to non-synonymous amino

Table 3
Occurrence of pseudogenes[a] in pathogenic uncultivable treponemal strains. Because of missing data in the draft genomes, only complete genome sequences were analyzed. TPA strains harbor different subsets of pseudogenes compared to TPE strains.

Strain/Pseudogene	TP0009	TP0012	TP0067	TP0082a	TP0132	TP0135	TP0136	TP0180	TP0266	TP0316	TP0318	TP0609	TP0896	TP1029	TP1030
TPA Nichols	+									+					
TPA SS14	+									+					
TPA DAL-1	+		+				+			+		+			
TPA Mexico A	+									+					
TPE Samoa D					+	+			+		+				+
TPE CDC-2					+	+			+		+				+
TPE Gauthier					+	+			+		+				+
TPE Fribourg-Blanc		+			+				+		+		+		+
TPE CDC 2575					+	+			+		+				+
TPE Ghana-051					+	+			+		+				+
TEN Bosnia A				+	+	+			+	+	+			+	+

[a] Pseudogenes resulting from variable numbers of G or C homopolymers (see Section 2.4) are not listed.

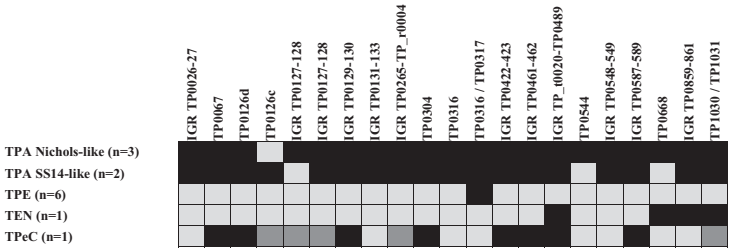


Fig. 2. Conserved signature indels (CSIs) identified in different treponemal subspecies/groups*. Different grey colors correspond to different CSI while identical grey colors indicate highly similar sequences. *Recombinant genes were omitted from the analysis (for recombinant genes see Table 4).

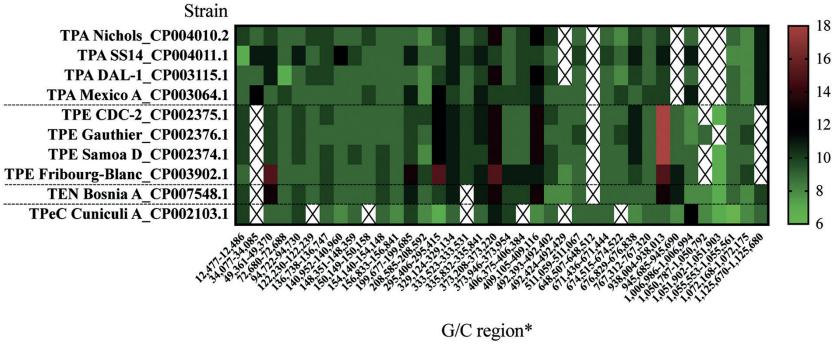


Fig. 3. Genetic differences in the length of homopolymeric tracts (G/C-regions) among and within treponemal strains. Details concerning the affected genes, ORFs and promoters are summarized in the Table 6. *Only homopolymeric tracts showing differences in the number of G/C residues (≥7) are reported; coordinates are based on to the TPA Nichols genome (CP004010.2; Pětrošová et al., 2013). Color codes, reflecting the the length of homopolymeric tracts are shown right to the figure. White color indicates absence of homopolymeric tract.

Table 4
A set of putative recombinant genes identified in previous studies among TP of human origin.

Gene[a]		Annotation	Recombination type	Reference
TP0117	*tprC*	Tpr protein C	Intra-strain	Gray et al. (2006)
TP0131	*tprD*	Tpr protein D	Intra-strain	Gray et al. (2006)
TP0133		hypothetical protein	Inter-strain	Štaudová et al. (2014)
TP0317	*tprG*	Tpr protein G	Intra-strain	Gray et al. (2006)
TP0326	*tp92*	BamA orthologue and rare outer membrane protein	Inter-strain	Harper et al. (2008), Pětrošová et al. (2012), Štaudová et al. (2014)
TP0488	*mcp2*	Methyl-accepting chemotaxis protein	Inter-strain	Pětrošová et al. (2012), Štaudová et al. (2014), Mikalová et al. (2017a), Noda et al. (2018)
TP0548		Putative membrane protein	Inter-strain	Mikalová et al. (2017a)
TP0577		Putative membrane protein	Inter-strain	Štaudová et al. (2014)
TP0620	*tprI*	Tpr protein I	Intra-strain	Gray et al. (2006)
TP0621	*tprJ*	Tpr protein J	Intra-strain	Gray et al. (2006)
TP0858		putative lipoprotein	Inter-strain	Štaudová et al. (2014)
TP0865		outer membrane protein	Inter-strain	Arora et al. (2016)
TP0897	*tprK*	Tpr protein K	Intra-strain	Centurion-Lara et al. (2004), Giacani et al. (2010), Giacani et al. (2012)
TP0968		hypothetical protein	Inter-strain	Štaudová et al. (2014)
TP1031	*tprL*	Tpr protein L	Inter-strain	Štaudová et al. (2014)
tprD/tprD2 alternation			Intra-strain	Centurion-Lara et al. (2013)
rDNA loci - tRNA spacer alternation			Intra-strain	Čejková et al. (2013)

[a] Many genes listed in this table belong to the following 5 paralogous gene families: (TP0040, TP0488, TP0639, TP0640); (TP0136, TP0133, TP0134 TP0462); (TP0548, TP0858, TP0859, TP0865); (TP0966, TP0967, TP0968, TP0969); (*tpr* genes).

Table 5
Treponemal genes[a] with intrastrain heterogeneity leading to non-synonymous amino acid changes.

Gene #	Gene name	Gene function	Identified in treponemal strain/isolate	No of identified sites	References
TP0006		Hypothetical gene	TPA Nichols	Read through stop codon	Čejková et al. (2015)
TP0038	*pfoR*	Probable regulatory protein	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0051	*prfA*	Translation	TPA Nichols	1	Čejková et al. (2015)
TP0065		SAM dependent methyl transferase	TPA DAL-1	1	Čejková et al. (2015)
TP0117	*tprC*	Virulence	TPA SS14	8	Matějková et al. (2008)
			TPA SS14	2	Čejková et al. (2015)
			TPE CDC 2575, TPE Ghana-051	1	Strouhal et al. (2017)
TP0122	*pckG*	Phosphoenolpyruvate carboxykinase (GTP)	TPA SS14-like clinical isolate	Stop codon	Pinto et al. (2016)
TP0131	*tprD*	Virulence	TPA SS14	*tprD/tprD2*[b]	Pětrošová et al. (2013)
TP0134		Outer membrane protein[c]	TPE Samoa D	1	Čejková et al. (2015)
			TPA SS14-like clinical isolate	2	Pinto et al. (2016)
TP0136		Fibronectin-binding protein	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0140		Trk family potassium transporter, membrane protein	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0144	*tbpA*	Thiamine ABC superfamily ATP binding cassette transporter	TPE CDC 2575	1	Strouhal et al. (2017)
TP0151	*rnfD*	NADH dehydrogenase (ubiquinone), subunit RnfD	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0179		Fe-only hydrogenase[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0215	*grpE*	Chaperone GrpE	TPE Ghana-051	1	Strouhal et al. (2017)
TP0216	*dnaK*	Chaperone DnaK	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0219	*rsbU2*	Probable sigma factor regulatory protein	TPA SS14-like clinical isolate	Stop codon	Pinto et al. (2016)
TP0222		Methyltransferase domain protein[c]	TPA Nichols	1	Čejková et al. (2015)
TP0237	*rplK*	Ribosomal protein L11	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0294	*prs*	Ribose-phosphate diphosphokinase	TPA SS14-like clinical isolate	2	Pinto et al. (2016)
TP0307		Penicillin binding- serine/threonine kinase[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0314		Subtilisin-like protein[c]	TEN Bosnia A	2	Čejková et al. (2015)
TP0316	*tprGI*[d]	Virulence	TEN Bosnia A	1	Čejková et al. (2015)
TP0324		Outer membrane protein[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0341	*murC*	Peptidoglycan synthesis	TPA SS14	1	Pětrošová et al. (2013), Čejková et al. (2015)
TP0363	*cheA*	Sensor histidine kinase	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0394	*topA*	DNA metabolism	TPA SS14	1	Čejková et al. (2015)
TP0402		ATP synthase	TPA SS14	2	Matějková et al. (2008), Čejková et al. (2015)
TP0417	*int*	Apolipoprotein N-acyltransferase	TPA SS14-like clinical isolate	Stop codon	Pinto et al. (2016)
TP0439	*cheW2*	Chemotaxis protein CheW	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0455		Fibronectin type III[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0456		Clavaminate synthase-like protein[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0471		Tetratricopeptide repeat containing protein[c]	TPA Nichols	2	Čejková et al. (2015)
TP0488	*mcp-2*	Methyl-accepting chemotaxis protein	TPE CDC 2575, TPE Ghana-051	1	Strouhal et al. (2017)
TP0515	*ostA*	Organic solvent tolerance protein	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0572		FMN-binding domain protein, putative membrane protein	TPE Ghana-051	1	Strouhal et al. (2017)
TP0575	*ptsI*	EI family phosphotransferase system enzyme I	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0578	*ftsY*	Sec family Type I general secretory pathway protein	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0586	*leuS*	Leucine—tRNA ligase	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0600		Putative membrane-associated zinc protease	TPE Ghana-051	1	Strouhal et al. (2017)
TP0610	*tprH*	Virulence	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0615		SUF system FeS assembly protein	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0620	*tprI*	Virulence	TPA SS14	18	Matějková et al. (2008)
			TPA SS14	21	Pětrošová et al. (2013)
TP0621	*tprJ*	Virulence	TPA SS14	3	Matějková et al. (2008)
			TPA SS14	3	Pětrošová et al. (2013)
TP0639	*mcp-3*	Methyl-accepting chemotaxis protein	TPE CDC 2575, TPE Ghana-051	1	Strouhal et al. (2017)
TP0691	*scpA*	Segregation and condensation protein ScpA	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0705	*mrcA*	Bifunctional membrane carboxypeptidase/penicillin-binding protein	TPA SS14-like clinical isolate	2	Pinto et al. (2016)
TP0714	*flhA*	IIISP family Type III (virulence-related) secretory pathway protein	TPA SS14-like clinical isolate	Stop codon	Pinto et al. (2016)
TP0720	*cheC/fliY*	Motility, chemotaxis protein	TPA DAL-1	2	Čejková et al. (2015)
TP0731	*nudE*	ADP-ribose diphosphatase	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0735	*gltD*	Glutamate synthase (NADPH)	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0744		Ribosomal protein[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0755	*ptsN*	PTS family fructose/mannitol (*fru*) porter component IIA	TPE Ghana-051	1	Strouhal et al. (2017)
TP0762		PLP-dependent transferase [e]	TPA SS14-like clinical isolate	2	Pinto et al. (2016)
TP0773	*htrA1*	S1 family peptidase	TPE Ghana-051	1	Strouhal et al. (2017)
TP0780	*nadE*	NAD(+) synthase	TPA SS14-like clinical isolate	2	Pinto et al. (2016)
TP0817	eno	Phosphopyruvate hydratase	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0853	fleN2	Probable flagellar synthesis regulator FleN	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0864		Tetratricopeptide repeat containing protein[c]	TPA SS14-like clinical isolate	2	Pinto et al. (2016)
TP0882		ABC transporter[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0899		NYN domain protein, limkain-b1-type	TPA SS14-like clinical isolate	1	Pinto et al. (2016)

(*continued on next page*)

Table 5 (*continued*)

Gene #	Gene name	Gene function	Identified in treponemal strain/isolate	No of identified sites	References
TP0903	*murD*	UDP-*N*-acetylmuramoyl-L-alanine—D-glutamate ligase	TPE Ghana-051	1	Strouhal et al. (2017)
TP0908	*trmD*	tRNA (guanine-N(1)-)-methyltransferase	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0919		Thioredoxin	TPE CDC 2575, TPE Ghana-051	1	Strouhal et al. (2017)
TP0939		Pyruvate synthase	TPA SS14-like clinical isolate	2	Pinto et al. (2016)
TP0954		Tetratricopeptide repeat containing protein[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0956		"Helical backbone" metal receptor[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0958	dctM	TRAP-T family tripartite ATP-independent periplasmic transporter, membrane protein	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP0967	*orpN*	Outer membrane protein	TPA SS14	3	Pětrošová et al. (2013)
			TEN Bosnia A	3	Čejková et al. (2015)
TP0971		Tp34 lipoprotein	TPA SS14	1	Matějková et al. (2008)
TP0983		Iron ABC transporter ATP-binding protein[c]	TPA SS14-like clinical isolate	1	Pinto et al. (2016)
TP01006	*gyrB*	DNA topoisomerase (ATP-hydrolyzing) subunit B	TPA SS14-like clinical isolate	2	Pinto et al. (2016)

[a] Out of annotated treponemal genes, *tprK* was found to be variable in multiple studies and is not included in this overview. Similarly, intrastrain variability in the intergenic regions is not listed.

[b] The *tprD2* and *tprD* alleles differ at 328 nucleotide positions (Centurion-Lara et al., 2013).

[c] Predictions by Naqvi et al. (2015).

[d] Centurion-Lara et al. (2013).

acid changes being shown. Most observed nucleotide changes showing genetic heterogeneity have resulted in non-synonymous amino acid changes (Čejková et al., 2015; Strouhal et al., 2017; Pinto et al., 2016). With the exception of the highly variable *tprK* gene, that was not analyzed in these studies, genetic heterogeneity was found in 71 additional treponemal genes with 114 variable positions that led to non-synonymous amino acid changes. Most of the variable positions were found within the *tpr* gene family (*tprC, D*, *tprGI*, *tprH, I, J*; the *tprGI* is a gene identified only in TEN Bosnia A and TPeC Cuniculi A genomes), which account for 34 (29.8%) of the variable sites. Besides intrastrain heterogeneity detected in six *tpr* genes, four genes that enable chemotaxis, six genes encoding chaperons and gene regulators, eleven genes for outer membrane proteins, and eleven genes encoding transporters were found to contain such sites. These genes account for 38 (53.5%) out of the 71 analyzed genes. The remaining genes mostly encoded for metabolic functions.

The presence of intrastrain heterogenous sites among *tpr* genes, genes for outer membrane proteins, and genes encoding for transporters suggest possible mechanisms for immune evasion, while the presence of intrastrain heterogeneous sites in genes coding for regulators and enzymes could argue for adaptive changes during infection of different host tissues. Future studies on intrastrain heterogenous sites will need to address the question whether the number of intrastrain heterogenous sites is limited and if some of the changes co-occur in an individual genome and are therefore present together.

Another source of genetic variability within an infecting population of treponemes is represented by homopolymeric tracts, which is a specific example of simple sequence DNA repeats (SSR; microsatellites). Repetitive bacterial sequences were shown to be involved in phenotypic variation in many bacteria through mechanisms of polymerase slippage. The generation of variants in SSR loci has been shown to play an important role in rapid adaptation of pathogenic bacteria to novel environments (Moxon et al., 2006; Gemayel et al., 2010). Interestingly, this mechanism is quite common in bacteria with limited metabolic capabilities (e.g., in *Haemophilus influenzae*, *Neisseria gonorrhoeae*, *Helicobacter pylori*) and has also been proposed for TP (Giacani et al., 2007; Giacani et al., 2015).

Analyses of available sequence data have revealed variability in guanosine/cytosine homopolymeric tracts (G/C-regions) among and within individual pathogenic treponemal strains. Altogether, 38 chromosomal loci containing variable numbers of G/Cs (i.e., polyG/C tracts) have been found among different genomes of uncultivable human and animal infecting treponemes (Table 6). Most homopolymeric tracts have been found in the intergenic regions suggesting a possible role of these homopolymers in gene regulation. Out of 38 positions presented in Table 6, intrastrain genetic variability was found in 22 of the G/C regions, which argues for a positive correlation of interstrain variability and high intrastrain genetic variability. Furthermore, several regions in the genomes of TPA, TPE, and TEN strains suggest important differences in genome structure between the different TP subspecies. In a recent analysis of clinical syphilis isolates from Portugal (Pinto et al., 2016), 31 G/C variable homopolymeric tracts were located within predicted genes (n = 11) as well as in intergenic regions (n = 20). While intrastrain genetic variability in the length of homopolymeric tracts located within genes frequently leads to protein truncation, variable lengths of homopolymeric tracts in the intergenic regions likely result in changes in the transcription rate (Giacani et al., 2015). In both cases, *tpr* genes and genes encoding for outer membrane proteins or lipoproteins are clearly enriched, which likely helps to promote phenotypic diversity that is required for host immune evasion. Currently, there are no analyses of variability in the length of adenine/thymine homopolymeric tracts (A/T-regions) in uncultivable pathogenic treponemes (Table 7).

2.5. Mutation rate in uncultivable treponemes that originated from humans and the evolution of treponematoses in humans

Mutation rates in uncultivable treponemes cannot be estimated from *ex vivo* experiments. They must therefore be extrapolated from available genome data of laboratory maintained (rabbit or hamster propagated) strains or from clinical samples. A general estimation has come from studying mutation rates in microbes with DNA-based chromosomes (Drake et al., 1998). In this study it was shown that for many microorganisms the observed mutation rate was close to 1/300 per genome per replication which would result in a proposed mutation rate of 2.92×10^{-9} per site per generation in TP. However, the mutation rate varies among different microorganisms (Ochman et al., 1999; Eppinger et al., 2006) and data for uncultivable pathogenic treponemes were not available. A more recent study conducted by de Mělo et al. (2010) performed an analysis of genetic differences in 21 genetic regions obtained from available genomic data of TPA, TPE, and TEN strains. This was done in the context of paleo-pathological findings and the authors concluded that the most probable estimated evolutionary rate of syphilis treponemes was 8.82×10^{-8} substitutions per site per year. Under the assumption that TPA has an average doubling time of 30 h (Magnuson et al., 1948; Cumberland and Turner, 1949), this

Table 6
A list of genes, ORFs and promoters showing genetic differences in the length of homopolymeric tracts (poly-G/C-regions) among and within 10 treponemal strains listed in Fig. 3.

G/C region[a]	Affected gene/encoded protein[b]	Intrastrain variability[c]	ORF/Promoter (P)
12,477-12,486	TP0012 (TonB dependent receptor[d]), TP0013 (HP)	+	ORF/P
34,077-34,085	TP0025 (Peptidase M16[d]), TP0026 (*fliG*), TP0027 (*hlyC*), TP0028 (*hlyC*)	+	P
49,361-49,370	TP0040 (*mcp*), TP0041 (Thymidylate synthase-complementing protein[d]), TP0042 (Peptidoglycan binding protein[d]), TP0043 (HP)	+	ORF/P
72,680-72,688	TP0067 (Tetratricopeptide repeat containing protein[d])	+	ORF
94,722-94,730	TP0084 (Thioredoxin[d]), TP0085 (HP)	+	P
122,230-122,239	TP0106 (*betT*), TP0107 (*licC*)	+	P
136,738-136,747	TP0117 (*tprC*)	+	P
140,952-140,960	TP0120 (*metN*), TP0123 (Tetratricopeptide repeat containing protein[d])	+	N/A[e]
148,351-148,359	TP0126 (Outer membrane protein[d]), TP0127 (Hepatitis B viral capsid protein[d]), TP0128 (Nucleoplasmin-like protein[d]), TP0129 (glutamate 5-kinase[d])	+	ORF/P
150,149-150,158	TP0127 (Hepatitis B viral capsid protein[d]), TP0128 (Nucleoplasmin-like protein[d]), TP0129 (Glutamate 5-kinase[d])	+	ORF/P
154,140-154,148	TP0130 (Repeat protein K[d]), TP0131 (*tprD*)	–	P
156,833-156,841	TP0133 (Outer membrane protein[d]), TP0134 (Outer membrane protein[d]), TP0135 (Sec7 domain protein[d])	+	ORF/P
199,677-199,685	TP0179 (Fe-only hydrogenase[d]), TP0181 (Septum formation initiator[d]), TP0182 (Telomere recombination protein[d])	+	P
208,585-208,592	TP0193 (*rpsS*), TP0194 (*rplV*), TP0195 (*rpsC*)	–	P
295,406-295,415	TP0279 (*rpsA*)	+	ORF
329,124-329,134	TP0313 (*tprE*)	+	P
333,523-333,531	TP0314 (Subtilisin-like protein[d]), TP0315 (Type I L-asparaginase[d]), TP0316 (*tprF*)	+	P
335,832-335,841	TP0316 (*tprF*), TP0317 (*tprG*)	–	P
373,208-373,220	TP0347 (DNA-binding domain protein[d])	–	ORF
373,946-373,954	TP0348 (Heptaprenyl diphosphate synthase[d])	–	ORF
406,375-406,384	TP0379 (*secA*)	–	P
409,105-409,116	TP0380 (DNA repair helicase), TP0381 (Integral membrane protein[d]), TP0382 (Cell division protein[d]), TP0383 (HP), TP0384 (S-adenosyl-L-methionine-dependent methyltransferase[d])	+	P
492,393-492,402	TP0461 (putative transcriptional regulator), TP0462 Subtilisin-like protein[d]	+	ORF/P
492,424-492,429	TP0462 (Subtilisin-like protein[d])	–	P
511,059-511,067	TP0479 (ABC transporter ATP-binding protein[d])	+	ORF
648,507-648,512	TP0594 (Signal peptide protein[d])	–	P
671,436-671,444	TP0617 (Pyrogenic exotoxin B[d]), TP0618 (Proto-oncogene c-Rel protein[d])	+	ORF/P
674,515-674,522	TP0619 (Fe, Mn superoxide dismutase[d]), TP0620 (*tprI*)	–	P
676,829-676,838	TP0620 (*tprI*), TP0621 (*tprJ*)	+	P
767,312-767,320	TP0696 (putative nicotinamidase)	–	P
938,004-938,013	TP0859 (Outer membrane protein[d]), TP0860 (HAMP domain-containing protein[d])	+	ORF/P
945,685-945,690	TP0865 (Tetratricopeptide repeat containing protein[d])	–	ORF
1,006,986–1,006,994	TP0925 (flavodoxin)	+	P
1,050,787–1,050,792	TP0964 (ABC transporter), TP0965 (*macA*), TP0966 (Outer membrane protein[d])	–	P
1,051,902–1,051,903	TP0967 (Immunoglobulin-like beta-sandwich protein[d])	–	ORF
1,055,553–1,055,561	TP0969 (Outer membrane efflux protein[d]),	–	ORF
1,072,168–1,072,175	TP0986 (HP)	–	P
1,125,670–1,125,680	TP1031 (*tprL*)	–	P

[a] Only homopolymeric tracts showing differences in the number of G/C residues (≥7) are reported (Pinto et al., 2016); coordinates are based on the TPA Nichols genome (CP004010.2; Pětrošová et al., 2013).

[b] HP - hypothetical protein; a set of treponemal ORFs associated with poly-G/C-repeats (also described in Giacani et al., 2007).

[c] Genetic intrastrain heterogeneity in G/C-tracts detected in at least one strain.

[d] Predictions by Naqvi et al. (2015).

[e] N/A - not applicable; G/C-region located downstream of the genes.

would correspond to a value of 3.3×10^{-10} mutations per site per generation and thus estimates the mutation rate for TP to be about one order of magnitude lower than the rate coming from the work published by Drake et al. (1998). Another possible way to estimate the mean evolutionary rate of TPA is to compare whole genome sequences of clinical isolates or laboratory strains taken from humans or experimental animals at different time points. Based on sample isolation dates, a birth-death serial skyline model can be applied (Stadler et al., 2013). This method was used in a recently published study by Arora et al. (2016), where the mean evolutionary rate of TPA was estimated to be 6.6×10^{-7} substitutions per site per year, which corresponds to a value of 2.26×10^{-9} per nucleotide site per generation. Recent work of Strouhal et al. (2017) revealed in two TPE strains isolated from patients in Ghana over a time span of seven years and three months, an identical genomic consensus sequence. The elapsed time between the isolation of the two samples and absence of fixed nucleotide changes corresponded to a maximum mutation rate of 1.21×10^{-7} per TPE nucleotide site per year or lower or 4.1×10^{-10} per site per generation. Although the applied methods give estimates of the TPA and TPE mutation rates, they all fall within one order of magnitude. However, the mutation rate in uncultivable pathogenic treponemes is probably even below 4.1×10^{-10} per site per generation since only the maximum mutation rate was estimated (Strouhal et al., 2017) and since other slow growing human pathogens including *Mycobacterium tuberculosis* have mutation rates that are even lower (Ford et al., 2011).

With about 2,000 nucleotide differences between TPE and TPA strains (Čejková et al., 2012), and since most of these difference are neither accumulated in a particular chromosomal regions nor were they result of recombination events, it likely that the most recent common ancestor of the syphilis and yaws treponemes dates back more than ten thousand years. Moreover, if the mutation rate estimated for TPE strains is applied to TPA strains, the genetic differences between the Nichols and SS14 clade would date back to a time well before Christ. The sudden emergence of syphilis followed by the return of Columbus

Table 7
An overview of studies that performed CDC typing and/or ECDC on TPA isolates of human origin.

Country	Number of CDC investigated isolates	Number of ECDC investigated isolates	References
Belgium	9	–	Mikalová et al., 2017b
Czech Republic	38	–	Flasarová et al., 2012
Denmark	197	–	Salado-Rasmussen et al., 2016
France	–	71[a]	Grange et al., 2013
Italy	–	49	Giacani et al., 2018
Ireland	–	10	Marra et al., 2010
Portugal	42	–	Florindo et al., 2008
Russia	-	77	Khairullin et al., 2016
UK	36	5	Cole et al., 2009; Tipple et al., 2011
Canada	36	–	Martin et al., 2010
USA	260	273	Pillay et al., 1998; Sutton et al., 2001; Pope et al., 2005; Katz et al., 2010; A2058G Prevalence Workgroup, 2012
Argentina	–	23	Gallo Vaulet et al., 2017
Cuba	–	37	Noda et al., 2016
Colombia	6	–	Cruz et al., 2010
Peru	–	7	Flores et al., 2016
China	36	1262	Martin et al., 2009; Marra et al., 2010; Dai et al., 2012; Peng et al., 2012; Tian et al., 2014; Xiao et al., 2016; Tong et al., 2017b; Zhu et al., 2016; Zhang et al., 2017; Li et al., 2017
Taiwan	–	131	Wu et al., 2012; Wu et al., 2014
Madagascar	21	20	Pillay et al., 1998; Marra et al., 2010
South Africa	274	77	Pillay et al., 1998; Pillay et al., 2002; Molepo et al., 2007; Müller et al., 2012
Australia	88[a]	191	Azzato et al., 2012; Read et al., 2016

[a] When more isolates from a single patient were available, number of patients are shown.

from the Americas could thus suggest the introduction of the strains belonging to one of the clades not present in Europe (either Nichols or SS14) rather that the rapid evolution of a new TPA pathogen. However, no data showing differences in infectivity or virulence of SS14 and Nichols clades are available. The accumulating genetic and genomic data and the recent estimates of mutation rates in treponemes are more consistent with a slow evolution of uncultivable pathogenic treponemes, possibly even reflecting the evolution of modern humans (Strouhal et al., 2017).

2.6. Molecular typing of syphilis-causing treponemes

Accumulation of genetic data on uncultivable pathogenic treponemes and deciphering of the first treponemal genome (Fraser et al., 1998) coincided with the development of the molecular typing system created by the Centers for Disease Control and Prevention (CDC typing) (Pillay et al., 1998). CDC typing determines the number of 60-bp tandem repeats within the *arp* (TP0433) gene in combination with restriction fragment length polymorphism analysis of the genes *tprE* (TP0313), *tprG* (TP0317), and *tprJ* (TP0621) genes, which can be amplified in a single PCR reaction (Pillay et al., 1998). Although the CDC typing scheme has been supplemented by detection of the number of poly-G repetitions within the *rpsA* gene (Katz et al., 2010), or by a sequencing analysis of an 83-nt long gene fragment of the TP0548 gene (enhanced CDC typing scheme [ECDC]; Marra et al., 2010), the principal limitations of this scheme remained unsolved. Briefly, there is no control during multiplex amplification of the genes *tprE, G,* and *J* genes, indicating that all three genes are amplified with equal efficiency. In addition, polymerases may introduce errors in the number of long repeat regions and the repeat region length estimation in the *arp* gene is inaccurate due to the limited use of sequencing for this purpose. Furthermore, genetic instability of the *arp* and *tpr* loci has been shown during the examination of parallel samples taken from the same patients indicating limited applicability of the CDC typing system in epidemiological studies (Mikalová et al., 2013). Finally, another principal limitation of this method is the fact that the detected CDC subtypes do not correlate with the phylogeny of TPA strains and isolates (see also Fig. 1; Flasarová et al., 2012; Grillová et al., 2014).

Though a considerable amount of genetic variability has been detected in these studies (for reviews see Peng et al., 2011; Ho and Lukehart, 2011; Tipple and Graham, 2015), few associations with other data or with patient parameters have been found, including the prevalent 14d/f genotype among neurosyphilis (Marra et al., 2010) and macrolide resistance (Read et al., 2016), HIV-status (Grimes et al., 2012), or the increased risk of serofast status (Zhang et al., 2017). Subsequently, as an alternative for CDC typing, sequencing-based molecular typing was introduced (Flasarová et al., 2006; Woznicová et al., 2007; Flasarová et al., 2012). This typing system is based on sequencing of variable but relatively stable genetic loci (stable at least for several years (Flasarová et al., 2012; Strouhal et al., 2017)) and include portions of TP0136 and TP0548 genes together with restriction analysis or sequencing of both 23S rRNA loci where mutations at nt positions 2,058 or 2,059 (numbered based on *E. coli*) lead to macrolide resistance (Stamm and Bergen, 2000a; Matějková et al., 2009). Since most clinical isolates in Europe, China, and North America belong to the clade of TPA SS14-like strains (see below), this typing system needs to be augmented by sequencing of additional loci that increase the resolution of SS14-like isolates. However, few studies have been published that make use of sequencing-based molecular typing, e.g. Czech Republic (Flasarová et al., 2006; Flasarová et al., 2012; Grillová et al., 2014), Belgium (Mikalová et al., 2017b), and Argentina (Gallo Vaulet et al., 2017). Sequencing based typing has been able to reveal dynamic changes in genotype distribution (Grillová et al., 2014) and association between several genotypes with patients age, geographical origin of samples and men having sex with men (MSM) status (Grillová et al., 2014). Since 2010, when CDC typing has been supplemented by sequencing of an 83-bp fragment of the TP0548 gene (Marra et al., 2010), the two typing schemes partially overlap. The sequencing data from the TP0548 gene can be used to distinguish clinical isolates belonging to either the SS14- or the Nichols-like clade. Enhanced CDC types of TP0548 “a, b, c, d, h, and o” are related to the TPA Nichols clade, while sequence types “e, f, g, i, j, k, m, n, p, and q” are related to the TPA SS14 clade (Fig. 4).

From the 1,989 clinical samples that can be classified with respect to Nichols-like or SS14-like cluster, only 117 (5.9%) worldwide clinical isolates are Nichols-like, while the remaining 94.1% of isolates are SS14-like (Fig. 5). This finding was surprising in the light of the fact that most of the laboratory maintained TPA strains are Nichols-like (Šmajs et al., 2016). However, while most of the tested TPA clinical isolates were collected recently, most of the reference laboratory strains were isolated in the USA during the last century. One possible

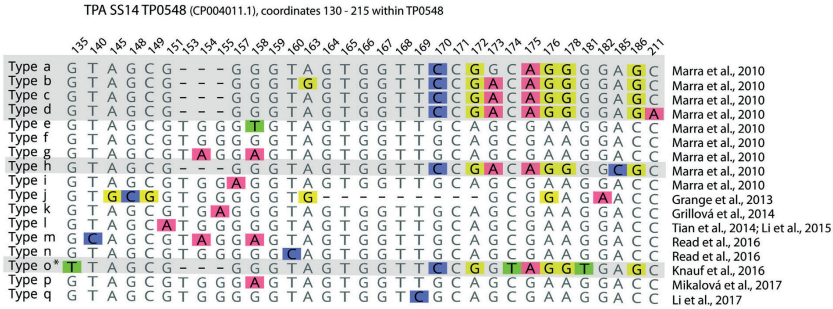


Fig. 4. An overview of enhanced CDC types in TP0548 denoted by alphabetical letters. Enhanced CDC types related to TPA Nichols clade are shaded.
*The type "o", described by Li et al. (2017), should be renamed as type "q" since two new types including "o" (Knauf et al., 2016) and "p" (Mikalová et al., 2017b) were published before the study of Li et al., 2017, apppeared.

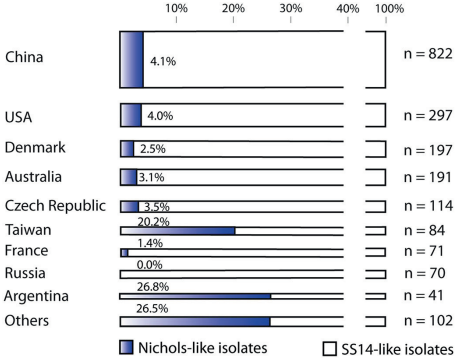


Fig. 5. An overview of studies mapping the occurrence of SS14- (shown in white) and Nichols-like TPA strains (shown in grey). Smaller sample sizes show marked differences in the occurrence of SS14- and Nichols-like TPA strains.

explanation for this phenomenon involves a change in the relative abundance of Nichols-like and SS14-like strains linked to early acquisition of macrolide resistance by SS14-like clade and the subsequent expansion of SS14-like macrolide-resistant TPA strains (Arora et al., 2016). An alternative explanation, involving more effective infection of Nichols-like isolates during inoculation of experimental rabbits or a selective advantage for SS14-like strains over Nichols-like strains in human infections, should be considered. However, no data exist that supports a higher infectivity or virulence of Nichols-like strains relative to rabbits or higher infectivity or virulence of SS14-like strains relative to humans. In general, the existence of SS14- and Nichols-like groups of TPA strains/isolates suggests a historical period, during which there was independent evolution of both TPA clades, likely in syphilis-infected human populations (Šmajs et al., 2016).

3. Uncultivable pathogenic treponemes causing diseases in animals

3.1. Uncultiviable treponemal pathogens in lagomorphs

Although the first description of a rabbit (*Oryctolagus cuniculus*) infecting spirochete (*Spirochaeta paralues-cuniculi*; Jacobsthal, 1920) appeared soon after the description of the syphilis treponeme, surprisingly little data accumulated on treponematoses naturally infecting lagomorphs. *Spirochaeta paralues-cuniculi* was subsequently renamed *Treponema paraluiscuniculi* and was recently reclassified as *T. paraluisleporidarum* ecovar Cuniculus (TPeC) (Lumeij et al., 2013). Rabbit treponematosis can be sexually transmitted and symptoms like erythema, edema, and crusting ulcers can be found in the anogenital region. Although TPeC can be transmitted from mother to neonates during parturition, in-utero infection seems to be uncommon (Posautz et al., 2014). Interestingly, infection of the nose, eyelids, lips, and paws can also occur (Smith and Pesetsky, 1967), mimicking some of the clinical features seen in human yaws. While TP and TPeC are genetically distinct species, there is a high degree of similarity relative to the clinical manifestations they produce. Since the first description of a syphilis-like infection in European brown hares (*Lepus europaeus*) in 1957 (Jakšić, 1957) it has now been reported (based on serology) in several European countries (Table 8).

Although detailed epidemiological data are lacking, infection is not limited to European brown hares since it can also be found in mountain hares (*Lepus timidus*) (Mörner, 1999; Knauf et al., unpublished results). Clinical infection with *T. paraluisleporidarum* ecovar Lepus (TPeL) in hares is often reported during the hunting season when infected animals are shot. Nonetheless, the impact of TPeL on hare populations is currently unknown. Clinical manifestations of hare infection include orofacial and anogenital proliferative crusty skin lesions. However, the infection seems to be most often clinically inapparent, since clinical signs of infection are rare compared to the frequency of seroconversion (Horvath et al., 1980; Lumeij et al., 1994). No data are available with respect to vertical transmission of TPeL and congenital hare syphilis.

In 2011, the complete genome sequence of *T. paraluisleporidarum*

Table 8
The reported serological evidence for hare syphilis in Europe.

Country	Seroprevalence of hare syphilis (%)	No. of hares tested	Reference
Serbia	20	NA[a]	Jakšić, 1957
Hungary	28	202	Horvath et al., 1980
Netherlands	60	100	Lumeij et al., 1994; Lumeij, 2010
Sweden	1	1118	Mörner, 1999
Italy	33	154	Verin et al., 2012
Austria	64	112	Posautz et al., 2014
Germany	44	69	Posautz et al., 2014

[a] NA, not available.

ecovar Cuniculus (formerly named *T. paraluiscuniculi*) strain Cuniculi A was published (Šmajs et al., 2011). A 98.1% sequence identity between the TPA Nichols and TPeC genomes was found (Šmajs et al., 2012), which indicates a close relatedness of TPeC to human infecting TP. Compared to TPA, TPE, and TEN, the TPeC Cuniculi A genome was about 6 kb (0.5% of total genome size) smaller with an overall identical gene synteny. Several deletions, insertions, and other prominent sequence changes in the TPeC Cuniculi A genome were found in 38 Cuniculi A gene homologs (Strouhal et al., 2007). These homologs comprised mainly *tpr* genes and the genes in the vicinity of these loci, suggesting their possible role in the host range and pathogenicity of TPeC. Interestingly, most of the prominent insertions identified in the TPeC genome indicated unique sequences with no homologous sequences identified during a BLAST search. In contrast to TPA, TPE, and TEN strains, the *tprD2*-like allele was identified in both *tprC/D loci*. In addition, the TP0136 gene homolog shares a 99.5% identity with the TP0133 locus, which is an unussually high homology between these paralogous genes. Moreover, the TPeC genome included a relatively high number of pseudogenes and gene fragments (n = 51) compared to TPA and TPE strains (Šmajs et al., 2012). The affected genes were mainly involved in cell envelope biosynthesis and structure as well as in DNA recombination, cell signaling, and gene regulation. Since downsizing of the genome and accumulation of pseudogenes is common for bacteria adapting to simpler host-associated niches (Pallen and Wren, 2007), the genome decay found in the TPeC genome may be an adaptation of TPeC to the rabbit host. At this stage, infectivity to humans was most probably lost, which suggests that TPeC could be a possible descendant of human pathogenic TP (Šmajs et al., 2011). While more research is needed in this field, it is currently believed that TPeC is an exclusive lagomorph pathogen.

The causative agent of hare syphilis was shown to be closely related to TPeC, although only four nucleotide differences were found during sequencing of the 16S rRNA gene (Lumeij et al., 2013). Analysis of additional chromosomal loci (Fig. 6; D. Šmajs, unpublished data) revealed clear relatedness between both rabbit and hare pathogens compared to TP subspecies that infect humans. Future analysis of genomic structure and diversity of the agents of hare and rabbit syphilis could be of great importance in clarifying the evolution of syphilis treponemes and for identification of virulence determinants that allow infections of human and non-human primate hosts.

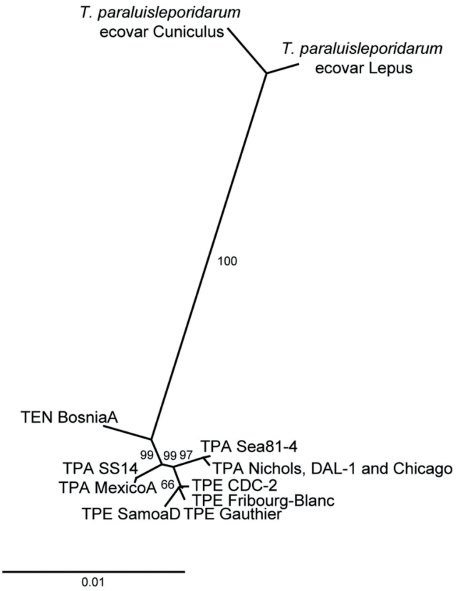


Fig. 6. A phylogenetic tree constructed from concatenated nucleotide regions of TP0011, TP0072, TP0319, TP0488, TP0511, TP0548, TP0610, and the 16S rRNA sequenced in *T. paraluisleporidarum* ecovar Lepus. The corresponding sequences of other treponemal strains were taken from published whole genome sequences (GenBank accession numbers: CP004010.2, CP004011.1, CP003115.1, CP003064.1, CP001752.1, CP003679.1, CP002374.1, CP002375.1, CP002376.1, CP003902.1, CP002103.1, CP007548.1). The maximum likelihood, heuristic search, and 1,000 bootstrap replicates were used for tree construction. The numbers close to branches represent bootstrap support (%). The agent of hare syphilis is similar but distinct from the agent of rabbit syphilis. The bar scale represents the number of substitutions per target site.

3.2. Uncultivable treponemal pathogens in nonhuman primates (NHPs)

Infection research on *Treponema* has a long history in NHPs. As early as the 20th century, infection experiments with human-isolated TP were conducted in chimpanzees (Metchnikoff and Roux, 1903; Metchnikoff and Roux, 1904; Metchnikoff and Roux, 1905) and monkeys (Ashbury and Craig, 1907; Castellani, 1907; Nichols, 1910). Since then few NHP models for human treponematoses have been established with limited success (Turner and Hollander, 1957; Clark and Yobs, 1968; Elsas et al., 1968; Sepetjian et al., 1969; Sepetjian et al., 1972; Marra et al., 1998). All of the models, however, indicated the potential for cross-species transmission of human TP. In the 1960s, the first reports of natural infections of NHPs with TP appeared (Fribourg-Blanc et al., 1963; Fribourg-Blanc et al., 1966; Fribourg-Blanc and Mollaret, 1969). Table 9 summarizes these early findings in African monkeys and great apes. In Guinea baboons, clinical signs of infection were described as mild with small keratotic lesions and ulcers on the muzzle, eyelids, and in the armpits (Baylet et al., 1971b). However, most animals appeared clinically unaffected. At the same time and compared to what was found in African NHPs, several investigations using Asian and South American monkeys showed only few and questionable (weak reactive) test results for antibodies against TP (Table 9). This suggested that in Asia and the new world, monkeys were (if ever) not comparably infected with TP, compared to their African counterparts.

From these early investigations, genetic proof of TP has only been obtained from a single Guinea baboon (*Papio papio*) in 1966, from West Africa (Fribourg-Blanc and Mollaret, 1969). Primary studies of infectivity of this isolate in hamsters and rabbits constituted the basis for the long-term availability of a laboratory-maintained TP strain of NHP origin, i.e. the Fribourg-Blanc simian strain (Fribourg-Blanc and Mollaret, 1969). The morphology of this strain was described to be identical to that seen in human TPA and TPE strains (Smith et al., 1971). Ethically questionable inoculation experiments in humans confirmed that the Fribourg-Blanc strain was capable of causing sustainable infection when inoculated into human skin (Smith et al., 1971).

Although gorillas with yaws-like facial skin lesions have been reported since the 1950s (Denis, 1963; Cousins, 1984), a detailed scientific description of clinical manifestations in gorillas from the Republic of Congo in 2007 indicated the similarity between lesions seen in gorillas and the lesions seen in human yaws (Levrero et al., 2007). The possibility of a geographically wide spread infection of African great apes with TP was further supported by yaws-like bone deformations that were documented in 14% of 126 investigated gorilla and 20% of 102 chimpanzee skeletons collected in central Africa (Lovell et al., 2000). There are few reports of animals that have been tested positive for antibodies against TP (*Gorilla gorilla*, six positive of nine tested) (Cousins, 1972; Karesh, 2000; Karesh personal communication). While genetic confirmation of gorilla infection with TP is still unavailable,

Table 9
Summary of the first reported TP cases in NHPs in the 1960-70s. Further details can be found in Harper and Knauf, 2013. We note here that three baboons described by Baylet et al. (1971b) had ulcerative skin lesions. All other tested animals were either reported clinically healthy (no skin lesions) or the information was missing.

Origin	Species	Number of tested animals	Number of TP positive serology	References
Africa	Baboon (*Papio* sp.)	1,085	317	(Fribourg-Blanc and Mollaret, 1969; Baylet et al., 1971a; Baylet et al., 1971b)
	Gueneons (*Chlorocebus* sp.)	72	34	(Baylet et al., 1971a; Felsenfeld and Wolf, 1971)
	Patas monkey (*Erythrocebus patas)*	141	11	(Fribourg-Blanc and Mollaret, 1969;, Felsenfeld and Wolf, 1971)
	Colobus monkey (species not determined)	1	1	(Fribourg-Blanc and Mollaret, 1969)
	Chimpanzee (*Pan troglodytes*)	278	52	(Fribourg-Blanc and Mollaret, 1969, Kuhn, 1970, Felsenfeld and Wolf, 1971)
Asia	Macaques (*Macaca* sp.)	1,309	1[a]	(Felsenfeld and Wolf, 1971)
South America	Owl monkey (*Aotus* sp.)	96	2[a]	(Levine et al., 1970; Felsenfeld and Wolf, 1971)
	Red-bellied titi (*Callicebus molochus*)	25	5[a]	(Levine et al., 1970; Felsenfeld and Wolf, 1971)
	Squirrel monkey (*Saimiri sciureus*)	91	4[a]	(Levine et al., 1970; Felsenfeld and Wolf, 1971)

[a] While the authors report positive serological reactions, they do note that positivity was questionable due to weak reactivity.

there is rudimentary evidence of chimpanzee infection with TP from bone samples (Gogarten et al., 2016).

In the late 1980s, intense investigations started when olive baboons (*P. anubis*) with severe genital ulcers were reported from Tanzania (Wallis and Lee, 1999; Wallis, 2000; Mlengeya, 2004). Genetic analysis confirmed an association between skin ulcers and TP infections in baboons with yaws-like treponemes at the Lake Manyara National Park and Serengeti National Park (Harper et al., 2012; Knauf et al., 2012). At about the same time, our understanding of the treponemes that infect NHPs entered a new stage when the first whole genome sequence data of the Fribourg-Blanc simian strain became available in 2013 (Zobaníková et al., 2013). The genome showed a striking genetic similarity to TPE strains of human origin. Specific differences (i.e. deletions, insertions, and substitutions) that separated the strain from human yaws-causing TPE consisted of 185 nucleotides in 68 genes (Zobaníková et al., 2013). The strain was much closely related to TPE than any of the known human TPA and TEN strains. As a consequence, it was proposed that the Fribourg-Blanc strain be classified as TPE (Zobaníková et al., 2013). This was further supported by a study that made eight more whole genome sequences of simian strains available (Knauf et al., 2017). All eight simian strains, which originated from Sooty mangabeys (*Cercocebus atys*) and African green monkeys (*Chlorocebus sabaeus*), from three different locations in West Africa and olive baboons from Lake Manyara NP in Tanzania, clustered with human TPE strains. This indicated that infection in NHPs was caused by the TP subsp. *pertenue*. Based on the phylogeny, it is believed that the ancestor of TPE must have crossed species barriers more than once (Knauf et al., 2017). Neither human- nor NHP-infecting TPE strains formed host species-specific clades (Knauf et al., 2017). Moreover, the observed star-like phylogenetic branching pattern with low bootstrap support for short basal branches appears to argue for a rapid initial radiation of TPE among humans and NHPs (Knauf et al., 2017).

The remarkable geographic overlap of historic and present endemic areas for human yaws and nonhuman primate infection with TPE argues for a spatial and temporal connectivity between multiple infected NHP species and humans (Knauf et al., 2013). However, important questions on disease maintenance and feasible transmission routes between humans and nonhuman primates remain unanswered (Hallmaier-Wacker et al., 2017). Based on the reports from Tanzania (Wallis and Lee, 1999; Wallis, 2000; Mlengeya, 2004; Harper et al., 2012; Knauf et al., 2012; Knauf et al., 2016), TPE is maintained in wild NHP populations, as well as there is high genetic and functional similarity of the pathogen in NHPs and humans (Zobaníková et al., 2013; Knauf et al., 2017). Feasible direct transmission routes between NHPs and humans exist (Mossoun et al., 2015; Klegarth et al., 2017), as do indirect transmission through insects has also been discussed (Knauf et al., 2016). If transmission between humans and NHPs occurs, bush meat hunting should be considered a likely source of interspecies disease transmission since TPE is mainly transmitted through direct contact with the exsudate of infectious lesions (Richard et al., 2017). Importantly, disease transmission under natural circumstances should also be considered bidirectional (NHP-to-human and human-to-NHP). This was indicated by a scenario where pet monkeys in Asia, in human yaws endemic areas, were tested serologically positive for antibodies against TP, whereas their wild counterparts were serologically negative (Klegarth et al., 2017).

Infections with other uncultivable *Treponema* pathogens in NHPs have rarely been studied. *Treponema carateum* has been used to infect chimpanzees (Varela, 1969; Kuhn III et al., 1970; Chandler Jr. et al., 1972) and the resulting lesions were strikingly similar to corresponding lesions in humans. However, the endstage of chimpanzee pinta was hyperpigmentation or normal pigmentation of the affected skin and not hypopigmentation as commonly seen in humans (Chandler Jr. et al., 1972). Today, only a few hundred *T. carateum* cases in humans, from Mexico and South America, have been reported (Gideon Informatics Inc., http://web.gideononline.com; last visited 17th Aug 2017), but nothing is known about infection in wild NHPs.

3.3. Yaws eradication and One Health

Accumulating data on NHP infections with TP (Knauf et al., 2013) and the recent finding that all whole genome sequenced simian TP strains fall paraphyletic with human yaws-causing TPE strains (Knauf et al., 2017) have not yet changed the strategy to follow a One Health approach. The One Health approach supports the idea that within an ecological context, human health is wedded to animal and environmental health. The close genetic relationship of simian and human yaws-causing strains (Zobaníková et al., 2013; Knauf et al., 2017) as well as the growing nonhuman primate-human interface substantiates the need for further One Health oriented investigations. Sustainable yaws eradication can only be reached when TPE is understood in the context of its natural ecosystem (Hallmaier-Wacker et al., 2017), which may involve the possibility of a bidirectional transmission, as recently described for Asiatic macaques kept as pet monkeys (Klegarth et al., 2017). One Health strategies must therefore find and eliminate possible niches where pathogens can hide and survive eradication attempts. The large geographic distribution and high prevalence of human yaws has led to two mass treatment campaigns of which the second campaign is currently ongoing. From 1952-1964, the first campaign, which was supported by WHO and the United Nations Children's Fund (UNICEF), treated over 50 million yaws cases in 46 countries with injectable penicillin (Asiedu et al., 2014). This reduced the global burden of endemic

treponematoses, of which yaws was the dominant disease, to a historic low of only 2.5 million cases worldwide (Asiedu et al., 2014). While this was considered to be one of the greatest public health achievements, it was unfortunate that the implemented screening and treatment programs came to a stop or were insufficient to tackle the remaining 5% of yaws cases (Asiedu et al., 2014). Since then, human yaws has re-emerged in West Africa, Southern Asia, and the Pacific region (Asiedu et al., 2014). In 2012, the efficacy of a single oral dose of azithromycin to cure yaws (Mitja et al., 2012) has triggered the second and currently ongoing eradication campaign. In a 12-month long treatment trial in a yaws endemic area in Papua New Guinea, the prevalence of active infectious yaws dropped from 2.4% to 0.3% (Mitja et al., 2015). The combination of an initial round of total community treatment followed by the subsequent seeking out of active cases and total target treatment of new cases and their contacts was thought to be able to ultimately eradicate human yaws globally (Marks et al., 2017). Subsequently, the Morges Strategy on yaws eradication was announced in 2012 and declared that its goal was to eradicate yaws by 2020 (The World Health Organization, 2012). Although the envisaged time frame seemed unlikely to be met globally (Marks, 2016), there are yaws-endemic countries (e.g., India) that have already reached countrywide yaws elimination (Narain et al., 2015). A key research question for the epidemiology and eradication of human yaws was formulated based on a dearth of information on the number of active yaws cases in countries that currently do not report human yaws, but where human yaws has been historically reported to be endemic (Marks et al., 2015). Moreover, in clinical settings, the current diagnosis of yaws is based on interpretation of clinical manifestations in combination with serology (Mitja et al., 2013). Other pathogens, of which *Haemophilus ducreyi* is the most relevant, can cause yaws-like skin ulcers and create diagnostic uncertainty (Mitja et al., 2014). Additionally, serological tests are unable to distinguish between infections with different TP subspecies (Centurion-Lara et al., 2006). A low cost and easy to perform rapid test with high specificity for TPE was understood to be essential to optimize target treatment as well as for the post-zero-case surveillance phase (Marks et al., 2015). Multiplex loop-mediated isothermal amplification (LAMP) assays were seen as a way to significantly shorten the time to a TPE diagnosis.

Acknowledgements

This work was supported by the Grant Agency of the Czech Republic [grant numbers GA17-25455S, J17-25589Y] and by a grant of the Ministry of Health of the Czech Republic [grant number 17-31333A]. It was also partially supported by the German Research Foundation (DFG) [grant numbers KN1097/3-1, KN1097/4-1]. We thank to Prof. Thomas Secrest (Secrest Editing, Ltd.) for English editing of the manuscript.

Declarations of interest: none.

References

A2058G Prevalence Workgroup, 2012. Prevalence of the 23S rRNA A2058G point mutation and molecular subtypes in *Treponema pallidum* in the United States, 2007 to 2009. Sex. Transm. Dis. 39, 794–798.

Arora, N., Schuenemann, V.J., Jäger, G., Peltzer, A., Seitz, A., Herbig, A., Strouhal, M., Grillová, L., Sánchez-Busó, L., Kühnert, D., Bos, K.I., Davis, L.R., Mikalová, L., Bruisten, S., Komericki, P., French, P., Grant, P.R., Pando, M.A., Vaulet, L.G., Fermepin, M.R., Martinez, A., Centurion Lara, A., Giacani, L., Norris, S.J., Šmajs, D., Bosshard, P.P., González-Candelas, F., Nieselt, K., Krause, J., Bagheri, H.C., 2016. Origin of modern syphilis and emergence of a pandemic *Treponema pallidum* cluster. Nat. Microbiol. 2, 16245.

Ashbury, P.M., Craig, C.F., 1907. Observations upon *Treponema pertenuis* (Castellani) of yaws and the experimental production of the disease in monkeys. Philip. J. Sci. B 441–467.

Asiedu, K., Fitzpatrick, C., Jannin, J., 2014. Eradication of yaws: historical efforts and achieving WHO's 2020 target. PLoS Negl. Trop. Dis. 8, e3016.

Azzato, F., Ryan, N., Fyfe, J., Leslie, D.E., 2012. Molecular subtyping of *Treponema pallidum* during a local syphilis epidemic in men who have sex with men in Melbourne. Australia. J. Clin. Microbiol. 50, 1895–1899.

Baylet, R., Thivolet, J., Sepetjian, M., Bert, J., 1971a. Seroepidemiological studies on primate treponematosis in Senegal. Bull. Soc. Pathol. Exot. Filiales 64, 836–841.

Baylet, R., Thivolet, J., Sepetjian, M., Nouhouay, Y., Baylet, M., 1971b. Natural open treponematosis in the *Papio papio* baboon in Casamance. Bull. Soc. Pathol. Exot. Filiales 64, 842–846.

Castellani, A., 1907. Experimental Investigations on Framboesia Tropica (Yaws). J. Hyg. 7, 558–569.

Čejková, D., Zobaníková, M., Chen, L., Pospíšilová, P., Strouhal, M., Qin, X., Mikalová, L., Norris, S.J., Muzny, D.M., Gibbs, R.A., Fulton, L.L., Sodergren, E., Weinstock, G.M., Šmajs, D., 2012. Whole genome sequences of three *Treponema pallidum* ssp. *pertenue* strains: yaws and syphilis treponemes differ in less than 0.2% of the genome sequence. PLoS Negl. Trop. Dis. 6, e1471.

Čejková, D., Zobaníková, M., Pospíšilová, P., Strouhal, M., Mikalová, L., Weinstock, G.M., Šmajs, D., 2013. Structure of *rrn* operons in pathogenic non-cultivable treponemes: sequence but not genomic position of intergenic spacers correlates with classification of *Treponema pallidum* and *Treponema paraluiscuniculi* strains. J. Med. Microbiol. 62, 196–207.

Čejková, D., Strouhal, M., Norris, S.J., Weinstock, G.M., Šmajs, D., 2015. A retrospective study on genetic heterogeneity within *Treponema* Strains: subpopulations are genetically distinct in a limited number of positions. PLoS Negl. Trop. Dis. 9, e0004110.

Centurion-Lara, A., Castro, C., Barrett, L., Cameron, C., Mostowfi, M., Van Voorhis, W.C., Lukehart, S.A., 1999. *Treponema pallidum* major sheath protein homologue TprK is a target of opsonic antibody and the protective immune response. J. Exp. Med. 189, 647–656.

Centurion-Lara, A., Godornes, C., Castro, C., Van Voorhis, W.C., Lukehart, S.A., 2000. The *tprK* gene is heterogeneous among *Treponema pallidum* strains and has multiple alleles. Infect. Immun. 68, 824–831.

Centurion-Lara, A., LaFond, R.E., Hevner, K., Godornes, C., Molini, B.J., Van Voorhis, W.C., Lukehart, S.A., 2004. Gene conversion: a mechanism for generation of heterogeneity in the *tprK* gene of *Treponema pallidum* during infection. Mol. Microbiol. 52, 1579–1596.

Centurion-Lara, A., Molini, B.J., Godornes, C., Sun, E., Hevner, K., Van Voorhis, W.C., Lukehart, S.A., 2006. Molecular Differentiation of *Treponema pallidum* Subspecies. J. Clin. Microbiol. 44, 3377–3380.

Centurion-Lara, A., Giacani, L., Godornes, C., Molini, B.J., Brinck Reid, T., Lukehart, S.A., 2013. Fine analysis of genetic diversity of the *tpr* gene family among treponemal species, subspecies and strains. PLoS Negl. Trop. Dis. 7, e2222.

Chandler Jr., F.W., Kaufmann, A.F., Kuhn III, U.S., 1972. The histopathology of experimental pinta in the chimpanzee. J. Investig. Dermatol. 58, 103–108.

Clark Jr., J.W., Yobs, A.R., 1968. Observations of the pathogenesis of syphilis in *Aotus trivirgatus*. Br. J. Vener. Dis. 44, 208–215.

Cole, M.J., Chisholm, S.A., Palmer, H.M., Wallace, L.A., Ison, C.A., 2009. Molecular epidemiology of syphilis in Scotland. Sex. Transm. Infect. 85, 447–451.

Cousins, D.O.N., 1972. Diseases and injuries in wild and captive gorillas: *Gorilla gorilla*. International Zoo Yearbook 12, 211–218.

Cousins, D., 1984. Notes on the occurrence of skin infections in gorillas (Gorilla gorilla). ool. Garten NF Jena 54, 333–338.

Cox, D.L., Luthra, A., Dunham-Ems, S., Desrosiers, D.C., Salazar, J.C., Caimano, M.J., Radolf, J.D., 2010. Surface immunolabeling and consensus computational framework to identify candidate rare outer membrane proteins of *Treponema pallidum*. Infect. Immun. 78, 5178–5194.

Crosby-Durrani, H.E., Clegg, S.R., Singer, E., Angell, J.W., Evans, N.J., Carter, S.D., Blundell, R.J., Duncan, J.S., 2016. Severe foot lesions in dairy goats associated with digital dermatitis treponemes. J. Comp. Pathol. 154, 283–296.

Cruz, A.R., Pillay, A., Zuluaga, A.V., Ramirez, L.G., Duque, J.E., Aristizabal, G.E., Fiel-Gan, M.D., Jaramillo, R., Trujillo, R., Valencia, C., Jagodzinski, L., Cox, D.L., Radolf, J.D., Salazar, J.C., 2010. Secondary syphilis in Cali, Colombia: new concepts in disease pathogenesis. PLoS Negl. Trop. Dis. 4, e690.

Cumberland, M.C., Turner, T.B., 1949. Rate of multiplication of *Treponema pallidum* in normal and immune rabbits. Am. J. Syphilis 33, 201–212.

Dai, T., Li, K., Lu, H., Gu, X., Wang, Q., Zhou, P., 2012. Molecular typing of *Treponema pallidum*: a 5-year surveillance in Shanghai, China. J. Clin. Microbiol. 50, 3674–3677.

Denis, A., 1963. On safari: the story of my life. Collins.

Drake, J.W., Charlesworth, B., Charlesworth, D., Crow, J.F., 1998. Rates of spontaneous mutation. Genetics 148, 1667–1686.

Duncan, J.S., Angell, J.W., Carter, S.D., Evans, N.J., Sullivan, L.E., Grove-White, D.H., 2014. Contagious ovine digital dermatitis: an emerging disease. Vet. J. 201, 265–268.

Elsas, F.J., Smith, J.L., Israel, C.W., Gager, W.E., 1968. Late syphilis in the primate. Br. J. Vener. Dis. 44, 267–273.

Engelkens, H.J., Oranje, A.P., Stolz, E., 1989. Early yaws, imported in The Netherlands. enitourin. Med. 65, 316–318.

Eppinger, M., Baar, C., Linz, B., Raddatz, G., Lanz, C., Keller, H., Morelli, G., Gressmann, H., Achtman, M., Schuster, S.C., 2006. Who ate whom? Adaptive *Helicobacter* genomic changes that accompanied a host jump from early humans to large felines. PLoS Genet. 2 (e120.M).

Felsenfeld, O., Wolf, R.H., 1971. Serological reactions with treponemal antigens in non-human primates and the natural history of treponematosis in man. Folia Primatol. (Basel) 16, 294–305.

Flasarová, M., Šmajs, D., Matějková, P., Woznicová, V., Heroldová-Dvořáková, M., Votava, M., 2006. Molecular detection and subtyping of *Treponema pallidum* subsp. *pallidum* in clinical specimens. Epidemiol. Mikrobiol. Imunol. 55, 105–111.

Flasarová, M., Pospíšilová, P., Mikalová, L., Vališová, Z., Dastychová, E., Strnadel, R.V., Kuklová, I., Woznicová, V., Zákoucká, H., Šmajs, D., 2012. Sequencing-based molecular typing of *Treponema pallidum* strains in the Czech Republic: all identified genotypes are related to the sequence of the SS14 strain. Acta Derm. Venereol. 92, 669–674.

Flores, J.A., Vargas, S.K., Leon, S.R., Perez, D.G., Ramos, L.B., Chow, J., Konda, K.A.,

Calvo, G.M., Salvatierra, H.J., Klaussner, J.D., Caceres, C.F., 2016. *Treponema pallidum pallidum* genotypes and macrolide resistance status in syphilitic lesions among patients at 2 sexually transmitted infection clinics in Lima, Peru. Sex. Transm. Dis. 43, 465–466.

Florindo, C., Reigado, V., Gomes, J.P., Azevedo, J., Santo, I., Borrego, M.J., 2008. Molecular typing of *Treponema pallidum* clinical strains from Lisbon, Portugal. J. Clin. Microbiol. 46, 3802–3803.

Ford, C.B., Lin, P.L., Chase, M.R., Shah, R.R., Iartchouk, O., Galagan, J., Mohaideen, N., Ioerger, T.R., Sacchettini, J.C., Lipsitch, M., Flynn, J.L., Fortune, S.M., 2011. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. Nat. Genet. 43, 482–486.

Fraser, C.M., Norris, S.J., Weinstock, G.M., White, O., Sutton, G.G., Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A., Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J., Khalak, H., Richardson, D., Howell, J.K., Chidambaram, M., Utterback, T., MacDonald, L., Artiach, P., Bowman, C., Cotton, M.D., Fujii, C., Garland, S., Hatch, B., Horst, K., Roberts, K., Sandusky, M., Weidman, J., Smith, H.O., Venter, J.C., 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science 281, 375–388.

Fribourg-Blanc, A., Mollaret, H.H., 1969. Natural treponematosis of the African primate. Primates Med. 3, 113–121.

Fribourg-Blanc, A., Niel, G., Mollaret, H.H., 1963. Note on some immunological aspects of the African *Cynocephalus*. 1. Antigenic relationship of its gamma globulin with human gamma globulin. 2. Guinean endemic focus of treponematosis. Bull. Soc. Pathol. Exot. Filiales 56, 474–485.

Fribourg-Blanc, A., Mollaret, H.H., Niel, G., 1966. Serologic and microscopic confirmation of treponemosis in Guinea baboons. Bull. Soc. Pathol. Exot. Filiales 59, 54–59.

Gallo Vaulet, L., Grillová, L., Mikalová, L., Casco, R., Rodríguez Fermepin, M., Pando, M.A., Šmajs, D., 2017. Molecular typing of *Treponema pallidum* isolates from Buenos Aires, Argentina: frequent Nichols-like isolates and low levels of macrolide resistance. PLoS One 12, e0172905.

Gastinel, P., Vaisman, A., Hamelin, A., Dunoyer, F., 1963. Study of a recently isolated strain of *Treponema pertenue*. Ann. Dermatol. Syphiligr. Paris 90, 155–161.

Gemayel, R., Vinces, M.D., Legendre, M., Verstrepen, K.J., 2010. Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu. Rev. Genet. 44, 445–477.

Giacani, L., Lukehart, S.A., 2014. The endemic treponematoses. Clin. Microbiol. Rev. 27, 89–115.

Giacani, L., Lukehart, S., Centurion-Lara, A., 2007. Length of guanosine homopolymeric repeats modulates promoter activity of subfamily II tpr genes of *Treponema pallidum* ssp. *pallidum*. FEMS Immunol. Med. Microbiol. 51, 289–301.

Giacani, L., Jeffrey, B.M., Molini, B.J., Le, H.T., Lukehart, S.A., Centurion-Lara, A., Rockey, D.D., 2010. Complete genome sequence and annotation of the *Treponema pallidum* subsp *pallidum* Chicago strain. J. Bacteriol. 192, 2645–2646.

Giacani, L., Brandt, S.L., Puray-Chavez, M., Reid, T.B., Godornes, C., Molini, B.J., Benzler, M., Hartig, J.S., Lukehart, S.A., Centurion-Lara, A., 2012. Comparative investigation of the genomic regions involved in antigenic variation of the TprK antigen among treponemal species, subspecies, and strains. J. Bacteriol. 194, 4208–4225.

Giacani, L., Iverson-Cabral, S.L., King, J.C., Molini, B.J., Lukehart, S.A., Centurion-Lara, A., 2014. Complete Genome Sequence of the *Treponema pallidum* subsp. *pallidum* Sea81-4 Strain. Genome Announc. 2, e00333-14.

Giacani, L., Brandt, S.L., Ke, W., Reid, T.B., Molini, B.J., Iverson-Cabral, S., Ciccarese, G., Drago, F., Lukehart, S.A., Centurion-Lara, A., 2015. Transcription of TP0126, *Treponema pallidum* putative OmpW homolog, is regulated by the length of a homopolymeric guanosine repeat. Infect. Immun. 83, 2275–2289.

Giacani, L., Ciccarese, G., Puga-Salazar, C., Dal Conte, I., Colli, L., Cusini, M., Ramoni, S., Delmonte, S., D'Antuono, A., Gaspari, V., Drago, F., 2018. Ehanced Molecular Typing of *Treponema pallidum subsp. pallidum* strains from 4 Italian hospitals shows geographical differences in strain type heterogeneity, widespread resistance to macrolides, and lack of mutations associated with doxycycline resistance. Sex. Transm. Dis. 45, 237–242.

Gogarten, J.F., Dux, A., Schuenemann, V.J., Nowak, K., Boesch, C., Wittig, R.M., Krause, J., Calvignac-Spencer, S., Leendertz, F.H., 2016. Tools for opening new chapters in the book of *Treponema pallidum* evolutionary history. Clin. Microbiol. Infect. 22, 916–921.

Golubchik, T., Batty, E.M., Miller, R.R., Farr, H., Young, B.C., Larner-Svensson, H., Fung, R., Godwin, H., Knox, K., Votintseva, A., Everitt, R.G., Street, T., Cule, M., Ip, C.L., Didelot, X., Peto, T.E., Harding, R.M., Wilson, D.J., Crook, D.W., Bowden, R., 2013. Within-host evolution of *Staphylococcus aureus* during asymptomatic carriage. PLoS One 8, e61319.

Grange, P.A., Allix-Beguec, C., Chanal, J., Benhaddou, N., Gerhardt, P., Morini, J.P., Deleuze, J., Lassau, F., Janier, M., Dupin, N., 2013. Molecular subtyping of *Treponema pallidum* in Paris, France. Sex. Transm. Dis. 40, 641–644.

Grange, P.A., Mikalová, L., Gaudin, C., Strouhal, M., Janier, M., Benhaddou, N., Šmajs, D., Dupin, N., 2016. *Treponema pallidum* 11q subtype may correspond to a *Treponema pallidum* subsp. *endemicum* strain. Sex. Transm. Dis. 43, 517–518.

Gray, R.R., Mulligan, C.J., Molini, B.J., Sun, E.S., Giacani, L., Godornes, C., Kitchen, A., Lukehart, S.A., Centurion-Lara, A., 2006. Molecular evolution of the *tprC, D, I, K, G*, and *J* genes in the pathogenic genus *Treponema*. Mol. Biol. Evol. 23, 2220–2233.

Grillová, L., Pětrošová, H., Mikalová, L., Strnadel, R., Dastychová, E., Kuklová, I., Kojanová, M., Kreidlová, M., Vaňousová, D., Hercogová, J., Procházka, P., Zákoucká, H., Krchňáková, A., Vašků, V., Šmajs, D., 2014. Molecular typing of *Treponema pallidum* in the Czech Republic during 2011 to 2013: increased prevalence of identified genotypes and of isolates with macrolide resistance. J. Clin. Microbiol. 52, 3693–3700.

Grimes, M., Sahi, S.K., Godornes, B.C., Tantalo, L.C., Roberts, N., Bostick, D., Marra, C.M., Lukehart, S.A., 2012. Two mutations associated with macrolide resistance in *Treponema pallidum*: increasing prevalence and correlation with molecular strain type in Seattle, Washington. Sex. Transm. Dis. 39, 954–958.

Hallmaier-Wacker, L.K., Munster, V.J., Knauf, S., 2017. Disease reservoirs: from conceptual frameworks to applicable criteria. Emerg. Microbes Infect. 6, e79.

Harper, K.N., Knauf, S., 2013. *Treponema pallidum* infection in primates: clinical manifestations, epidemiology, and evolution of a stealthy pathogen. In: Brinkworth, J.F., Pechenkina, K. (Eds.), Primates, Pathogens, and Evolution. Developments in Primatology: Progress and Prospects. 38. Springer, New York, pp. 189–219.

Harper, K.N., Liu, H., Ocampo, P.S., Steiner, B.M., Martin, A., Levert, K., Wang, D., Sutton, M., Armelagos, G.J., 2008. The sequence of the acidic repeat protein (*arp*) gene differentiates venereal from nonvenereal *Treponema pallidum* subspecies, and the gene has evolved under strong positive selection in the subspecies that causes syphilis. FEMS Immunol. Med. Microbiol. 53, 322–332.

Harper, K.N., Fyumagwa, R.D., Hoare, R., Wambura, P.N., Coppenhaver, D.H., Sapolsky, R.M., Alberts, S.C., Tung, J., Rogers, J., Kilewo, M., Batamuzi, E.K., Leendertz, F.H., Armelagos, G.J., Knauf, S., 2012. *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS One 7, e50882.

Hazlett, K.R., Sellati, T.J., Nguyen, T.T., Cox, D.L., Clawson, M.L., Caimano, M.J., Radolf, J.D., 2001. The TprK protein of *Treponema pallidum* is periplasmic and is not a target of opsonic antibody or protective immunity. J. Exp. Med. 193, 1015–1026.

Heymans, R., Kolader, M.E., van der Helm, J.J., Coutinho, R.A., Bruisten, S.M., 2009. TprK gene regions are not suitable for epidemiological syphilis typing. Eur. J. Clin. Microbiol. Infect. Dis. 28, 875–878.

Ho, E.L., Lukehart, S.A., 2011. Syphilis: using modern approaches to understand an old disease. J. Clin. Invest. 121, 4584–4592.

Horvath, I., Kemenes, F., Molnar, L., Szeky, A., Racz, I., 1980. Experimental syphilis and serological examination for treponematosis in hares. Infect. Immun. 27, 231–234.

Jacobsthal, E., 1920. Untersuchungen über eine syphilisähnliche Spontanerkrankungen des Kaninchens (*Paralues cuniculi*). Derm. Wschr. 71, 569–571.

Jakšić, B.L., 1957. Brucelose et syphilis du lievre. Vet. Glas. 11, 423–425.

Karesh, W., 2000. Suivi de la santé des gorilles au Nord-Congo. Canopée 18, 16–17.

Katz, K.A., Pillay, A., Ahrens, K., Kohn, R.P., Hermanstyne, K., Bernstein, K.T., Ballard, R.C., Klausner, J.D., 2010. Molecular epidemiology of syphilis-San Francisco, 2004–2007. Sex. Transm. Dis. 37, 660–663.

Khairullin, R., Vorobyev, D., Obukhov, A., Kuular, U.H., Kubanova, A., Kubanov, A., Unemo, M., 2016. Syphilis epidemiology in 1994–2013, molecular epidemiological strain typing and determination of macrolide resistance in *Treponema pallidum* in 2013-2014 in Tuva Republic, Russia. APMIS 124, 595–602.

Klegarth, A.R., Ezeonwu, C.A., Rompis, A., Lee, B.P.Y., Aggimarangsee, N., Chalise, M., Cortes, J., Feeroz, M., Molini, B.J., Godornes, B.C., Marks, M., Schillaci, M., Engel, G., Knauf, S., Lukehart, S.A., Jones-Engel, L., 2017. Survey of Treponemal Infections in Free-Ranging and Captive Macaques, 1999-2012. Emerg. Infect. Dis. 23, 816–819.

Knauf, S., Batamuzi, E.K., Mlengeya, T., Kilewo, M., Lejora, I.A., Nordhoff, M., Ehlers, B., Harper, K.N., Fyumagwa, R., Hoare, R., Failing, K., Wehrend, A., Kaup, F.J., Leendertz, F.H., Matz-Rensing, K., 2012. *Treponema* infection associated with genital ulceration in wild baboons. Vet. Pathol. 49, 292–303.

Knauf, S., Liu, H., Harper, K.N., 2013. Treponemal infection in nonhuman primates as possible reservoir for human yaws. Emerg. Infect. Dis. 19, 2058–2060.

Knauf, S., Raphael, J., Mitja, O., Lejora, I.A., Chuma, I.S., Batamuzi, E.K., Keyyu, J.D., Fyumagwa, R., Luert, S., Godornes, C., Liu, H., Schwarz, C., Šmajs, D., Grange, P., Zinner, D., Roos, C., Lukehart, S.A., 2016. Isolation of treponema DNA from necrophagous flies in a natural ecosystem. EBioMedicine 11, 85–90.

Knauf, S., Gogarten, J., Schuenemann, V.J., De Nys, H.M., Duex, A., Strouhal, M., Mikalova, L., Bos, K.I., Armstrong, R., Batamuzi, E.K., Chuma, I.S., Davoust, B., Diatta, G., Fyumagwa, R.D., Kazwala, R.R., Keyyu, J.D., Lejora, I.A.V., Levasseur, A., Liu, H., Mayhew, M.A., Mediannikov, O., Raoult, D., Wittig, R.M., Roos, C., Leendertz, F.H., Smajs, D., Nieselt, K., Krause, J., Calvignac-Spencer, S., 2017. African Nonhuman Primates Are InfectedAbs With The Yaws Bacterium Treponema pallidum subsp. pertenue. (bioRxiv).

Krieg, N.R., Staley, J.T., Brown, D.R., Hedlund, B.P., Paster, B.J., Ward, N.L., Ludwig, W., Whitman, W.B., 2010. Bergey's Manual of Systematic Bacteriology, second ed. Springer, New York, pp. 501–531.

Kuhn, U.S.G., 1970. The Treponematoses. The Chimpanzee, Karger, Basel, München, New York.

Kuhn III, U.S., Medina, R., Cohen, P.G., Vegas, M., 1970. Inoculation pinta in chimpanzees. Br. J. Vener. Dis. 46, 311–312.

LaFond, R.E., Centurion-Lara, A., Godornes, C., Rompalo, A.M., Van Voorhis, W.C., Lukehart, S.A., 2003. Sequence diversity of Treponema pallidum subsp. pallidum tprK in human syphilis lesions and rabbit-propagated isolates. J. Bacteriol. 185, 6262–6268.

LaFond, R.E., Centurion-Lara, A., Godornes, C., Van Voorhis, W.C., Lukehart, S.A., 2006. TprK sequence diversity accumulates during infection of rabbits with *Treponema pallidum* subsp. *pallidum* Nichols strain. Infect. Immun. 74, 1896–1906.

Leader, B.T., Hevner, K., Molini, B.J., Barrett, L.K., Van Voorhis, W.C., Lukehart, S.A., 2003. Antibody responses elicited against the *Treponema pallidum* repeat proteins differ during infection with different isolates of *Treponema pallidum* subsp. *pallidum*. Infect. Immun. 71, 6054–6057.

Levine, B.M., Smith, J.L., Israel, C.W., 1970. Serology of normal primates. Br. J. Vener. Dis. 46, 307–310.

Levrero, F., Gatti, S., Gautier-Hion, A., Menard, N., 2007. Yaws disease in a wild gorilla population and its impact on the reproductive status of males. Am. J. Phys. Anthropol. 132, 568–575.

Li, Z., Wang, C., Xiao, H., Zhao, W., Li, Z., Zheng, R., Hou, J., Huang, N., Tian, H., 2017. Enhanced molecular typing of *Treponema pallidum* identified a new tp0548 gene type in Shandong, China. APMIS 125, 937–939.

Liska, S.L., Perine, P.L., Hunter, E.F., Crawford, J.A., Feeley, J.C., 1982. Isolation and transportation of *Treponema pertenue* in golden hamsters. Curr. Microbiol. 7, 41–43.
Lovell, N., Jurmain, R., Kilgore, L., 2000. Skeletal evidence of probable treponemal infection in free-ranging African apes. Primates 41, 275–290.
Lukehart, S.A., Shaffer, J.M., Baker-Zander, S.A., 1992. A subpopulation of *Treponema pallidum* is resistant to phagocytosis: possible mechanism of persistence. J. Infect. Dis. 166, 1449–1453.
Lumeij, J.T., 2010. Widespread treponemal infections of hare populations (*Lepus europaeus*) in the Netherlands. Eur. J. Wildl. Res. 57, 183–186.
Lumeij, J.T., de Koning, J., Bosma, R.B., van der Sluis, J.J., Schellekens, J.F., 1994. Treponemal infections in hares in The Netherlands. J. Clin. Microbiol. 32, 543–546.
Lumeij, J.T., Mikalová, L., Šmajs, D., 2013. Is there a difference between hare syphilis and rabbit syphilis? Cross infection experiments between rabbits and hares. Vet. Microbiol. 164, 190–194.
Magnuson, H.J., Eagle, H., Fleischmann, R., 1948. The minimal infectious inoculum of *Spirochaeta pallida* (Nichols strain), and a consideration of its rate of multiplication *in vivo*. Am. J. Syph. Gonorrhea Vener. Dis. 32, 1–18.
Marks, M., 2016. Yaws: towards the WHO eradication target. Trans. R. Soc. Trop. Med. Hyg. 110, 319–320.
Marks, M., Mitja, O., Vestergaard, L.S., Pillay, A., Knauf, S., Chen, C.Y., Bassat, Q., Martin, D.L., Fegan, D., Taleo, F., Kool, J., Lukehart, S., Emerson, P.M., Solomon, A.W., Ye, T., Ballard, R.C., Mabey, D.C., Asiedu, K.B., 2015. Challenges and key research questions for yaws eradication. The Lancet Infect. Dis. 15, 1220–1225.
Marks, M., Mitja, O., Fitzpatrick, C., Asiedu, K., Solomon, A.W., Mabey, D.C., Funk, S., 2017. Mathematical Modeling of Programmatic Requirements for Yaws Eradication. Emerg. Infect. Dis. 23, 22–28.
Marra, C.M., Castro, C.D., Kuller, L., Dukes, A.C., Centurion-Lara, A., Morton, W.R., Lukehart, S.A., 1998. Mechanisms of clearance of *Treponema pallidum* from the CSF in a nonhuman primate model. Neurology 51, 957–961.
Marra, C.M., Sahi, S.K., Tantalo, L.C., Godornes, C., Reid, T., Behets, F., Rompalo, A., Klausner, J.D., Yin, Y.P., Mulcahy, F., Golden, M.R., Centurion-Lara, A., Lukehart, S.A., 2010. Enhanced molecular typing of *Treponema pallidum*: geographical distribution of strain types and association with neurosyphilis. J. Infect. Dis. 202, 1380–1388.
Martin, I.E., Gu, W., Yang, Y., Tsang, R.S., 2009. Macrolide resistance and molecular types of *Treponema pallidum* causing primary syphilis in Shanghai, China. Clin. Infect. Dis. 49, 515–521.
Martin, I.E., Tsang, R.S., Sutherland, K., Anderson, B., Read, R., Roy, C., Yanow, S., Fonseca, K., White, W., Kandola, K., Kouadjo, E., Singh, A.E., 2010. Molecular typing of *Treponema pallidum* strains in western Canada: predominance of 14d subtypes. Sex. Transm. Dis. 37, 544–548.
Matějková, P., Strouhal, M., Šmajs, D., Norris, S.J., Palzkill, T., Petrosino, J.F., Sodergren, E., Norton, J.E., Singh, J., Richmond, T.A., Molla, M.N., Albert, T.J., Weinstock, G.M., 2008. Complete genome sequence of *Treponema pallidum* ssp. *pallidum* strain SS14 determined with oligonucleotide arrays. BMC Microbiol. 8, 76.
Matějková, P., Flasarová, M., Zákoucká, H., Bořek, M., Křemenová, S., Arenberger, P., Woznicová, V., Weinstock, G.M., Šmajs, D., 2009. Macrolide treatment failure in a case of secondary syphilis: a novel A2059G mutation in the 23S rRNA gene of *Treponema pallidum* subsp. *pallidum*. J. Med. Microbiol. 58, 832–836.
de Mello, F.L., de Mello, J.C., Fraga, A.M., Nunes, K., Eggers, S., 2010. Syphilis at the crossroad of phylogenetics and paleopathology. PLoS Negl. Trop. Dis. 4, e575.
Metchnikoff, E., Roux, E., 1903. Études experimentales sur la syphilis. Ann. Inst. Pasteur 17, 808–821.
Metchnikoff, E., Roux, E., 1904. Études experimentales sur la syphilis. Ann. Inst. Pasteur 18, 1–6.
Metchnikoff, E., Roux, E., 1905. Études expÉrimentales sur la syphilis. Ann. Inst. Pasteur 19, 673–698.
Mikalová, L., Pospíšilová, P., Woznicová, V., Kuklová, I., Zákoucká, H., Šmajs, D., 2013. Comparison of CDC and sequence-based molecular typing of syphilis treponemes: *tpr* and *arp* loci are variable in multiple samples from the same patient. BMC Microbiol. 13, 178.
Mikalová, L., Strouhal, M., Oppelt, J., Grange, P.A., Janier, M., Benhaddou, N., Dupin, N., Šmajs, D., 2017a. Human *Treponema pallidum* 11q/j isolate belongs to subsp. *endemicum* but contains two loci with a sequence in TP0548 and TP0488 similar to subsp. *pertenue* and subsp. *pallidum*, respectively. PLoS Negl. Trop. Dis. 11, e0005434.
Mikalová, L., Grillová, L., Osbak, K., Strouhal, M., Kenyon, C., Crucitti, T., Šmajs, D., 2017b. Molecular typing of syphilis-causing strains among human immunodeficiency virus-positive patients in Antwerp, Belgium. Sex. Transm. Dis. 44, 376–379.
Mitja, O., Hays, R., Rinaldi, A.C., McDermott, R., Bassat, Q., 2012. New treatment schemes for yaws: the path toward eradication. Clin. Infect. Dis. 55, 406–412.
Mitja, O., Asiedu, K., Mabey, D., 2013. Yaws. Lancet 381, 763–773.
Mitja, O., Lukehart, S.A., Pokowas, G., Moses, P., Kapa, A., Godornes, C., Robson, J., Cherian, S., Houinei, W., Kazadi, W., Siba, P., de Lazzari, E., Bassat, Q., 2014. *Haemophilus ducreyi* as a cause of skin ulcers in children from a yaws-endemic area of Papua New Guinea: a prospective cohort study. Lancet Glob. Health 2, e235–241.
Mitja, O., Houinei, W., Moses, P., Kapa, A., Paru, R., Hays, R., Lukehart, S., Godornes, C., Bieb, S.V., Grice, T., Siba, P., Mabey, D., Sanz, S., Alonso, P.L., Asiedu, K., Bassat, Q., 2015. Mass treatment with single-dose azithromycin for yaws. N. Engl. J. Med. 372, 703–710.
Mlengeya, T.D.K., 2004. Distribution pattern of a sexually transmitted disease (STD) of Olive Baboon in Lake Manyara National Park, Tanzania. College of African Wildlife Management, Moshi.
Molepo, J., Pillay, A., Weber, B., Morse, S.A., Hoosen, A.A., 2007. Molecular typing of *Treponema pallidum* strains from patients with neurosyphilis in Pretoria, South Africa. Sex. Transm. Infect. 83, 189–192.
Mörner, T., 1999. Monitoring disease in wildlife – A review of diseases in the orders lagomorpha and rodentia in Sweden. Verh. ber Erkrg. Zootiere 39, 255–262.
Mossoun, A., Pauly, M., Akoua-Koffi, C., Couacy-Hymann, E., Leendertz, S.A., Anoh, A.E., Gnoukpoho, A.H., Leendertz, F.H., Schubert, G., 2015. Contact to non-human primates and risk factors for zoonotic disease emergence in the tai region, Cote d'Ivoire. EcoHealth 12, 580–591.
Moxon, R., Bayliss, C., Hood, D., 2006. Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. Annu. Rev. Genet. 40, 307–333.
Müller, E.E., Paz-Bailey, G., Lewis, D.A., 2012. Macrolide resistance testing and molecular subtyping of *Treponema pallidum* strains from southern Africa. Sex. Transm. Infect. 88, 470–474.
Naqvi, A.A., Shahbaaz, M., Ahmad, F., Hassan, M.I., 2015. Identification of functional candidates amongst hypothetical proteins of *Treponema pallidum* ssp. *pallidum*. PLoS One 10, e0124177.
Narain, J.P., Jain, S.K., Bora, D., Venkatesh, S., 2015. Eradicating successfully yaws from India: The strategy & global lessons. Indian J. Med. Res. 141, 608–613.
Nichols, H.J., 1910. Experimental yaws in the monkey and rabbit. J. Exp. Med. 12, 616–622.
Nichols, H.J., Hough, W.H., 1913. Demonstration of *Spirochaeta pallida* in the cerebrospinal fluid. JAMA-J. Am. Med. Assoc. 60, 108–110.
Noda, A.A., Matos, N., Blanco, O., Rodríguez, I., Stamm, L.V., 2016. First report of the 23S rRNA gene A2058G point mutation associated with macrolide resistance in *Treponema pallidum* from syphilis patients in Cuba. Sex. Transm. Dis. 43, 332–334.
Noda, A.A., Grillová, L., Lienhard, R., Blanco, O., Rodríguez, I., Šmajs, D., 2018. Bejel in Cuba: molecular identification of *Treponema pallidum* subsp. *endemicum* in patients diagnosed with venereal syphilis. Clin. Microbiol. Infect. http://dx.doi.org/10.1016/j.cmi.2018.02.006.
Noordhoek, G.T., Engelkens, H.J., Judanarso, J., van der Stek, J., Aelbers, G.N., van der Sluis, J.J., van Embden, J.D., Stolz, E., 1991. Yaws in West Sumatra, Indonesia: clinical manifestations, serological findings and characterisation of new *Treponema* isolates by DNA probes. Eur. J. Clin. Microbiol. Infect. Dis. 10, 12–19.
Ochman, H., Elwyn, S., Moran, N.A., 1999. Calibrating bacterial evolution. Proc. Natl. Acad. Sci. U. S. A. 96, 12638–12643.
Pallen, M.J., Wren, B.W., 2007. Bacterial pathogenomics. Nature 449, 835–842.
Palmer, G.H., Bankhead, T., Lukehart, S.A., 2009. "Nothing is permanent but change"-antigenic variation in persistent bacterial pathogens. Cell. Microbiol. 11, 1697–1705.
Peng, R.R., Wang, A.L., Li, J., Tucker, J.D., Yin, Y.P., Chen, X.S., 2011. Molecular typing of *Treponema pallidum*: a systematic review and meta-analysis. PLoS Negl. Trop. Dis. 5, e1273.
Peng, R.R., Yin, Y.P., Wei, W.H., Wang, H.C., Zhu, B.Y., Liu, Q.Z., Zheng, H.P., Zhang, J.P., Huang, S.J., Chen, X.S., 2012. Molecular typing of *Treponema pallidum* causing early syphilis in China: a cross-sectional study. Sex. Transm. Dis. 39, 42–45.
Pětrošová, H., Zobaníková, M., Čejková, D., Mikalová, L., Pospíšilová, P., Strouhal, M., Chen, L., Qin, X., Muzny, D.M., Weinstock, G.M., Šmajs, D., 2012. Whole genome sequence of *Treponema pallidum* ssp. *pallidum*, strain Mexico A, suggests recombination between yaws and syphilis strains. PLoS Negl. Trop. Dis. 6, e1832.
Pětrošová, H., Pospíšilová, P., Strouhal, M., Čejková, D., Zobaníková, M., Mikalová, L., Sodergren, E., Weinstock, G.M., Šmajs, D., 2013. Resequencing of *Treponema pallidum* ssp. *pallidum* strains Nichols and SS14: correction of sequencing errors resulted in increased separation of syphilis treponeme subclusters. PLoS One 8, e74319.
Pillay, A., Liu, H., Chen, C.Y., Holloway, B., Sturm, A.W., Steiner, B., Morse, S.A., 1998. Molecular subtyping of *Treponema pallidum* subspecies *pallidum*. Sex. Transm. Dis. 25, 408–414.
Pillay, A., Liu, H., Ebrahim, S., Chen, C.Y., Lai, W., Fehler, G., Ballard, R.C., Steiner, B., Sturm, A.W., Morse, S.A., 2002. Molecular typing of *Treponema pallidum* in South Africa: Cross-Sectional Studies. J. Clin. Microbiol. 40, 256–258.
Pinto, M., Borges, V., Antelo, M., Pinheiro, M., Nunes, A., Azevedo, J., Borrego, M.J., Mendonça, J., Carpinteiro, D., Vieira, L., Gomes, J.P., 2016. Genome-scale analysis of the non-cultivable *Treponema pallidum* reveals extensive within-patient genetic variation. Nat. Microbiol. 2, 16190.
Plummer, P.J., Krull, A., 2017. Clinical perspectives of digital dermatitis in dairy and beef cattle. Vet. Clin. North Am. Food Anim. Pract. 33, 165–181.
Pope, V., Fox, K., Liu, H., Marfin, A.A., Leone, P., Seña, A.C., Chapin, J., Fears, M.B., Markowitz, L., 2005. Molecular subtyping of *Treponema pallidum* from North and South Carolina. J. Clin. Microbiol. 43, 3743–3746.
Posautz, A., Leidinger, E., Knauer, F., Hoffman, D., Suchentrunk, F., Walzer, C., Kubber-Heiss, A., 2014. Seroprevalence of *Treponema* sp. in European brown hares (*Lepus europaeus*) in Austria and Germany. Wien. Tierarztl. Monat. 101, 281–285.
Radolf, J.D., Kumar, S., 2017. The *Treponema pallidum* outer membrane. Curr. Top. Microbiol. Immunol. http://dx.doi.org/10.1007/82_2017_44. in press.
Radolf, J.D., Deka, R.K., Anand, A., Šmajs, D., Norgard, M.V., Yang, X.F., 2016. *Treponema pallidum*, the syphilis spirochete: making a living as a stealth pathogen. Nat. Rev. Microbiol. 14, 744–759.
Read, P., Tagg, K.A., Jeoffreys, N., Guy, R.J., Gilbert, G.L., Donovan, B., 2016. *Treponema pallidum* strain types and association with macrolide resistance in Sydney, Australia: new TP0548 gene types identified. J. Clin. Microbiol. 54, 2172–2174.
Richard, M., Knauf, S., Lawrence, P., Mather, A.E., Munster, V.J., Muller, M.A., Smith, D., Kuiken, T., 2017. Factors determining human-to-human transmissibility of zoonotic pathogens via contact. Curr. Opin. Virol. 22, 7–12.
Salado-Rasmussen, K., Cowan, S., Gerstoft, J., Larsen, H.K., Hoffmann, S., Knudsen, T.B., Katzenstein, T.L., Jensen, J.S., 2016. Molecular typing of *Treponema pallidum* in Denmark: a nationwide study of syphilis. Acta Derm. Venereol. 96, 202–206.
Sepetjian, M., Guerraz, F.T., Salussola, D., Thivolet, J., Monier, J.C., 1969. Contribution to the study of the treponeme isolated from monkeys by A. Fribourg-Blanc. Bull. World Health Organ. 40, 141–151.
Sepetjian, M., Thivolet, J., Salussola, D., Tissot Guerraz, F., Monier, J.C., 1972. Comparative study of quantitative FTA and FTA ABS tests in man and animals at

various stages of syphilis. Pathol. Biol. (Paris) 20, 449–455.
Šmajs, D., McKevitt, M., Wang, L., Howell, J.K., Norris, S.J., Palzkill, T., Weinstock, G.M., 2002. BAC library of *T. pallidum* DNA in *E. coli*. Genome Res. 12, 515–522.
Šmajs, D., McKevitt, M., Howell, J.K., Norris, S.J., Cai, W.W., Palzkill, T., Weinstock, G.M., 2005. Transcriptome of *Treponema pallidum*: gene expression profile during experimental rabbit infection. J. Bacteriol. 187, 1866–1874.
Šmajs, D., Zobaníková, M., Strouhal, M., Čejková, D., Dugan-Rocha, S., Pospíšilová, P., Norris, S.J., Albert, T., Qin, X., Hallsworth-Pepin, K., Buhay, C., Muzny, D.M., Chen, L., Gibbs, R.A., Weinstock, G.M., 2011. Complete genome sequence of *Treponema paraluiscuniculi*, strain Cuniculi A: the loss of infectivity to humans is associated with genome decay. PLoS One 6, e20415.
Šmajs, D., Norris, S.J., Weinstock, G.M., 2012. Genetic diversity in *Treponema pallidum*: implications for pathogenesis, evolution and molecular diagnostics of syphilis and yaws. Infect. Genet. Evol. 12, 191–202.
Šmajs, D., Paštěková, L., Grillová, L., 2015. Macrolide resistance in the syphilis spirochete, *Treponema pallidum* ssp. *pallidum*: can we also expect macrolide-resistant yaws strains? Am. J. Trop. Med. Hyg. 93, 678–683.
Šmajs, D., Mikalová, L., Strouhal, M., Grillová, L., 2016. Why there are two genetically distinct syphilis-causing strains? Forum Immun. Dis. Ther. 7, 181–190.
Smith, J.L., Pesetsky, B.R., 1967. The current status of *Treponema cuniculi*: review of the literature. Br. J. Venereol. Dis. 43, 117–127.
Smith, J.L., David, N.J., Indgin, S., Israel, C.W., Levine, B.M., Justice Jr., J., McCrary 3rd, J.A., Medina, R., Paez, P., Santana, E., Sarkar, M., Schatz, N.J., Spitzer, M.L., Spitzer, W.O., Walter, E.K., 1971. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br. J. Vener. Dis. 47, 226–251.
Stadler, T., Kühnert, D., Bonhoeffer, S., Drummond, A.J., 2013. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). Proc. Natl. Acad. Sci. U. S. A. 110, 228–233.
Stamm, L.V., Bergen, H.L., 2000a. A point mutation associated with bacterial macrolide resistance is present in both 23S rRNA genes of an erythromycin-resistant *Treponema pallidum* clinical isolate. Antimicrob. Agents Chemother. 44, 806–807.
Stamm, L.V., Bergen, H.L., 2000b. The sequence-variable, single-copy *tprK* gene of *Treponema pallidum* Nichols strain UNC and Street strain 14 encodes heterogeneous TprK proteins. Infect. Immun. 68, 6482–6486.
Stamm, L.V., Kerner Jr., T.C., Bankaitis, V.A., Bassford Jr., P.J., 1983. Identification and preliminary characterization of *Treponema pallidum* protein antigens expressed in *Escherichia coli*. Infect. Immun. 41, 709–721.
Štaudová, B., Strouhal, M., Zobaníková, M., Čejková, D., Fulton, L.L., Chen, L., Giacani, L., Centurion-Lara, A., Bruisten, S.M., Sodergren, E., Weinstock, G.M., Šmajs, D., 2014. Whole genome sequence of the *Treponema pallidum* subsp. *endemicum* strain Bosnia A: the genome is related to yaws treponemes but contains few loci similar to syphilis treponemes. PLoS Negl. Trop. Dis. 8, e3261.
Strouhal, M., Šmajs, D., Matějková, P., Sodergren, E., Amin, A.G., Howell, J.K., Norris, S.J., Weinstock, G.M., 2007. Genome differences between *Treponema pallidum* subsp. *pallidum* strain Nichols and *T. paraluiscuniculi* strain Cuniculi A. Infect. Immun. 75, 5859–5866.
Strouhal, M., Mikalová, L., Havlíčková, P., Tenti, P., Čejková, D., Rychlík, I., Bruisten, S., Šmajs, D., 2017. Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Ghana, Africa: Identical genome sequences in samples isolated more than 7 years apart. PLoS Negl. Trop. Dis. 11, e0005894.
Strouhal, M., Oppelt, J., Mikalová, L., Arora, N., Nieselt, K., González-Candelas, F., Šmajs, D., 2018. Reanalysis of Chinese *Treponema pallidum* samples: all Chinese samples cluster with SS14-like group of syphilis-causing treponemes. BMC Res. Notes. 11, 16.
Sun, J., Meng, Z., Wu, K., Liu, B., Zhang, S., Liu, Y., Wang, Y., Zheng, H., Huang, J., Zhou, P., 2016. Tracing the origin of *Treponema pallidum* in China using next-generation sequencing. Oncotarget 7, 42904–42918.
Sutton, M.Y., Liu, H., Steiner, B., Pillay, A., Mickey, T., Finelli, L., Morse, S., Markowitz, L.E., St Louis, M.E., 2001. Molecular subtyping of *Treponema pallidum* in an Arizona County with increasing syphilis morbidity: use of specimens from ulcers and blood. J. Infect. Dis. 183, 1601–1606.
The World Health Organization, 2012. Eradication of yaws-the Morges strategy. In: Releve epidemiologique hebdomadaire/Section d'hygiene du Secretariat de la Societe des Nations, Weekly epidemiological record/Health Section of the Secretariat of the League of Nations. 87. pp. 189–194.
Tian, H., Li, Z., Li, Z., Hou, J., Zheng, R., Li, F., Liu, R., Liu, B., Wang, C., Zhang, F., 2014. Molecular typing of *Treponema pallidum*: identification of a new sequence of Tp0548 gene in Shandong, China. Sex. Transm. Dis. 41, 551.
Tipple, C., Graham, P.T., 2015. Syphilis testing, typing, and treatment follow-up: a new era for an old disease. Curr. Opin. Infect. Dis. 28, 53–60.
Tipple, C., McClure, M.O., Graham, P.T., 2011. High prevalence of macrolide resistant *Treponema pallidum* strains in a London centre. Sex. Transm. Infect. 87, 486–488.
Tong, M.L., Zhao, Q., Liu, L.L., Zhu, X.Z., Gao, K., Zhang, H.L., Lin, L.R., Niu, J.J., Ji, Z.L., Yang, T.C., 2017a. Whole genome sequence of the *Treponema pallidum* subsp. pallidum strain Amoy: An Asian isolate highly similar to SS14. PLoS One 12, e0182768.
Tong, M.L., Zhang, H.L., Zhu, X.Z., Fan, J.Y., Gao, K., Lin, L.R., Liu, L.L., Li, S.L., Lin, H.L., Lin, Z.F., Niu, J.J., Zheng, W.H., Yang, T.C., 2017b. Re-evaluating the sensitivity of the rabbit infectivity test for *Treponema pallidum* in modern era. Clin. Chim. Acta 464, 136–141.
Turner, T.B., Hollander, D.H., 1957. Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. Monogr. Ser. World Health Organ. 35, 3–266.
Varela, G., 1969. Experimental pinta in the chimpanzee. Salud Publica Mex. 11, 37–38.
Verin, R., Pestelli, M., Poli, A., 2012. Treponemal infection in free-ranging European brown hares (*Lepus europaeus*) in central Italy: serology and epidemiology. J. Wildl. Dis. 48, 1079–1082.
Visser, M.B., Ellen, R.P., 2011. New insights into the emerging role of oral spirochaetes in periodontal disease. Clin. Microbiol. Infect. 17, 502–512.
Wallis, J., 2000. Prevention of disease transmission in primate conservation. Ann. N. Y. Acad. Sci. 916, 691–693.
Wallis, J., Lee, D.R., 1999. Primate Conservation: the Prevention of Disease Transmission. Int. J. Primatol. 20, 803–826.
Weinstock, G.M., Šmajs, D., Hardham, J., Norris, S.J., 2000. From microbial genome sequence to applications. Res. Microbiol. 151, 151–158.
Wendel Jr., G.D., Sanchez, P.J., Peters, M.T., Harstad, T.W., Potter, L.L., Norgard, M.V., 1991. Identification of *Treponema pallidum* in amniotic fluid and fetal blood from pregnancies comlicated by congenital syphilis. Obstet. Gynecol. 78, 890–895.
van der Woude, M.W., Bäumler, A.J., 2004. Phase and antigenic variation in bacteria. Clin. Microbiol. Rev. 17, 581–611.
Woznicová, V., Šmajs, D., Wechsler, D., Matějková, P., Flasarová, M., 2007. Detection of *Treponema pallidum* subsp. *pallidum* from skin lesions, serum, and cerebrospinal fluid in an infant with congenital syphilis after clindamycin treatment of the mother during pregnancy. J. Clin. Microbiol. 45, 659–661.
Wu, H., Chang, S.Y., Lee, N.Y., Huang, W.C., Wu, B.R., Yang, C.J., Liang, S.H., Lee, C.H., Ko, W.C., Lin, H.H., Chen, Y.H., Liu, W.C., Su, Y.C., Hsieh, C.Y., Wu, P.Y., Hung, C.C., 2012. Evaluation of macrolide resistance and enhanced molecular typing of *Treponema pallidum* in patients with syphilis in Taiwan: a prospective multicenter study. J. Clin. Microbiol. 50, 2299–2304.
Wu, B.R., Yang, C.J., Tsai, M.S., Lee, K.Y., Lee, N.Y., Huang, W.C., Wu, H., Lee, C.H., Chen, T.C., Ko, W.C., Lin, H.H., Lu, P.L., Chen, Y.H., Liu, W.C., Yang, S.P., Wu, P.Y., Su, Y.C., Hung, C.C., Chang, S.Y., 2014. Multicentre surveillance of prevalence of the 23S rRNA A2058G and A2059G point mutations and molecular subtypes of *Treponema pallidum* in Taiwan, 2009–2013. Clin. Microbiol. Infect. 20, 802–807.
Xiao, Y., Liu, S., Liu, Z., Xie, Y., Jiang, C., Xu, M., Zhao, F., Zeng, T., Yu, J., Wu, Y., 2016. Molecular subtyping and surveillance of resistance genes in *Treponema pallidum* DNA from patients with secondary and latent syphilis in Hunan, China. Sex. Transm. Dis. 43, 310–316.
Zhang, R.L., Wang, Q.Q., Zhang, J.P., Yang, L.J., 2017. Molecular subtyping of *Treponema pallidum* and associated factors of serofast status in early syphilis patients: identified novel genotype and cytokine marker. PLoS One 12, e0175477.
Zhu, B., Bu, J., Li, W., Zhang, J., Huang, G., Cao, J., Tang, Z., Gan, Q., Wei, P., 2016. High resistance to azithromycin in clinical samples from patients with sexually transmitted diseases in Guangxi Zhuang Autonomous Region, China. PLoS One 11, e0159787.
Zobaníková, M., Mikolka, P., Čejková, D., Pospíšilová, P., Chen, L., Strouhal, M., Qin, X., Weinstock, G.M., Šmajs, D., 2012. Complete genome sequence of *Treponema pallidum* strain DAL-1. Stand. Genomic Sci. 10, 12–21.
Zobaníková, M., Strouhal, M., Mikalová, L., Čejková, D., Ambrožová, L., Pospíšilová, P., Fulton, L.L., Chen, L., Sodergren, E., Weinstock, G.M., Šmajs, D., 2013. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl. Trop. Dis. 7, e2172.

7.2 Nonhuman Primates as a Potential Reservoir for Human Yaws Infection

In the late 1980s, sightings of genital ulcerated olive baboons (*P. anubis*) in Tanzania (Wallis and Lee, 1999) triggered new investigations in treponematoses of NHPs Despite the early evidence of *Treponema* infection in West African NHPs (as reviewed in (Harper et al., 2013)) the most comprehensive proof of naturally occurring *T. pallidum* infection came from Lake Manyara National Park in northern Tanzania (Harper et al., 2012; Knauf et al., 2011). Today, infection in Tanzanian monkeys is the best-characterized treponematosis in NHPs (Chuma et al., 2018; Harper et al., 2012; Knauf et al., 2018, 2015, 2011). Although the whole genome of the baboon infecting Fribourg-Blanc strain, an isolate that originates from West Africa (Fribourg-Blanc and Mollaret, 1969), shows already high genetic similarity with human yaws causing strains (Zobanikova et al., 2013), it took several more years until sufficient data from multiple whole genomes of *T. pallidum* strains of NHP origin became available that allowed to conclude that humans are not the exclusive host of the yaws bacterium (Knauf et al., 2018) (genetic similarity, chapter 4). An ethically questionable study conducted infection experiments and showed that the Fribourg-Blanc baboon strain can cause sustainable infection in humans (Smith et al., 1971) (functional similarity, chapter 4). There is therefore basic evidence that *T. pallidum* strains of NHP origin have zoonotic potential. It is, however, currently unknown whether human and NHP infecting strains are epidemiological connected.

7.2.1 Challenges and Key Research Questions for Yaws Eradication

The notion that African NHPs might play a role in the epidemiology of human yaws infection (Knauf et al., 2013) has led the World Health Organization to discuss this matter along with other challenges and key research questions during the ‘Third Consultation on Yaws Eradication’ in Geneva, Switzerland, 2017. The results of this meeting have been summarized and published in the following publication. The inclusion of NHPs as a potential reservoir and obstacle for human yaws eradication must be seen as a milestone, since previous eradication efforts, despite early reports of *T. pallidum* infection in NHPs (Fribourg-Blanc et al., 1963; Fribourg-Blanc and Mollaret, 1969), have not considered further research in this particular field as relevant to yaws eradication.

Challenges and key research questions for yaws eradication

Michael Marks, Oriol Mitjà, Lasse S Vestergaard, Allan Pillay, Sascha Knauf, Cheng-Yen Chen, Quique Bassat, Diana L Martin, David Fegan, Fasihah Taleo, Jacob Kool, Sheila Lukehart, Paul M Emerson, Anthony W Solomon, Tun Ye, Ronald C Ballard, David C W Mabey, Kingsley B Asiedu

Lancet Infect Dis 2015; 15: 1220–25

Published Online Septemper 9, 2015 http://dx.doi.org/10.1016/S1473-3099(15)00136-X

Clinical Research Department, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK (M Marks MBBS, A W Solomon PhD, Prof D C W Mabey PhD); **Hospital for Tropical Diseases, University College London Hospitals NHS Trust, London, UK** (M Marks, A W Solomon, Prof D C W Mabey); **ISGlobal, Barcelona Centre for International Health Research, Hospital Clínic—Universitat de Barcelona, Barcelona, Spain** (O Mitjà PhD, Q Bassat PhD); **Lihir Medical Centre, International SOS, Newcrest Mining, Lihir Island, Papua New Guinea** (O Mitjà); **Division for Communicable Diseases, Regional Office for the Western Pacific, World Health Organization, Manila, Philippines** (L S Vestergaard PhD); **Molecular Diagnostics and Typing Laboratory, Laboratory Reference and Research Branch, Division of STD Prevention** (A Pillay PhD, C-Y Chen PhD), **Division of Parasitic Diseases and Malaria** (D L Martin PhD), **and Center for Global Health** (T Ye PhD, R C Ballard PhD), **Centers for Disease Control and Prevention, Atlanta, GA, USA; German Primate Center, Leibniz-Institute for Primate Research, Pathology Unit, Working Group Neglected Tropical Diseases, Göttingen, Germany** (S Knauf PhD); **Centro de Investigação em Saúde de Manhiça, Maputo, Mozambique** (Q Bassat); **World Health Organization Consultant, Springhill, Brisbane, Queensland, Australia** (D Fegan FRCP); **Ministry of Health, Port Vila, Vanuatu** (F Taleo BSc); **World Health Organization, Vanuatu Country Office, Port Vila, Vanuatu** (J Kool MD);

Yaws is endemic in west Africa, southeast Asia, and the Pacific region. To eradicate yaws by 2020, WHO has launched a campaign of mass treatment with azithromycin. Progress has been made towards achievement of this ambitious goal, including the validation of point-of-care and molecular diagnostic tests and piloting of the strategy in several countries, including Ghana, Vanuatu, and Papua New Guinea. Gaps in knowledge need to be addressed to allow refinement of the eradication strategy. Studies exploring determinants of the spatial distribution of yaws are needed to help with the completion of baseline mapping. The finding that *Haemophilus ducreyi* causes lesions similar to yaws is particularly important and further work is needed to assess the effect of azithromycin on these lesions. The integration of diagnostic tests into different stages of the eradication campaign needs investigation. Finally, studies must be done to inform the optimum mass-treatment strategy for sustainable interruption of transmission.

Introduction

Yaws, a disease caused by *Treponema pallidum* subspecies *pertenue*,[1] is one of the three endemic non-venereal treponemal diseases.[2] Yaws predominantly affects children living in poor, remote communities and results in lesions of the skin, bone, and cartilage. If untreated, the disease can progress to cause destructive lesions of bones and cartilage. Previously, yaws was widespread throughout the tropics,[3] but a series of control efforts based on mass treatment with intramuscular penicillin and case finding led by WHO and UNICEF in the 20th century is estimated to have reduced the burden of cases worldwide by up to 95%.[4] Despite these efforts, regional incidence has rebounded in west and central Africa, the Pacific region, and southeast Asia, and yaws remains common in some of the poorest countries of the world.

Momentum for a new campaign against yaws was stimulated by the 2012 publication of a study[5] that showed single-dose azithromycin was clinically highly effective and non-inferior to penicillin. The availability of a single-dose that is oral, well tolerated, and a proven treatment prompted WHO to develop a new strategy on the basis of total-community mass treatment with azithromycin, identification of cases, and targeted treatment to eradicate the disease worldwide by 2020 (the Morges strategy).[6]

In the past 3 years, new diagnostic tools to potentially help to support the WHO eradication effort have been developed. At the same time, key areas of research have been identified to ensure the international community can overcome some of the challenges related to the delivery of this goal. We review major developments and present the priority research questions in epidemiology, diagnostics, and treatment that must be addressed for the successful completion of this eradication campaign.

Epidemiology and mapping

A major obstacle to interruption of yaws transmission is the paucity of information about where cases still occur. Before the mid-20th century eradication programmes, 99 countries reported yaws endemics (panel 1).[3] Of these countries, two—India and Ecuador—have reported successful elimination of yaws after large, mass-treatment programmes sponsored by their respective governments.[7,8] 13 countries are now known to be endemic for yaws, with the major documented foci being in the Pacific region, west Africa, and southeast Asia. For the other 84 countries that had a yaws endemic, little epidemiological information about yaws is available. To improve this situation, a substantial increase in the scale and speed of yaws mapping is needed and these improved epidemiological data are crucial to inform decisions about the interventions and resources necessary to successfully undertake eradication efforts.

To develop a successful mapping strategy, several questions need to be answered. A key question is: what are the appropriate evaluation and treatment units for yaws? The WHO eradication strategy[6] defines the treatment unit as the endemic village or community. Experience with other disease programmes has shown that village-level surveys are very labour intensive. Cases of yaws are known to cluster at both the household and the village level.[9,10] Climatic factors such as rainfall[11] have been proposed to explain this spatial heterogeneity but few data are available to explain the clustered nature of the disease, of which an increased understanding would help to refine strategies for defining evaluation units. Detailed fine-scale mapping at a very high spatial resolution of people with yaws might be necessary in some areas to allow this question to be explored in more detail. As the treatment programme scales up, decision making about treatment will probably be impractical at the village level and treatment, in the first round, at a larger level might be more appropriate (panel 2).

Initial mapping efforts should be focused on the 13 countries known to be endemic, to inform local efforts about how to interrupt transmission (table). Mapping should also be done in the countries that were formerly classified as endemic, but for which no present data are available. Creation of centralised systems to store and display data could help to standardise approaches for their collection and analysis, as has happened for other diseases such as trachoma,

Panel 1: **Key research questions for the epidemiology, diagnosis, and treatment of yaws**

Epidemiology

1 What is the appropriate evaluation unit for a survey before mass treatment?
2 What is the at-risk population and distribution of cases in the 13 known endemic countries?
3 What is the status of the 84 formerly endemic countries?
4 Which factors affect the spatial heterogeneity of yaws?

Diagnostics

1 Can the Dual Path Platform point-of-care assay replace traditional serological tests for all phases of the yaws eradication programme?
2 Can an instrument be developed to allow diagnosis of active yaws from a dried blood spot—facilitating integration of mapping activities with other programmes?
3 What is the appropriate role of PCR in the implementation and post-zero-cases surveillance phases of the yaws eradication programme?

Treatment

1 How many rounds of mass treatment should be done to interrupt transmission?
2 Is a recommended dose of azithromycin of 20 mg/kg non-inferior to a recommended dose of 30 mg/kg?
3 Does resistance to azithromycin emerge in *Treponema pallidum* subspecies *pertenue* after mass treatment?
4 What is the effect of community mass treatment with azithromycin on other ulcers, including those caused by *Haemophilus ducreyi*?
5 What are the best approaches to mobilise and sustain communities' support for yaws eradication?

Panel 2: **Key definitions for endemic status, implementation unit, and case definitions**

Endemic status

- Endemic village: a village containing at least one indigenous confirmed case
- Endemic country: a country with at least one indigenous confirmed case
- Formerly endemic country: a country that formerly reported yaws but that has either eliminated the disease or no data are available

Implementation unit

- Initial total community treatment: the implementation unit will be flexible covering a population of 100 000–250 000 people living in a region in which endemic villages are known
- Subsequent treatment rounds: the implementation unit at the level of the individual endemic village

Case definitions

- Suspected case in the implementation phase: individual with clinical signs consistent with yaws
- Confirmed case in the implementation phase: a suspected case with dual positive serology (either Dual Path Platform dually-positive or *Treponema pallidum* particle agglutination assay + rapid plasma reagin positive); PCR might be used during the implementation phase to monitor for resistance but is not an essential part of the case definition
- Confirmed case in the post-zero-cases surveillance phase: suspected case with both dual positive serological tests and positive PCR of lesion material for *T pallidum* subspecies *pertenue*

as well as the rapid implementation of public health interventions in regions or countries in which yaws is established.

Clinical diagnosis of yaws alone is unlikely to be reliable enough to inform eradication programmes[14,15] and therefore development of a strategy integrating robust diagnostics such as point-of-care serological tests and new molecular tests[16–20] into national health systems is needed. The case definition used and requirements for mapping might differ between settings before community mass treatment (in which a less specific method might be acceptable), and after mass treatment (in which high sensitivity and specificity will be needed to confirm eradication; panel 2).

Finally, efforts should be undertaken to learn from and integrate with other large-scale mapping projects such as the Global Trachoma Mapping Project[21] and the Atlas of Human African Trypanosomiasis.[22] These efforts led to large numbers of field staff being trained to internationally recognised standards and development of standard platforms for data capture, analysis, and sharing. Development of similar methods to map yaws should be a priority. Successful integration of yaws and trachoma mapping has already been accomplished in the Solomon Islands and Vanuatu[9] and similar efforts should be considered elsewhere.

Diagnostics

The renewal of the yaws eradication programme has stimulated advances in diagnostic methods for treponemal infections (panel 1). The selection of the most appropriate diagnostic algorithm at each stage of the eradication effort will be crucial for success of the programme. Historically, clinical and serological diagnostics have been used at all stages of programmes to certify elimination of the disease, although a refined approach will probably be needed in the post-mass drug administration environment to ensure that transmission has been interrupted.

Several advances in yaws diagnostics might contribute to the mapping and confirmation of the elimination of this disease. A major focus has been on the validation of a new point-of-care assay—the Dual Path Platform

Departments of Medicine and Global Health, University of Washington, Seattle, WA, USA (Prof S Lukehart PhD); **International Trachoma Initiative, The Task Force for Global Health, Decatur, GA, USA** (P M Emerson PhD); **and Department of Control of Neglected Tropical Diseases, World Health Organization, Geneva, Switzerland** (A W Solomon, K B Asiedu MPH)

Correspondence to:
Dr Michael Marks, Clinical Research Department, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK
michael.marks@lshtm.ac.uk

	2008	2009	2010	2011	2012	2013
Benin	..	..	45	..	11	..
Côte d'Ivoire	..	..	3740	3343	3092	2256
Cameroon	..	..	802	133	59	97
Central African Republic	243	..	..	..	230	..
Congo (Brazzaville)	..	646	..	..	197	..
Democratic Republic of the Congo	383	..	..	..	..	..
Ghana	20525	35248	18157	9674	8980	18702
Indonesia	6083	7751	6178	6631	4360	2043
Papua New Guinea	28989	25822	29061	28989	17560	19710
Solomon Islands	..	..	20635	..	12372	14909
Timor-Leste*	..	..	..	..	..	..
Togo	..	..	15	..	..	..
Vanuatu	1972	2432	1593	2331	2514	1198

Data taken from the WHO Global Observatory Data Repository.[12] *No recent routine reporting data exist from Timor-Leste, but reports from both the Ministry of Health and WHO Country office suggest that the disease is still known to be endemic in the country.[13]

Table: **Number of cases of yaws reported by country**

(DPP) Syphilis Screen-and-Confirm assay (Chembio Diagnostic Systems, Medford, NY, USA), which is based on simultaneous detection of antibodies to treponemal and non-treponemal antigens. This point-of-care test has now been shown to be accurate for community screening in yaws-endemic communities and for confirmation of clinically suspected cases.[18,19] Replacement of central laboratory rapid plasma reagin testing with the DPP assay might improve the implementation of surveillance in setting before and after mass treatment. Although DPP has shown high sensitivity in strongly seropositive individuals, the sensitivity is reduced in individuals with low-titre serological results.[19] Strategies such as repeat or confirmatory testing might therefore be necessary in some settings.

A large-scale surveillance method would be useful to identify communities needing focal mapping. To this end, studies are underway to validate serological testing for yaws on a multiplexed platform that could be integrated with other elimination or national health programmes.[23] Ideally, this strategy would help to differentiate previous infections from current ones, the latter of which, applied in appropriate age groups, might suggest continuing transmission.[24]

Molecular instruments specific to subspecies that accurately identify the presence or absence of *T pallidum* subspecies *pertenue*, and the mutations associated with azithromycin resistance, have also become available.[16,18,20,25–28] Data showing that lesion exudate samples collected with filter paper in the field generate reliable results, will help to overcome a major logistical obstacle previously precluding programmatic application of PCR, allowing these techniques to be used even for samples from the most remote communities in which yaws is identified.[20] Additionally, the development of other DNA-based amplification techniques, such as a loop-mediated isothermal amplification test, as a simple and rapid screening method in the field or with point-of-care tests could be useful in low-resource settings in which a sophisticated molecular testing might be impractical. The application of molecular diagnostics, genotyping, and microbiome analysis with whole genome sequencing might also allow a detailed understanding of the biological mechanisms of pathogens in cutaneous lesions and the epidemiology of yaws that will enable careful assessment of the efficacy of eradication efforts.

Instruments to differentiate patients who had unsuccessful treatment from individuals who remain seropositive after successful treatment (serofast status) need to be developed. This development is particularly important in view of the increasing recognition that other organisms might cause chronic skin ulcers in communities endemic for yaws.[14,26] So far, no available diagnostic test is shown to be of value for this role. Detection of treponemal DNA sequences in blood specimens by PCR has been achieved sporadically in patients with early syphilis,[29] but not in patients with yaws.[30]

Individuals who had been treated for clinical disease yet remain serofast will probably need pragmatic retreatment, possibly with an alternative drug such as injectable penicillin. If treatment of these individuals has been successful then such cases will cease to contribute to transmission of the disease and the number of clinical cases should decrease. Because yaws can relapse for up to 5–10 years, disease surveillance in these individuals will need to be maintained for a prolonged period to confirm that transmission has truly been interrupted. As the number of clinical cases falls, the emphasis of surveillance will switch to serological surveillance of children aged 5 years or younger to detect the presence or absence of newly acquired yaws infections. Use of this measure should avoid the complexity of monitoring serofast individuals.

Laboratory capacity building is needed for distribution and deployment of new diagnostic instruments worldwide. Technology transfer efforts and the identification of appropriate regional and national reference laboratories will be very important to support national control programmes for yaws. Roll out of new diagnostic techniques to local laboratories will need the development and implementation of appropriate quality control and assessment structures.

Non-treponemal causes of yaws-like lesions

The differential diagnosis of ulcerative lesions in tropical countries is broad and includes treponemal infections, pyoderma, and polymicrobial tropical ulcers. *Haemophilus ducreyi* in particular has now been shown to be an important cause of ulcerative skin lesions clinically similar to yaws, in Papua New Guinea, the

Solomon Islands, Vanuatu, and Ghana.[14,26] Although some clinical phenotypes—particularly bony disease—are likely to be more yaws-specific, the clinical similarity of non-yaws lesions compared with common skin features of yaws complicates clinical case reporting by national surveillance programmes[15] and shows the need for integration of point-of-care assays into surveillance strategies to confirm suspected cases of yaws.

An additional complication of yaws-like ulcers that should be considered is that these might affect community perceptions of the efficacy of mass treatment of azithromycin. Communities might expect mass treatment to be a solution for all skin diseases as opposed to a specific programme targeting a particular pathogen. The best way to educate communities about the intended benefits of yaws eradication programmes should be an area of active research, and will need input from social scientists and medical anthropologists. Integration of case management for all skin ulcers, irrespective of their cause, into mass treatment programmes should be considered.

Experimental models[31] suggest that azithromycin will also be effective in treating non-genital lesions caused by *H ducreyi*, but further studies will be needed to assess this in the field. Data from pilot studies[32,33] of community mass treatment for yaws eradication showed that the absolute number of lesions caused by *H ducreyi* was also reduced, although not as substantially as those caused by *T pallidum* subspecies *pertenue*. Consideration should also be given to a syndromic approach to the management of skin ulcers with an integration of basic skin care interventions such as provision of soap and ulcer dressings, alongside mass treatment with azithromycin, a strategy likely to have collateral benefits for community engagement and general skin health.

Treatment

Until 2012, long-acting injectable penicillin (typically benzathine benzylpenicillin) had been the mainstay of treatment for yaws.[11] A landmark study[5] in Papua New Guinea showed that single-dose oral azithromycin at a recommended dose of 30 mg/kg (maximum 2 g) was equivalent to treatment with penicillin in both primary and secondary yaws with 106 (95%) of 110 patients clinically and serologically cured at 6 months. Treatment with azithromycin is now central to the WHO eradication strategy and should be more readily accepted by children and their guardians than parenteral penicillin.[6]

The present WHO eradication strategy is initial total community treatment with single-dose oral azithromycin and subsequent resurveys with targeted treatment of patients with residual yaws and their contacts. Selection of the appropriate treatment unit will be important to make this strategy both effective and efficient (panel 1). In earlier campaigns, failure to treat contacts and latent cases were thought to contribute to rapid return of the disease.[34] This risk should theoretically be reduced by the new mass-treatment recommendation.

Pilot studies[33] have shown that implementation of the WHO eradication strategy results in a reduction of clinical and latent cases of yaws and might potentially interrupt transmission.[32] For syphilis, mass treatment with azithromycin was temporarily effective in reducing the frequency of disease in high-risk groups;[35] however, in these studies syphilis subsequently rebounded to premass treatment levels, emphasising the importance of follow-up treatment to clear infection from cases missed from detection and interrupt transmission. The number of rounds of mass treatment and the population coverage needed to interrupt transmission for yaws is unknown. In the successful elimination campaign in India, community surveillance and targeted treatment were done every 6 months for a period of 7 years.[36] Randomised trials are unlikely to be possible that answer the questions about the proportion of population coverage and numbers of rounds of treatment needed to eradicate yaws. Instead, observational data from pilot mass drug administration programmes and mathematical models should be reviewed to identify optimum community treatment strategies.

The recommended dose of azithromycin for patients with yaws is 30 mg/kg (maximum 2 g).[6] A major disadvantage of this regimen is the relatively high incidence of associated gastrointestinal side-effects of about 15%. Azithromycin is also used for several other indications, notably trachoma, at a lower recommended dose of 20 mg/kg bodyweight (maximum 1 g). Lower doses of azithromycin are likely to be slightly better tolerated than higher-dose treatment.[37] Use of a lower dose could confer substantial cost savings for health-care programmes, which is particularly important in view of the absence of a drug donation programme for yaws eradication, as well as offering synergies for countries in the Pacific region where yaws and trachoma are coendemic. Because three of the most heavily endemic countries are in the Pacific region, this low-dose treatment could be particularly beneficial. However, these potential benefits will need to be carefully balanced against the possibility of the development of drug resistance. A WHO-sponsored trial is planned for 2015 (NCT02344628), in Papua New Guinea and Ghana, to establish whether a 20 mg/kg dose of azithromycin is effective against yaws.

Emergence of macrolide resistance in *T pallidum* subspecies *pertenue*, which has already occurred with *T pallidum* subspecies *pallidum*, is a major concern.[27,38,39] Macrolide resistance in treponemes is mediated by point mutations in the 23S rRNA gene[25,27,28,40] and is associated with previous exposure to other macrolides.[41] Surveillance for known macrolide resistance mutations in a very small number of azithromycin-naive settings has not detected any azithromycin resistant yaws strains; however, close monitoring after mass treatment will be essential to ensure the success of the eradication strategy.

Other considerations

Human yaws is not known to have an animal reservoir. However, a substantial number of pathogenic *T pallidum* infections, asymptomatic infection, and clinical disease occur in our closest relatives—non-human primates in Africa.[42–46] So far, all the sequenced simian samples are closely related to human yaws-causing strains.[42,47] Experimental infection of human beings with the Fribourg-Blanc simian strain,[48] and reports on laboratory infections of non-human primates with human *T pallidum* isolates,[44] suggest that zoonotic transmission is theoretically possible.[48] The relevance of these findings to the yaws eradication campaign is unclear, but further study is warranted.

Although the focus of this Review has been on yaws, the two other endemic treponemal diseases, bejel and pinta, are also worth remembering.[11] Available data for these two diseases are even more restricted than for yaws, but the clinical and microbiological similarities among the diseases suggest that strategies developed for the control of yaws might be more broadly applicable. These considerations are particularly important when mapping strategies are formulated, so that opportunities to improve our understanding and control of these diseases are also grasped when feasible.

Although yaws eradication is likely to be a cost-effective intervention,[49] financial and political support must be secured at the national, regional, and international levels. Pilot projects[30,33] have shown the feasibility of the WHO strategy and should encourage donations of both funds and azithromycin. In addition to financial support, engagement of advocacy groups and endemic communities is needed to increase awareness of yaws and show the possibility of a successful eradication of the disease. Links to other control programmes for neglected tropical diseases should be actively sought to build synergies at all stages of the eradication campaign. A major burden of yaws is in west and central Africa, which has been severely affected by the Ebola virus disease epidemic. Although this outbreak did not predominantly affect countries in which yaws is a major problem, the regional disruption on health-care services and personnel caused by the epidemic might affect yaws eradication efforts. Finally, the eradication campaign must develop strategies to link into the wider public health systems of the affected countries.

Conclusions

Much progress has been made in the past 3 years in development strategies and methods for yaws eradication: a simple, cheap, and well tolerated oral treatment, azithromycin, is effective in the treatment of yaws; point of care diagnostic tests have been validated; and new molecular tests have become available. Studies on the epidemiology of yaws and pilot mass-treatment programmes have been rolled out in a few endemic countries. Additionally, interest and support has increased within the academic and global health communities for yaws eradication efforts, as shown in the increase in research output related to yaws. The WHO Morges strategy represents the cornerstone of efforts to eradicate yaws and international efforts should be focused on continuing to roll out this strategy in the known endemic countries, while simultaneously working to answer the research questions identified in this Review. Advocacy will be crucial in the upcoming years to transmit to the wider public health community the importance of addressing this neglected disease and to trigger permanent plans for action towards yaws eradication. The present yaws programme can benefit from the knowledge and experience of past yaws campaigns and of other international control initiatives for neglected tropical diseases. Important opportunities are available for synergy among disease control efforts to help to lead to a sustained improvement in the quality of life for some of the world's most neglected people.

Search strategy and selection criteria

We searched PubMed for studies published between January, 1900, and February, 2015, using the terms "yaws", "pian", and "*Treponema pallidum*". Reference lists of identified manuscripts were reviewed to identify additional relevant material. We reviewed literature and statistics held at the WHO and data presented at WHO consultative meetings on yaws eradication. No language restrictions were made for this Review.

Declaration of interests
We declare no competing interests.

Contributors
MM, LSV, OM, and KBA conceived of the project. MM wrote the first draft of the manuscript and made revisions. LSV, OM, SK, AP, C-YC, DLM, TY, QB, JK, FT, DF, SL, PME, AWS, DCWM, RCB, and KBA reviewed drafts of the manuscript, provided comments, critical review, and helped to revise the manuscript.

Acknowledgments
MM is supported by a Wellcome Trust Clinical Research Fellowship (WT102807). QB has a fellowship from the program Miguel Servet of the ISCIII (Plan Nacional de I+D+I 2008–2011, grant number: CP11/00269). The funder had no role in the preparation of the manuscript or the decision to submit it for publication. We thank the participants of the WHO yaws eradication and COR-NTD meetings held in 2014 for their valuable contributions to the topics discussed in this manuscript. KBA, AWS, JK, and LSV are employees of the WHO. AP, C-YC, DLM are employees of the Centers for Disease Control and Prevention (CDC) . The views expressed in this article are the views of the authors and may not necessarily reflect the views of the WHO or the CDC.

References

1 Mitjà O, Asiedu K, Mabey D. Yaws. *Lancet* 2013; **381:** 763–73.
2 Marks M, Solomon AW, Mabey DC. Endemic treponemal diseases. *Trans R Soc Trop Med Hyg* 2014; **108:** 601–07.
3 Kazadi WM, Asiedu KB, Agana N, Mitjà O. Epidemiology of yaws: an update. *Clin Epidemiol* 2014; **6:** 119–28.
4 Asiedu K, Fitzpatrick C, Jannin J. Eradication of yaws: historical efforts and achieving WHO's 2020 target. *PLoS Negl Trop Dis* 2014; **8:** e3016.

5 Mitjà O, Hays R, Ipai A, et al. Single-dose azithromycin versus benzathine benzylpenicillin for treatment of yaws in children in Papua New Guinea: an open-label, non-inferiority, randomised trial. *Lancet* 2012; **379:** 342–47.

6 The World Health Organization. Eradication of yaws—the Morges strategy. *Wkly Epidemiol Rec* 2012; **87:** 189–94.

7 Anselmi M, Moreira J-M, Caicedo C, Guderian R, Tognoni G. Community participation eliminates yaws in Ecuador. *Trop Med Int Health* 2003; **8:** 634–38.

8 Bora D, Dhariwal AC, Lal S. Yaws and its eradication in India—a brief review. *J Commun Dis* 2005; **37:** 1–11.

9 Marks M, Vahi V, Sokana O, et al. Mapping the epidemiology of yaws in the Solomon Islands: a cluster randomized survey. *Am J Trop Med Hyg* 2015; **92:** 129–33.

10 Mitjà O, Hays R, Ipai A, et al. Outcome predictors in treatment of yaws. *Emerg Infect Dis* 2011; **17:** 1083–85.

11 Perine PL, Hopkins DR, Niemel PLA, St John RK, Causse G, Antal GM. Handbook of endemic treponematoses : yaws, endemic syphilis and pinta. Geneva: World Health Organization, 1984.

12 WHO. Country Cooperation Strategy 2009–2013. Timor-Leste. Geneva: World Health Organization, 2009.

13 WHO. Global Health Observatory Data Repository. Status of endemicity for yaws. http://apps.who.int/gho/data/node.main.NTDYAWSEND?lang=en (accessed Oct 2, 2014).

14 Marks M, Chi K-H, Vahi V, et al. *Haemophilus ducreyi* associated with skin ulcers among children, Solomon Islands. *Emerg Infect Dis* 2014; **20:** 1705–07.

15 Ghinai R, El-Duah P, Chi K-H, et al. A cross-sectional study of 'yaws' in districts of Ghana which have previously undertaken azithromycin mass drug administration for trachoma control. *PLoS Negl Trop Dis* 2015; **9:** e0003496.

16 Centurion-Lara A, Molini BJ, Godornes C, et al. Molecular differentiation of *Treponema pallidum* subspecies. *J Clin Microbiol* 2006; **44:** 3377–80.

17 Centurion-Lara A, Giacani L, Godornes C, Molini BJ, Brinck Reid T, Lukehart SA. Fine analysis of genetic diversity of the tpr gene family among treponemal species, subspecies and strains. *PLoS Negl Trop Dis* 2013; **7:** e2222.

18 Ayove T, Houniei W, Wangnapi R, et al. Sensitivity and specificity of a rapid point-of-care test for active yaws: a comparative study. *Lancet Glob Health* 2014; **2:** e415–21.

19 Marks M, Goncalves A, Vahi V, et al. Evaluation of a rapid diagnostic test for yaws infection in a community surveillance setting. *PLoS Negl Trop Dis* 2014; **8:** e3156.

20 Chi K-H, Danavall D, Taleo F, et al. Molecular differentiation of *Treponema pallidum* subspecies in skin ulceration clinically suspected as yaws in Vanuatu using real-time multiplex PCR and serological methods. *Am J Trop Med Hyg* 2015; **92:** 134–38.

21 Kabore A, Solomon AW. Completing the baseline map of the preventive-chemotherapy Neglected Tropical Diseases (NTDs): the global trachoma mapping project and WHO's NTD mapping initiative. American Society of Tropical Medicine and Hygiene—62nd Annual Meeting; Washington, DC; November 13–17, 2013; 124.

22 Simarro PP, Cecchi G, Paone M, et al. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *Int J Health Geogr* 2010; **9:** 57.

23 Lammie PJ, Moss DM, Brook Goodhew E, et al. Development of a new platform for neglected tropical disease surveillance. *Int J Parasitol* 2012; **42:** 797–800.

24 Ijsselmuiden OE, Schouls LM, Stolz E, et al. Sensitivity and specificity of an enzyme-linked immunosorbent assay using the recombinant DNA-derived *Treponema pallidum* protein TmpA for serodiagnosis of syphilis and the potential use of TmpA for assessing the effect of antibiotic therapy. *J Clin Microbiol* 1989; **27:** 152–57.

25 Chen C-Y, Chi K-H, Pillay A, Nachamkin E, Su JR, Ballard RC. Detection of the A2058G and A2059G 23S rRNA gene point mutations associated with azithromycin resistance in *Treponema pallidum* by use of a TaqMan real-time multiplex PCR assay. *J Clin Microbiol* 2013; **51:** 908–13.

26 Mitjà O, Lukehart SA, Pokowas G, et al. *Haemophilus ducreyi* as a cause of skin ulcers in children from a yaws-endemic area of Papua New Guinea: a prospective cohort study. *Lancet Glob Health* 2014; **2:** e235–41.

27 Lukehart SA, Godornes C, Molini BJ, et al. Macrolide resistance in *Treponema pallidum* in the United States and Ireland. *N Engl J Med* 2004; **351:** 154–58.

28 Matejková P, Flasarová M, Zákoucká H, et al. Macrolide treatment failure in a case of secondary syphilis: a novel A2059G mutation in the 23S rRNA gene of *Treponema pallidum* subsp. *pallidum*. *J Med Microbiol* 2009; **58:** 832–36.

29 Marfin AA, Liu H, Sutton MY, Steiner B, Pillay A, Markowitz LE. Amplification of the DNA polymerase I gene of *Treponema pallidum* from whole blood of persons with syphilis. *Diagn Microbiol Infect Dis* 2001; **40:** 163–66.

30 Marks M, Katz S, Vahi V, et al. Failure of PCR to detect *Treponema pallidum* ssp *pertenue* DNA in blood in latent yaws. *PLoS Negl Trop Dis* 2015; **9:** e0003905.

31 Thornton AC, O'Mara EM Jr, Sorensen SJ, et al. Prevention of experimental *Haemophilus ducreyi* infection: a randomized, controlled clinical trial. *J Infect Dis* 1998; **177:** 1608–13.

32 Mitjà O, Houinei W, Moses P, et al. Mass treatment with single-dose azithromycin for yaws. *N Engl J Med* 2015; **372:** 703–10.

33 Marks M, Vahi V, Sokana O, et al. Impact of community mass treatment with azithromycin for trachoma elimination on the prevalence of yaws. *PLoS Negl Trop Dis* 2015; **9:** e0003988.

34 Hackett CJ, Guthe T. Some important aspects of yaws eradication. *Bull World Health Organ* 1956; **15:** 869–96.

35 Rekart ML, Patrick DM, Chakraborty B, et al. Targeted mass treatment for syphilis with oral azithromycin. *Lancet* 2003; **361:** 313–14.

36 World Health Organization. Elimination of yaws in India. *Wkly Epidemiol Rec* 2008; **83:** 125–32.

37 Campbell JP, Mkocha H, Munoz B, West SK. Randomized trial of high dose azithromycin compared to standard dosing for children with severe trachoma in Tanzania. *Ophthal Epidemiol* 2009; **16:** 175–80.

38 Chen X-S, Yin Y-P, Wei W-H, et al. High prevalence of azithromycin resistance to *Treponema pallidum* in geographically different areas in China. *Clin Microbiol Infect* 2013; **19:** 975–79.

39 Martin IE, Tsang RSW, Sutherland K, et al. Molecular characterization of syphilis in patients in Canada: azithromycin resistance and detection of *Treponema pallidum* DNA in whole-blood samples versus ulcerative swabs. *J Clin Microbiol* 2009; **47:** 1668–73.

40 Stamm LV, Bergen HL. A point mutation associated with bacterial macrolide resistance is present in both 23S rRNA genes of an erythromycin-resistant *Treponema pallidum* clinical isolate. *Antimicrob Agents Chemother* 2000; **44:** 806–07.

41 Marra CM, Colina AP, Godornes C, et al. Antibiotic selection may contribute to increases in macrolide-resistant *Treponema pallidum*. *J Infect Dis* 2006; **194:** 1771–73.

42 Harper KN, Fyumagwa RD, Hoare R, et al. *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. *PLoS One* 2012; **7:** e50882.

43 Knauf S, Batamuzi EK, Mlengeya T, et al. Treponema infection associated with genital ulceration in wild baboons. *Vet Pathol* 2012; **49:** 292–303.

44 Harper KN, Knauf S. *Treponema pallidum* infection in primates: clinical manifestations, epidemiology, and evolution of a stealthy pathogen. In: Brinkworth J, Pechenkina K, eds. Primates, pathogens, and evolution. Series: developments in primatology: progress and prospects. New York: Springer, 2013: 189–219.

45 Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. *Primates Med* 1969; **3:** 113–21.

46 Fribourg-Blanc A, Mollaret HH, Niel G. Serologic and microscopic confirmation of treponemosis in Guinea baboons. *Bull Société Pathol Exot Ses Fil* 1966; **59:** 54–59 (in French).

47 Zobaníková M, Strouhal M, Mikalová L, et al. Whole genome sequence of the Treponema Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. *PLoS Negl Trop Dis* 2013; **7:** e2172.

48 Smith JL, David NJ, Indgin S, et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. *Br J Vener Dis* 1971; **47:** 226–51.

49 Fitzpatrick C, Asiedu K, Jannin J. Where the road ends, yaws begins? The cost-effectiveness of eradication versus more roads. *PLoS Negl Trop Dis* 2014; **8:** e3165.

7.2.2 Treponemal Infection in Nonhuman Primates as Possible Reservoir for Human Yaws

The publication summarizes current knowledge and stresses the need to investigate NHPs as a potential reservoir for human yaws infection. The consolidation of reported *T. pallidum* cases in African NHPs shows a clear geographic overlap with areas that have been reported to be endemic for human yaws (spatial connectivity, chapter 4).

burnetii infection and related reproductive performance in high producing dairy cows. Res Vet Sci. 2012;93:67–73. http://dx.doi.org/10.1016/j.rvsc.2011.07.017
5. Berri M, Souriau A, Crosby M, Rodolakis A. Shedding of *Coxiella burnetii* in ewes in two pregnancies following an episode of *Coxiella* abortion in a sheep flock. Vet Microbiol. 2002;85:55–60. http://dx.doi.org/10.1016/S0378-1135(01)00480-1
6. Böttcher J, Vossen A, Janowetz B, Alex M, Gangl A, Randt A, et al. Insights into the dynamics of endemic *Coxiella burnetii* infection in cattle by application of phase-specific ELISAs in an infected dairy herd. Vet Microbiol. 2011;151:291–300. http://dx.doi.org/10.1016/j.vetmic.2011.03.007
7. Centers for Disease Control and Prevention. Notes from the field: Q fever outbreak associated with goat farms—Washington and Montana, 2011. MMWR Morb Mortal Wkly Rep. 2011;60:1393.
8. Runge M, Binder A, Schotte U, Ganter M. Investigations concerning the prevalence of *Coxiella burnetii* and *Chlamydia abortus* in sheep in correlation with management systems and abortion rate in Lower Saxony in 2004. Berl Munch Tierarztl Wochenschr. 2012;125:138–43.
9. Hansen MS, Rodolakis A, Cochonneau D, Agger JF, Christoffersen AB, Jensen TK, et al. *Coxiella burnetii* associated placental lesions and infection level in parturient cows. Vet J. 2011;190:e135–9. Epub 2011 Feb 2. http://dx.doi.org/10.1016/j.tvjl.2010.12.021
10. Christensen DR, Hartman LJ, Loveless BM, Frye MS, Shipley MA, Bridge DL, et al. Detection of biological threat agents by real-time PCR: comparison of assay performance on the R.A.P.I.D., the LightCycler, and the Smart Cycler platforms. Clin Chem. 2006;52:141–5. http://dx.doi.org/10.1373/clinchem.2005.052522

Address for correspondence: Samuel L. Yingst, United States Army Materiel Command, Armed Forces Research Institute of Medical Sciences, APO AP 96546, Bangkok, Thailand; email: samuel.yingst@afrims.org

Treponemal Infection in Nonhuman Primates as Possible Reservoir for Human Yaws

To the Editor: In 2012, the World Health Organization launched plans for a second campaign to eradicate the neglected tropical disease, yaws (*1*). The first campaign, conducted during the mid-20th century, was tremendously successful in terms of treatment and reduced the number of cases by 95%. However, it failed to eradicate the disease, and when local efforts to prevent new cases proved insufficient, yaws resurged in some areas. Comments on the new yaws eradication campaign have emphasized the need for sustained support and resources. Here we draw attention to an additional concern that could impede yaws eradication efforts.

The success of any eradication campaign depends on the absence of a nonhuman reservoir. Smallpox had no known animal reservoir, and polio and dracunculiasis (guinea worm disease), which are currently the focus of the World Health Organization eradication campaigns, also have none. By contrast, compelling evidence suggests that yaws exists in wild nonhuman primate populations residing in regions where humans are also infected (Figure).

The subspecies of the bacterium *Treponema pallidum* that cause the non–sexually transmitted diseases yaws (subsp. *pertenue* infection) and endemic syphilis (subsp. *endemicum* infection) and the sexually transmitted infection syphilis (subsp. *pallidum*) are close relatives. The 3 diseases cannot be distinguished serologically. Instead, the diseases they cause are usually differentiated by clinical characteristics and geographic distribution. Whereas syphilis is a venereal disease with a worldwide distribution, yaws primarily affects children in hot and humid areas of Africa and Asia, and endemic syphilis occurs in arid regions. Because methods available to differentiate between the *T. pallidum* subspecies were unavailable in the past, prevalence data for yaws were sometimes vague and inaccurate. Recently, molecular tests capable of distinguishing between the subspecies by using single nucleotide polymorphisms have been developed (*2*,*3*). These tests have enabled us to learn more about the *T. pallidum* strains that infect wild nonhuman primates.

During the 1960s, researchers reported that many baboons in West Africa were seropositive for treponemal infection (*4*). Since then, high levels of infection have been documented in other monkey species in West Africa and in great apes (*5*). Recently, we documented *T. pallidum* infection in olive baboons (*Papio anubis*) at Lake Manyara National Park in Tanzania (*6*). In West Africa, clinical signs of infection in nonhuman primates are usually mild, if present at all, consisting of small lesions around the muzzle, eyelids, and armpits (*4*). A recent survey in 2013 at Parc National du Niokolo-Koba, Senegal, revealed *T. pallidum* antibodies in Guinea baboons (*P. papio*) with no signs of infection (S. Knauf et al, unpub. data). By contrast, severe manifestations resembling tertiary-stage yaws have been reported in wild gorillas (*5*). In terms of genetic distance, studies thus far indicate that the organisms infecting baboons in West and East Africa closely resemble *T. pallidum* subsp. *pertenue,* the agent responsible for yaws in humans (*2*,*7*). In fact, the genome sequence of a *T. pallidum* strain collected from a baboon in Guinea indicates that it should be considered a *T. pallidum* subsp. *pertenue* strain (*8*). Infection has been confirmed by serologic tests in a variety of nonhuman primate species in the yaws belt of Africa and by PCR in baboons from East and West Africa (Figure).

The high prevalence of nonhuman primate infection in areas of tropical Africa where yaws is common in humans (Figure) suggests that cross-species infection may occur. Decades ago, researchers reported that the Fribourg-Blanc simian strain, collected in Guinea, can cause sustained infection in humans after inoculation (*9*). Such experiments are ethically questionable and the details given are scant, but this work suggests that simian strains have zoonotic potential. Additional research is needed to determine whether interspecies transmission of *T. pallidum* occurs under natural conditions. Bush meat preparation is common in many African countries and a major source of zoonotic infection. It involves frequent skin-to-skin contact, which is the preferred mode of transmission for yaws. Insects also have been proposed to be vectors of infection, although this has not been documented(*10*). If evidence of interspecies yaws transmission, either direct or by vector, is discovered, then nonhuman primates may be a major reservoir of infection for humans.

Additional studies comparing human and simian strains may show whether zoonotic transmission of *T. pallidum* occurs frequently, an important consideration with regard to disease eradication and the conservation of great apes and other endangered nonhuman primates. To eradicate yaws, all host species and any possible reservoirs need to be taken into account. We, like the rest of the world, want the second yaws eradication campaign to succeed and hope that nonhuman primate infection will be evaluated as a factor in disease transmission.

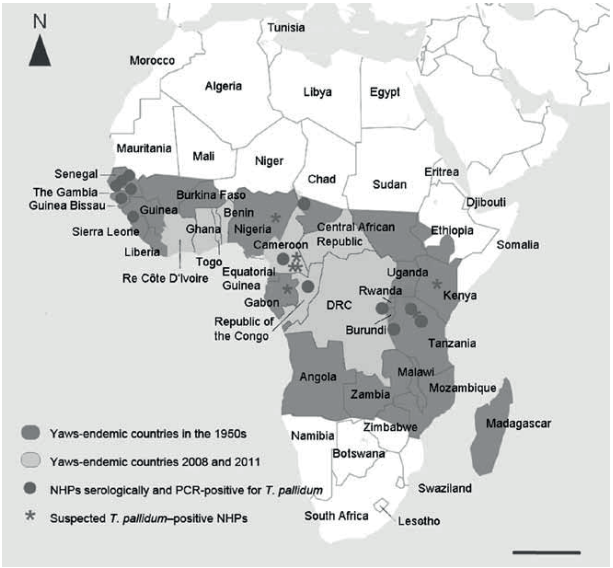


Figure. Geographic proximity between human yaws and endemic syphilis, as estimated by the World Health Organization, and locations in which treponemal infection has been identified in nonhuman primates (NHPs), Africa, 1990s. Red dots indicate infection in NHPs confirmed by sensitive and specific treponemal serologic tests (TPI/FTA-ABS/MHA-TP [Treponema-pallidum-immobilization reaction/fluorescence-Treponema-antibody-absorption test/Treponema pallidum microhemagglutination assay]) and, in some cases, PCR. Stars indicate suspected infection (i.e., sightings of NHPs with lesions consistent with infection). Sources include the following: 1) Cameroon: *Gorilla gorilla*, observation (W. Karesh, pers. comm.); *Pan troglodytes*, *G. gorilla*, and *Papio* sp., skeletal analysis and serology (*4*;*11* in online Technical Appendix: wwwnc.cdc.gov/EID/articlepdfs/19/12/13-0863-Techapp1.pdf). 2) Chad: *Erythrocebus patas*, serology (*4*). 3) Democratic Republic of Congo (DRC): *Pan troglodytes*, serology (*4*). 4) Gabon: *G. gorilla*, observation (W. Karesh, pers. comm.). 5) Guinea: *Papio* sp., serology and PCR (*4*,*8*). 6) Kenya: *Papio anubis* and *Chlorocebus* sp., observation and serology (J. Fischer, pers. comm.); *12* in online Technical Appendix). 7) Nigeria, *Papio anubis* (J. Wallis, pers. comm.). 8) Republic of Congo: *G. gorilla*, serology and observation (W. Karesh, unpub. data; *5*). 9) Tanzania: *P. anubis*; observation, serology, PCR (*6*,*7*; *13* in online Technical Appendix; S. Knauf, unpub. data). 10) Senegal: *Papio* sp., *Chlorocebus* sp., colobus monkeys, and *Erythrocebus patas*; serology (S. Knauf, unpub. data; *4*; *14* in online Technical Appendix). Scale bar = 1,000 km.

Sascha Knauf, Hsi Liu, and Kristin N. Harper

Author affiliations: German Primate Center, Göttingen, Germany (S. Knauf); Centers for Diseases Control and Prevention, Atlanta, Georgia, USA (H. Liu); and Columbia University Medical Center, New York, New York, USA (K.N. Harper)

DOI: http://dx.doi.org/10.3201/eid1912.130863

References

1. Maurice J. WHO plans new yaws eradication campaign. Lancet. 2012;379:1377–8. http://dx.doi.org/10.1016/S0140-6736(12)60581-9
2. Harper KN, Ocampo PS, Steiner BM, George RW, Silverman MS, Bolotin S, et al. On the origin of the treponematoses: a phylogenetic approach. PLoS Negl Trop Dis. 2008;2:e148. http://dx.doi.org/10.1371/journal.pntd.0000148
3. Centurion-Lara A, Castro C, Castillo R, Shaffer J. Voorhis Wv, Lukehart S. The flanking region sequences of the 15-kDa lipoprotein gene differentiate pathogenic treponemes. J Infect Dis. 1998;177:1036–40. http://dx.doi.org/10.1086/515247
4. Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. Primates Med. 1969;3:113–21.

5. Levréro F, Gatti S, Gautier-Hion A, Ménard N. Yaws disease in a wild gorilla population and its impact on the reproductive status of males. Am J Phys Anthropol. 2007;132:568–75. http://dx.doi.org/10.1002/ajpa.20560
6. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IA, Nordhoff M, et al. *Treponema* infection associated with genital ulceration in wild baboons. Vet Pathol. 2012;49:292–303. http://dx.doi.org/10.1177/0300985811402839
7. Harper KN, Fyumagwa R, Hoare R, Wambura P, Coppenhaver D, Sapolsky R, et al. *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS ONE. 2012;7:e50882. http://dx.doi.org/10.1371/journal.pone.0050882
8. Zobaníková M, Strouhal M, Mikalová L, Cejková D, Ambrožová L, Pospíšilová P, et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl Trop Dis. 2013;7:e2172. http://dx.doi.org/10.1371/journal.pntd.0002172
9. Smith JL, David NJ, Indgin S, Israel CW, Levine BM, Justice J Jr, et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br J Vener Dis. 1971;47:226–51.
10. Thomson JG, Lamborn WA. Mechanical transmission of trypanosomiasis, leishmaniasis, and yaws through the agency of non-biting haematophagous flies. BMJ. 1934;2:506–9. http://dx.doi.org/10.1136/bmj.2.3845.506

Address for correspondence: Sascha Knauf, Pathology Unit, German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany; email: sknauf@dpz.eu

Porcine Hokovirus in Domestic Pigs, Cameroon

To the Editor: Since 2005, new parvoviruses forming a novel genus of the proposed name *Partetravirus*, within the subfamily *Parvovirinae*, have been described (*1*). Human parvovirus 4 (PARV4) with 3 different genotypes globally infects humans (*2*). A related porcine virus, hokovirus (HoV or porcine partetravirus), was found in wild boar and domestic pig populations in Germany, Romania, China, and the United States, with prevalences of 12%–47%, forming 1 common genotype (*3*–*6*). Prevalence figures from sub-Saharan Africa are not available. Furthermore, no information about possibly region-associated genotypes is available for porcine HoV, although it is for human PARV4 from the same genus. We therefore used samples (collected during February–March 2012) from a study investigating hepatitis E virus (HEV) in pigs from Cameroon (*7*) to analyze the occurrence of porcine HoV in pigs in Africa and to determine the respective genotype.

Viral DNA was extracted from liver samples by using the RTP DNA/RNA Virus Mini Kit II (STRATEC-Molecular, Berlin, Germany) according to the manufacturer's instructions. DNA samples were pooled, with each pool containing 3 different samples. A total of 94 pooled samples from 282 animals originating from 3 districts in Cameroon (Doula, Yaoundé, and Bamenda) were investigated by using quantitative real-time PCR (*3*,*7*). Samples from pools that tested positive were analyzed individually.

We detected HoV in 65 (69%) of the 94 pooled samples: 2 (15%) of 13 from Bamenda, 39 (70%) of 56 from Douala, and 24 (96%) of 25 from Yaoundé. We used an online tool to estimate the individual prevalence from pooled samples for fixed pool size and perfect test with exact 5% upper and lower CIs (http://epitools.ausvet.com.au/content.php?page = PooledPrevalence). A pool size of 3 with a total of 94 pooled samples and 65 positive samples resulted in an estimated general prevalence of 32.4% (95% CI 27%–39%). For Bamenda, the estimated prevalence was 5.4% (95% CI 1%–16%); for Douala, 32.8% (95% CI 25%–41%); and for Yaoundé, 65.8% (95% CI 44%–87%).

From 94 positive pools, a total of 184 samples were available for individual testing: 6 from Bamenda, 110 from Douala, and 68 from Yaoundé; 12 were missing. Using the results from the negative tested pools and the individual testing, we found an estimated general prevalence of 47% (128/270). The regional prevalence was 10% (4/39) for Bamenda, 41% (65/160) for Douala, and 83% (59/71) for Yaoundé.

These prevalences are higher than the estimates, but lie within the regional estimates within the range of the CI determined with the online tool. The discrepancy in the total prevalence might be due to the missing samples for the individual testing. Our results show that pooled sample testing can yield a good approximation of the actual prevalence, at least for settings in Africa. The varying prevalence and inhomogeneous regional distribution of porcine HoV correspond to previous findings from Europe, China, and the United States in wild boar and domestic pigs (*3*,*5*,*6*). Overall, no general defined pig-breeding program is in place in Cameroon. Douala and Yaoundé are the main markets for pig trade. Yaoundé, the main town for pig purchase and slaughter, gets live pigs from northwestern (Bamenda), western, and northern Cameroon, and Douala receives pigs from northwestern (Bamenda), western, and southwestern Cameroon. To fully understand the observed regional prevalences, the presence of HoV needs to be investigated in detail in the southwest, west, and north, where intensive farming systems are in place and pig farming is of economic importance.

Near full-length genome data were generated from 3 positive samples, and partial sequence information was retrieved for 8 additional samples (Figure) as described (*3*). The phylogenetic analysis showed a very close relation, with 98%–99% homology between the porcine HoV isolates from Cameroon, Europe, the

7.3 Pathogen Maintenance in the Potential Nonhuman Reservoir

To define a functional reservoir, two important requirements must be fulfilled: pathogen maintenance in the reservoir system (chapter 7.3) and the existence of feasible transmission routes (chapter 7.4) (Hallmaier-Wacker et al., 2017). The following publications are investigating the epidemiology and maintenance of *T. pallidum* in NHPs and humans.

7.3.1 High Prevalence of Antibodies Against the Bacterium *Treponema pallidum* in Senegalese Guinea Baboons (*Papio papio*)

PLOS ONE

RESEARCH ARTICLE

High Prevalence of Antibodies against the Bacterium *Treponema pallidum* in Senegalese Guinea Baboons (*Papio papio*)

Sascha Knauf[1]*, Ulrike Barnett[2], Peter Maciej[2], Matthias Klapproth[2], Ibrahima Ndao[3], Sieghard Frischmann[4], Julia Fischer[2], Dietmar Zinner[2], Hsi Liu[5]

1 Work Group Neglected Tropical Diseases, Pathology Unit, German Primate Center, Leibniz-Institute for Primate Research, Göttingen, Germany, **2** Cognitive Ethology Laboratory, German Primate Center, Leibniz-Institute for Primate Research, Göttingen, Germany, **3** Direction du Parc National Niokolo Koba, Tambacounda, Senegal, **4** Mast Diagnostica GmbH, Reinfeld, Germany, **5** National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Diseases Control and Prevention, Atlanta, Georgia, United States of America

* sknauf@dpz.eu

Citation: Knauf S, Barnett U, Maciej P, Klapproth M, Ndao I, Frischmann S, et al. (2015) High Prevalence of Antibodies against the Bacterium *Treponema pallidum* in Senegalese Guinea Baboons (*Papio papio*). PLoS ONE 10(11): e0143100. doi:10.1371/journal.pone.0143100

Editor: Imtaiyaz Hassan, Jamia Millia Islamia, INDIA

Received: July 14, 2015

Accepted: October 30, 2015

Published: November 20, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The funder (Mast Diagnostica GmbH) provided support in the form of salaries for authors [SF], but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing Interests: SF is employed by Mast Diagnostica GmbH, a company that sells treponemal tests (Serodia TP-PA, Espline TP, Mastafluor FTA-

Abstract

The bacterium *Treponema pallidum* is known to cause syphilis (ssp. *pallidum*), yaws (ssp. *pertenue*), and endemic syphilis (ssp. *endemicum*) in humans. Nonhuman primates have also been reported to be infected with the bacterium with equally versatile clinical manifestations, from severe skin ulcerations to asymptomatic. At present all simian strains are closely related to human yaws-causing strains, an important consideration for yaws eradication. We tested clinically healthy Guinea baboons (*Papio papio*) at Parc National Niokolo Koba in south eastern Senegal for the presence of anti-*T. pallidum* antibodies. Since T. pallidum infection in this species was identified 50 years ago, and there has been no attempt to treat non-human primates for infection, it was hypothesized that a large number of West African baboons are still infected with simian strains of the yaws-bacterium. All animals were without clinical signs of treponematoses, but 18 of 20 (90%) baboons tested positive for antibodies against *T. pallidum* based on treponemal tests. Yet, Guinea baboons seem to develop no clinical symptoms, though it must be assumed that infection is chronic or comparable to the latent stage in human yaws infection. The non-active character is supported by the low anti-*T. pallidum* serum titers in Guinea baboons (median = 1:2,560) versus serum titers that are found in genital-ulcerated olive baboons with active infection in Tanzania (range of medians among the groups of initial, moderate, and severe infected animals = 1:15,360 to 1:2.097e+7). Our findings provide evidence for simian infection with *T. pallidum* in wild Senegalese baboons. Potentially, Guinea baboons in West Africa serve as a natural reservoir for human infection, as the West African simian strain has been shown to cause sustainable yaws infection when inoculated into humans. The present study pinpoints an area where further research is needed to support the currently on-going second WHO led yaws eradication campaign with its goal to eradicate yaws by 2020.

ABS IgG, and Mastablot TP IgG) that were used in this study. This does not alter the authors' adherence to all PLOS policies on sharing data and materials.

Introduction

In the mid-1960s, serological surveys demonstrated that the bacterium *Treponema pallidum* infected large numbers of nonhuman primates in Guinea Bissau, Senegal, and Cameroon [1–4]. While the bacterium, which belongs to the order of Spirochaetales, is known to cause syphilis (ssp. *pallidum*), yaws (ssp. *pertenue*, TPE), and endemic syphilis (ssp. *endemicum*) in humans, simian isolates have been reported to cause equally versatile clinical manifestations (reviewed in [5]). At present, all simian isolates are genetically most closely related to human yaws-causing strains [6, 7]. Currently the Fribourg-Blanc simian strain, which was isolated from a baboon in West Africa [4], is the most profoundly characterized simian isolate. It has been whole genome sequenced and due to its similar genetic characteristics, matching those of other TPE strains, it has been proposed to be renamed as *T. pallidum* ssp. *pertenue* strain Fribourg-Blanc [6]. It is furthermore the only simian strain that has been shown to infect humans when inoculated into skin [8], though it should be noted that the indicated study must be considered as ethically questionable. Nevertheless, the findings suggest that simian strains may successfully cross species barriers, an important observation for the ongoing second WHO led yaws eradication campaign [9]. In the 1960's, the isolated West African simian strain was described to cause mild skin lesions in some baboons that included small keratotic lesions and ulcers on the muzzle, eyelids, and armpits. However, most serologically positive animals were free of any clinical symptoms [2].

We tested clinically healthy Guinea baboons (*Papio papio*) at Parc National Niokolo Koba (PNNK) in south eastern Senegal for the presence of anti-*T. pallidum* antibodies. It was hypothesized that infection is present even five decades after its first description in West Africa, especially because the baboons had no history of treatment against the spirochete in this area.

Material and Methods

Ethical statement

All animal work was conducted according to relevant national and international guidelines. Baboon serum samples were taken with permission of the National Parks Direction and the Ministry of Environment and Sustainable Development of Senegal (Attestation 0383/24/03/2009, 0373/10/03/2011, and 1089/02/09/2013). In addition, the Animal Welfare and Ethics Committee of the German Primate Center approved the entire study. 'Good Veterinary Practice' rules were applied to all procedures whenever animals were handled, e.g. during blood sampling. A veterinarian closely monitored anaesthetized animals until they were fully recovered and able to make their way back to their group.

Study site and animals

PNNK is located in the south eastern part of Senegal and bordersGuinea to the south. While poaching threatens several species of native wildlife in the park and has reduced their numbers, the park's population of Guinea baboons is stable and on the rise [10]. The study area lies next to the 'Centre de Recherche de Primatologie' (CRP) at Simenti (GPS N13.026111, W13.294722), which is located next to the Gambia River. The field site is operated by the German Primate Center, and its main focus is on the behavioral ecology, social system, and cognition of Guinea baboons in their natural environment. The study population consists of ~300 baboons of which ~150 are habituated to the close presence of human observers [11]. The home range encompasses about 25 km^2 [11]. Since 2007, behavioral research has been ongoing with students and park rangers performing daily focal observations, following the habituated baboons from 6:00 AM– 12:00 PM and 4:00 PM– 7:00 PM. To track the whereabouts of the

Table 1. Spatial-, demography- and life-time data of animals that were sampled for blood. GPS data indicate the sampling site. All baboons were clinically healthy. The baboon ID ends with the date of sampling. n/m = not measured.

Case	Baboon ID	Sampling Site (GPS Data, Decimal Degrees, N and W)	Body Weight (kg)	Sex
1	3-OSM-25.04.13*	13.03314, -13.28127	20.0	Male
2	4-MST-25.04.13	13.02652, -13.29628	23.0	Male
3	5-MSA-26.04.13	13.02549, -13.29602	22.0	Male
4	6-AMT-28.04.13	13.02959, -13.27823	12.0	Female
5	7-HOK-28.04.13	13.02892, -13.27821	20.0	Male
6	8-SNE-29.04.13	13.01826, -13.28522	21.5	Male
7	9-NDO-30.04.13	13.00957, -13.27582	20.0	Male
8	10-JLA-01.05.13	13.02788, -13.28452	14.0	Female
9	11-BNT-02.05.13	13.03016, -13.28543	n/m	Female
10	15-JHN-09.05.13	13.02344, -13.28742	21.0	Male
11	16-MRM-10.05.13	13.03697, -13.27967	10.0	Female
12	18-FRD-12.05.13	13.02019, -13.28528	22.0	Male
13	1-RBT-23.11.14	13.03775, -13.31649	20.0	Male
14	2-ANT-24.11.14	13.02555, -13.29417	20.5	Male
15	3-FDL-25.11.14	13.02559, -13.29420	20.0	Male
16	4-NDR-26.11.14	13.02559, -13.29420	20.5	Male
17	5-DRK-30.11.14	13.02555, -1329421	19.5	Male
18	6-OSM-01.12.14*	13.02559, -13.29420	19.0	Male
19	9-JKY-07.12.14	13.02558, -13.29417	22.5	Male
20	10-BAA-11.12.14	13.02788, -13.28452	20.0	Male
21	11-FDR-12.12.14	13.01089, -13.26952	21.0	Male

*The same animal sampled in 2013 and 2014

doi:10.1371/journal.pone.0143100.t001

different study groups, a number of adult female and male baboons were collared e.g., with VHF radio transmitters collar devices (M2320, 130 g, ATS, Isanti, MN, USA). In order to collar, remove or exchange transmitters, 20 animals, 4 females and 16 males, underwent routine anesthesia, which allowed access for blood sampling in April to May 2013 (n = 12) and November to December 2014 (n = 9; Table 1, one baboon was captured in 2013 and 2014 and is counted in total as a single individual).

Continuous health data are available from all baboons of the habituated study group. Signs of discomfort (e.g., fatigue or reduced grooming behavior), lesions caused by trauma, or chronic skin alterations or rash were recorded and reported on a regular basis.

Anesthesia and sampling

Baboons were either short-term immobilized while they were ranging in their group, or when trapped in a custom-made baboon trap. Chemical immobilization was achieved using a mixture of 5.5 mg ketamine/kg body weight (bm) (Ketavet, Pfizer, Berlin, Germany), 1.1 mg xylazine/kg bw (Rompun, Bayer, Leverkusen, Germany), and 0.01 mg atropine/kg bm (Atropium sulfuricum, Eifelfango, Bad Neuenahr-Ahrweiler, Germany). Injection was done by intramuscular remote distance injection using blowpipe or cold-gas immobilization rifle.

Body weight was measured using a spring scale and animals were monitored for vital parameters, such as breath and heart frequency as well as internal body temperature, followed by a routine health check with special focus on skin lesions such as chronic skin ulcers, rash, or bone deformation.

Blood sampling was performed away from the other baboons and in the shade. After proper disinfection of the skin the femoral vein was punctured using a closed blood collection system with a 20G needle (S-Monovette-Kanüle 20Gx1/1.5, #85.1160) to protect the researcher from direct contact with blood. Briefly, a total of 27 ml whole blood was collected with one Lithium-Heparine (S-Monovette 9 ml LH, #02.1065.001), one EDTA (S-Monovette 9 ml K3E, #02.1066.001), and one serum collection tube (S-Monovette 9 ml Z, #02.1063). After removal of the needle, firm pressure was applied to the puncture side to prevent formation of hematoma. Blood samples were kept cool utilizing a recoolx-bag (Recoolx Sievers, Bramsche, Germany) until proper storage facilities were reached after return to the field camp. Samples were kept in upright position until sedimentation was achieved.

Sample processing occurred in the late evening of the sampling day. Plasma and serum samples were divided into 1.0-ml aliquots and transferred into 1.5-ml Protein LoBind tubes (Eppendorf, Hamburg, Germany). All steps were performed using sterile syringes and needles. Plasma and serum aliquots were stored at approximately -10°C using a gas-operated freezer (Dometic RC 2200 50mbar, Waeco, Emsdetten, Germany) until samples were exported to the German Primate Center, where they were stored at -80°C until they were used for the different serological tests.

Serology

Validation of serological tests and performance characteristics with baboon sera (sensitivity, specificity, positive and negative predictive values) excluding the detection of cardiolipin can be found elsewhere [12]. Due to sample loss and degradation during export, animals sampled in 2014 were only tested in the field with the treponemal test (TT) ESPLINE TP (Fujirebio Diagnostics Inc., Malvern, PA, USA), whereas baboons sampled in 2013 were tested extensively for the presence of anti-*T. pallidum* antibodies as it is described below.

Treponemal tests

Briefly, serum samples were tested for anti-*T. pallidum* antibodies in the field using the ESPLINE TP (Fujirebio Diagnostics Inc., Malvern, PA, USA) TT. Four other TTs were used to detect antibodies against *Treponema* in one of our German laboratories (SF): Syphilitop Optima (ALL. DIAG S.A.S., Strasbourg, France), Mastafluor FTA-ABS IgM and IgG (Mast Diagnostica, Reinfeld, Germany), Mastablot TP (Mast Diagnostica, Reinfeld, Germany), and Serodia TP-PA (Fujirebio Diagnostics Inc., Malvern, PA, USA). The latter was used to assess anti-*T. pallidum* titres. All TTs were performed as described elsewhere [12].

Non-treponemal tests

Non-treponemal tests (NTTs) were used to detect antibodies to lipid antigens and were performed at the German Primate Center (SK). While the RPR-100 (Biorad, Marnes, France) and the VDRLCHECK Charbon/RPR (ALL. DIAG S.A.S., Strasbourg, France) were used as described elsewhere [12], we also included a cardiolipin IgG (cut-off value $>$ 10 U) (Mast Diagnostica, Reinfeld, Germany) and IgM assay (cut-off value $>$ 8 U) (Mast Diagnostica, Reinfeld, Germany) to further evaluate the presence of IgG and IgM-class antibodies against cardiolipins. Cardiolipin IgG and IgM assays were run at Mast Diagnostica, Reinfeld, Germany (SF). Purified cardiolipin antigens of bovine heart origin were bound to an enzyme immunoassay (EIA) solid phase. The assay procedure as well as baboon sera dilution were performed according to the kit's instruction. For the differentiation between anti-IgG and anti-IgM-antibodies, specific conjugates were used to detect the different antibody-antigen immune complexes. The results were obtained by using an EIA plate reader (Sunrise, Tecan, Männedorf, Switzerland;

filter setting (450/620 nm)), calculating the serum concentration from a 4-parameter calibration curve, which was performed for the IgG and IgM assay separately.

Statistics

Statistical analyses were performed using Prism 6.0 (GraphPad Software, La Jolla, CA, USA). Endpoint titers of exponential scale were $\log_{10}$ transformed to reduce variance. In case of non-Gaussian distribution and log transformation, zero-titers were converted into 10e-14. Normal distribution was tested using the D'Agostino & Pearson omnibus normality test and the Shapiro-Wilk normality test. Antibody titers from this study were compared to the results of anti-*T. pallidum* titers from a different study [12], which were assorted into four different baboon categories: clinically non-affected, initially, moderately, and severely genital-ulcerated [13]. The non-parametric Kruskal-Wallis test was applied to the $\log^{10}$-transformed data sets of the antibody titers. Each mean rank was compared to the mean rank of every other group. Dunn's correction for multiple comparisons (significance without confidence intervals) was applied to the test. In all tests, $p \leq 0.05$ was considered statistically significant.

Results

Sero-positivity is not associated with clinical signs of infection

No signs of chronic skin ulceration or rash have been reported from baboons in the study area at PNNK since behavioral research started in 2007. Also, animals that were chemically immobilized were thoroughly clinically inspected and showed no abnormalities that could be associated with *T. pallidum* infection. The only skin lesions found were due to acute or partly healed skin wounds from trauma, especially in elderly males.

High sero-prevalence in Guinea baboons at PNNK

Although all animals were without clinical signs of treponematosis, 18 of 20 (90%) examined baboons were tested positive for antibodies against *T. pallidum* based on the outcome of the TTs. The only exception was one female of the 12 animals tested in 2013 and one male baboon of the 9 animals tested in 2014. The female was negative for antibodies against *T. pallidum* in all serological tests (case 4, Table 2, S1 Table), while the male baboon in 2014 was only tested using the ESPLINE TP.

Results of the Mastafluor FTA-ABS IgG were in accordance to the ESPLINE TP findings, while Mastafluor FTA-ABS IgM detected antibodies of the IgM class against *T. pallidum* in only a single baboon (8.3%, case 6, Table 2, S1 Table). The same finding was present when serum was tested with the Mastablot TP for IgG and IgM. Again, the female 6-AMT-28.04.13 (case 4, Table 2, S1 Table) was negative for IgG antibodies against the pathogen, whereas all other baboons tested positive. IgM, however, was detected in only a single animal using Mastablot TP IgM. This animal (case 10, Table 2, S1 Table) was different from the one that was identified to have IgM antibodies against *T. pallidum* in the Mastafluor FTA-ABS IgM test (case 6, Table 2, S1 Table).

VDRLCHECK and RPR-100 results differ in two cases (2 and 7, Table 2, S1 Table). In both cases the VDRLCHECK result was positive, while the RPR-100 test result was negative. The female that was negative in all TTs was equally negative in both NTTs. Antibodies of the IgM class directed against cardiolipins exceeded the cut-off value in 3 animals (case 1, 6, and 7, Table 2) and thus were considered to be positive. Only one animal had a positive outcome for IgG anti-cardiolipin antibodies against the spirochete (case 1, Table 2, S1 Table).

Table 2. Crosstable of the results of seven treponemal tests and four non-treponemal tests that were used to detect anti-*T. pallidum* antibodies in the Senegalese baboons sampled in 2013. Case 13–21 (Table 1) are not included.

		Treponemal Tests							Non-Treponemal Tests			
Case	Baboon ID	Espline TP	Syphili-top Optima	Serodia TP-PA (Titer 1:x)	Mastafluor FTA-ABS IgG	Mastafluor FTA-ABS IgM	Masta-blot TP IgG	Masta-blot TP IgM	VDRL-CHECK	RPR-100	Cardiolipin IgG (Co > 10 U)	Cardiolipin IgM (Co > 8 U)
1	3-OSM-25.04.13	+	+	1,280	+	-	+	-	+	+	11.27	8.82
2	4-MST-25.04.13	+	+	1,280	+	-	+	-	+	-	6.71	3.58
3	5-MSA-26.04.13	+	+	2,560	+	-	+	-	+	+	6.48	6.95
4	6-AMT-28.04.13	-	-	0	-	-	-	-	-	-	5.16	4.99
5	7-HOK-28.04.13	+	+	10,240	+	-	+	-	+	+	6.09	7.96
6	8-SNE-29.04.13	+	+	10,240	+	+	+	-	+	+	7.93	9.80
7	9-NDO-30.04.13	+	+	1,280	+	-	+	-	+	-	7.02	10.59
8	10-JLA-01.05.13	+	+	10,240	+	-	+	-	-	-	4.79	3.73
9	11-BNT-02.05.13	+	+	20,480	+	-	+	-	+	+	8.76	4.13
10	15-JHN-09.05.13	+	+	1,280	+	-	+	-	-	-	4.64	4.27
11	16-MRM-10.05.13	+	+	2,560	+	-	+	+	-	-	5.13	1.79
12	18-FRD-12.05.13	+	+	5,120	+	-	+	-	-	-	5.85	2.84

doi:10.1371/journal.pone.0143100.t002

Anti-*T. pallidum* antibody titers are not as high as in East African baboons

In all sero-positive baboons that were sampled in 2013, antibody titers had a median titer of 1:2,560 with a range of 1:1,280 to 20,480 (S1 Table). The $\log_{10}$-transformed data for the antibody titers obtained from the PNNK baboons did not follow a Gaussian distribution and thus were analyzed using a non-parametric test. The comparison of anti-*T. pallidum* antibody titers obtained from clinically non-affected (CNA) sero-positive Guinea baboons showed no significant difference to the results of CNA olive baboons at Lake Manyara National Park in Tanzania, as they are published elsewhere ([12], Fig 1). However, there is a tendency of higher antibody titers in the CNA baboons from Senegal with a median of 1:2,560 (range 0 to 1:20,480) compared to the titers of CNA Tanzanian baboons (median = 0, range 0 to 1:81,920). The same applies when compared to the group of initially (median = 1:15,360, range 0 to 1:81,920) and moderately genital-ulcerated and *T. pallidum* infected baboons of East Africa (median = 1:40,960, range 1:1,280 to 1.311e+6). In our analysis, there was a significantly lower antibody titer in the baboons sampled in 2013 at PNNK compared to severely infected (SEV) baboons at Lake Manyara National Park (p = 0.0004, SEV median = 1:81,920, range 1:20,480 to 2.097e+7; Fig 1).

Discussion

Transmission of *T. pallidum* in Senegalese baboons remains elusive

T. pallidum infection in baboons has been reported in many areas of tropical Africa [14]. Although the high number of *T. pallidum* sero-positive Guinea baboons at PNNK (90%, n = 18/20) was expected based on reports of the 1960s to 70s [1–4], it is currently not clear how *T. pallidum* is transmitted within a population of clinically non-affected baboons. In humans, yaws affects primarily young children and is generally transmitted through direct skin contact [15]. In primary yaws, lesions can be found commonly on the legs and ankles but can also be on arms and face. The primary lesion usually heals up within 3–6 months and is followed by secondary lesions that result from lymphatic and hematogenous spread (reviewed in [15]). These lesions are teeming with spirochetes and support the transmission of the pathogen by close contact with a susceptible new host. However, skin ulcers in baboons have not been reported from PNNK since behavioral research started in 2007. We have not yet detected treponemal DNA from blood collected in 2013 and 2014, though analysis is currently ongoing. The possibility to not only identify pathogen specific antibodies against *T. pallidum*, but also to study the genetic information of the simian strain(s) circulating in south east Senegal, would provide a better understanding of disease ecology and transmission pathways.

All simian T. pallidum strains analyzed thus far are genetically closely related to human yaws-causing strains [6, 7] and the Fribourg-Blanc simian strain must be considered as the most likely candidate to infect the baboons at PNNK. Historically, it was isolated from a baboon in West Africa and infection was not commonly associated with clinical manifestations [2]. It is possible that some baboons did sero-convert before macroscopic lesions became visible. This, however, is mostly unlikely since baboons are long-term monitored and all animals had no history of skin lesions before and after blood was sampled in 2013 and 2014. PNNK's habituated baboons allow researchers to approach them very closely, sometimes closer than 2 meters of distance. If present, skin lesions would have been identified during some of the daily focal observations.

In humans, *T. pallidum* is recognized as an obligate pathogen, which has a very low infectious dose of ~50 inoculated microorganisms (at least for the syphilis causing ssp. *pallidum*,

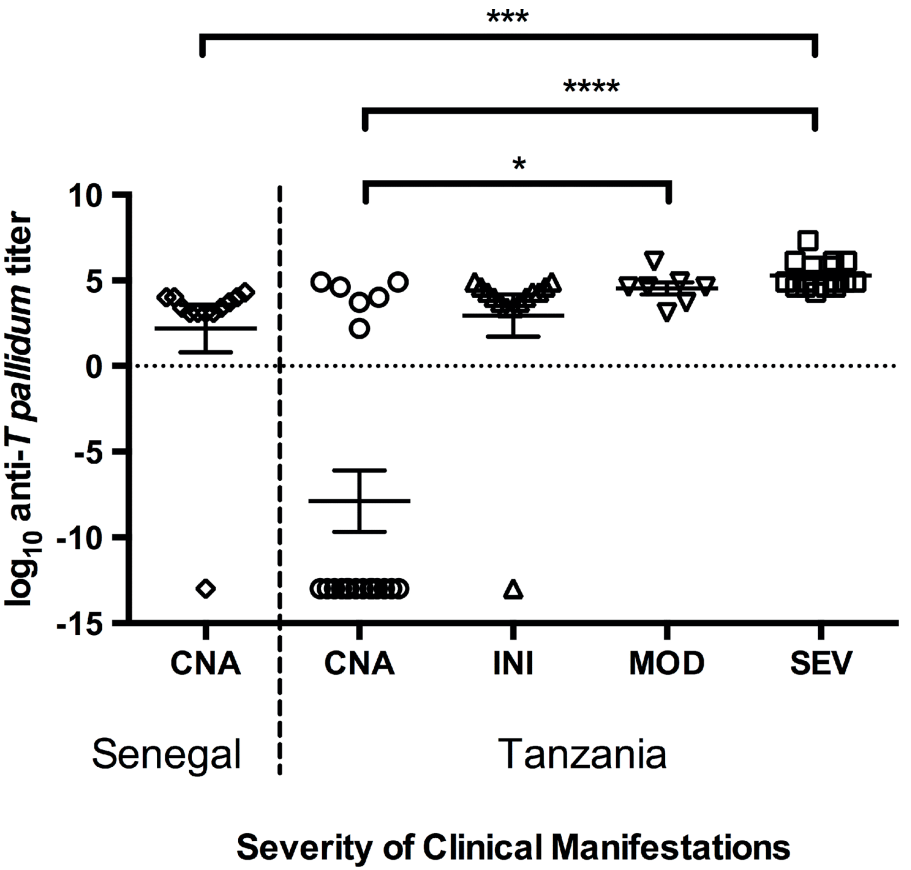


Fig 1. Overview of the anti-*T. pallidum* antibody titers of the Senegalese baboons sampled in 2013 compared to antibody titers obtained from *T. pallidum* infected baboons at Lake Manyara National Park in Tanzania [12]. Anti-*T. pallidum* antibody quantification was investigated using the Serodia TP-PA. Multiple comparison of log anti-*T. pallidum* titers in the group of baboons from PNNK in Senegal (CNA = clinically non-affected (n = 12), and four groups of baboons from Lake Manyara National Park in Tanzania with different stages of genital ulceration (CNA (n = 20), INI = initially (n = 14), MOD = moderately (n = 7), and SEV = severely genital-ulcerated (n = 16); data from [11], for stage definition see [12]). Kruskal-Wallis test using Dunn's correction for multiple comparison: CNA (Tanzania) vs. SEV mean rank diff. = -38.96, ****p < 0.0001; CNA (Senegal) vs. SEV mean rank diff. = -31.23, ***p = 0.0004; CNA (Tanzania) vs. MOD mean rank diff. = -26.97, *p = 0.0198. (Mean ± SEM)

doi:10.1371/journal.pone.0143100.g001

[16]). Congenital infection is a common feature in syphilis [17] but rarely reported in yaws [18]. It is not known whether simian strains can be transmitted from an infected mother via the placenta to an unborn child, but neither at the PNNK nor in the clinically affected East African baboon population has an increase in neonatal or sub-adult mortality rate been noted.

Future studies at PNNK should include representative sampling of all age classes and long-term serological monitoring of individuals, starting at an early age and considering pedigrees. This could help to identify possible patterns of sero-conversion and may contribute to the identification and modeling of possible transmission routes in Guinea baboons.

Noteworthy, there is some evidence that other monkey species may in addition play a role as a reservoir for *T. pallidum* in Senegal. Interestingly, some green monkeys (*Chlorocebus sabaeus*) at PNNK do show signs of clinical infection that represent classic yaws-induce lesions in the face (Fig 2). A major task for future research activities must therefore include sampling of clinically infected green monkeys. Transmission between the latter and baboons may occur when juveniles play together or in the rare event, when adult baboons prey on green monkeys, as described for other baboon species [19]. It is possible that different species develop different clinical manifestations upon infection with the spirochete. This again underscores the importance to enhance research activities on transmission routs. Baboons and green monkeys tend to share resources (e.g., food and water) with humans. Even if monkey consumption has no recent history in Senegal, nonhuman primates may often be chased and killed to minimize crop raiding. At the same time, hunting provides a major potential source of cross-species infection with direct skin contact. The likelihood that infection in nonhuman primates is wide spread across West Africa, plus the tradition of primate bush meat hunting in neighboring countries such as Guinea, may provide a potential hotspot of nonhuman primate to human infection.

Serology indicates a latent stage infection

All tests that were used in this study have been validated for the use in baboons [12]. Individual-based testing results of the TTs were constant for anti-*T. pallidum* antibodies of the IgG class in all assays. However, test results for IgM differed and were not always consistent. For example, one baboon was tested positive with the Mastafluor FTA-ABS IgM (case 6, Table 2, S1 Table), but tested negative with the Mastablot TP IgM. The opposite applied to case 11 (Table 2, S1 Table). All other animals were negative for IgM anti-*T. pallidum* antibodies. IgG antibodies against the spirochete are known to be stable and persistent throughout life [20], which correlates well with the reads of the IgG detecting TTs in this study. In contrast, IgM antibodies are the first to be produced post-infection but have a lower amplitude than IgG titers [21] and a reduced half-life compared to IgG [22]. In human treponematoses, antigen-specific IgM antibodies start to decline ~8 weeks post infection and may reach baseline levels after 2 years of infection [20]. Due to the high sero-prevalence of *T. pallidum* in baboons at PNNK, it can be assumed that infection is well established in the population and most animals are long-term infected with IgM titers that most likely reached baseline level already. However, it must be taken into consideration that at least some of the IgM results are false negative. IgM is sensitive to freeze-thaw cycles, which can contribute to the formation of cryoglobulins [23]. In this case, IgM may no longer be detected as it is precipitated. No free-floating detectable IgM will be found in such sera.

The inconsistent results of the cardiolipin EIA when compared to the corresponding NTT results (Table 2, S1 Table) may be due to the fact that cardiolipin is immobilized onto a solid phase using co-factors such as bovine sera. Co-factor related antigen structures, however, seem to be more specific for (human) phospholipid syndrome detection than for the use as a NTT. Thus, antibodies binding to VDRL and RPR may not be detected to the same level by an EIA.

Although Guinea baboons seem to develop no clinical symptoms of treponematosis, it is likely that infection is chronic or comparable to the latent stage in human yaws infection. This is supported by the low anti-*T. pallidum* serum titers in asymptomatic Guinea baboons

Fig 2. An adult green monkey (*Chlorocebus sabaeus*) with facial lesions at PNNK. The clinical manifestations resemble lesions known from tertiary human yaws infection.

doi:10.1371/journal.pone.0143100.g002

(median = 1:2,560) *versus* serum titers that are found in genital-ulcerated baboons with active infection (median INI = 1:15,360, MOD = 1:40,960, and SEV = 1:81,920, [12]). Yet, low but consistent antibody levels again raise the question, why Senegalese Guinea baboons do not

develop yaws- or syphilis-like lesions. This might be explained by strain specific genetic factors, which we do not understand yet and which are subject to functional analyses once the genomes of a substantial number of *T. pallidum* simian isolates have been sequenced and compared to human yaws and syphilis strains. In addition, host factors might be responsible for subtle changes in the immunological response, which is known to be a main driver for tissue damage during active infection in *T. pallidum* infection [24]. It is also possible that some species have coevolved with the bacterium and thus show different clinical manifestations through adaptation. However, Guinea baboons are parapatric with western olive baboons and hybridization among both species is likely [25–27]. It would thus be important to include western olive baboons in future investigations.

Conclusion

Five decades after its first description from West Africa, south-eastern Senegalese Guinea baboons are (still) infected with *T. pallidum*, which is demonstrated by the specific antibodies against the spirochete. Though molecular detection of spirochete DNA is currently under investigation, the most likely strain that infects baboons in West Africa is the Fribourg-Blanc simian strain. This strain has been whole genome sequenced and shows close relationship to human yaws-causing *T. pallidum* strains [6]. In addition, it has been demonstrated to be infectious to humans [8]. This underlines the theoretic potential for Guinea baboons in West Africa, especially in areas where bush meat is consumed, to serve as a natural reservoir of human yaws infection. The reported results from our investigation in West Africa are of importance for the on-going second WHO led yaws eradication campaign [28] with its goal to eradicate yaws by 2020 [29].

Supporting Information

S1 Table. Individual-based dataset on clinical manifestations and serological test results.
(XLSX)

Acknowledgments

We thank Wildlife Services of Senegal, National Parks Direction, Water and Forestry Direction and Direction of Veterinary Services. Heike Klensang (DPZ) is thanked for her support to import samples to Germany. Franziska Dahlmann and Simone Lüert are thanked for their support in the laboratory.

Author Contributions

Conceived and designed the experiments: SK PM SF HL. Performed the experiments: SK UB PM MK IN SF. Analyzed the data: SK SF JF DZ HL. Contributed reagents/materials/analysis tools: SK JF SF. Wrote the paper: SK IN SF JF DZ HL.

References

1. Baylet R, Thivolet J, Sepetjian M, Bert J. Seroepidemiological studies on primate treponematosis in Senegal. Bull Soc Pathol Exot Filiales. 1971; 64: 836–41. PMID: 5005249
2. Baylet R, Thivolet J, Sepetjian M, Nouhouay Y, Baylet M. Natural open treponematosis in the *Papio papio* baboon in Casamance. Bull Soc Pathol Exot Filiales. 1971; 64: 842–6. PMID: 4949450
3. Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. Primates Med. 1969; 3: 113–21. PMID: 5006024
4. Fribourg-Blanc A, Mollaret HH, Niel G. Serologic and microscopic confirmation of treponemosis in Guinea baboons. Bull Soc Pathol Exot Filiales. 1966; 59(1): 54–9. PMID: 5333741

5. Harper K, Knauf S. *Treponema pallidum* in primates: Clinical manifestations, epidemiology, and evolution of a stealthy pathogen. In: Brinkworth JF, Pechenkina K, editors. Primates, pathogens, and evolution. Developments in primatology: Progress and prospects. 38: Springer New York; 2013. pp. 189–219.
6. Zobanikova M, Strouhal M, Mikalova L, Cejkova D, Ambrozova L, Pospisilova P, et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl Trop Dis. 2013; 7: e2172. doi: 10.1371/journal.pntd.0002172 PMID: 23638193
7. Harper KN, Fyumagwa RD, Hoare R, Wambura PN, Coppenhaver DH, Sapolsky RM, et al. *Treponema pallidum* Infection in the wild baboons of East Africa: Distribution and genetic characterization of the strains responsible. PLoS ONE. 2012; 7: e50882. doi: 10.1371/journal.pone.0050882 PMID: 23284649
8. Smith JL, David NJ, Indgin S, Israel CW, Levine BM, Justice J Jr., et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br J Vener Dis. 1971; 47: 226–51. PMID: 4936861
9. Maurice J. WHO plans new yaws eradication campaign. Lancet. 2012; 379: 1377–8. PMID: 22509526
10. Galat-Luong A, Galat G, Hagell S. The social and ecological flexibility of Guinea baboons: implications for Guinea baboon social organization and male strategies. In: Swedell L, Leigh S, editors. Reproduction and fitness in baboons behavioral, ecological, and life history perspectives. New York: Springer; 2006. pp. 105–121.
11. Patzelt A, Kopp GH, Ndao I, Kalbitzer U, Zinner D, Fischer J. Male tolerance and male-male bonds in a multilevel primate society. Proc Natl Acad Sci U S A. 2014; 111: 14740–5. doi: 10.1073/pnas.1405811111 PMID: 25201960
12. Knauf S, Dahlmann F, Batamuzi EK, Frischmann S, Liu H. Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Negl Trop Dis. 2015; 9: e0003637. doi: 10.1371/journal.pntd.0003637 PMID: 25803295
13. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IA, Nordhoff M, et al. *Treponema* infection associated with genital ulceration in wild baboons. Vet Pathol. 2012; 49: 292–303. doi: 10.1177/0300985811402839 PMID: 21411621
14. Knauf S, Liu H, Harper KN. *Treponemal* infection in nonhuman primates may serve as a potential reservoir for human yaws. Emerg Infect Dis. 2013; 19: 2058–60. doi: http://dx.doi.org/10.3201/eid1912.130863 PMID: 24274094 doi: 10.3201/eid1912.130863
15. Mitja O, Asiedu K, Mabey D. Yaws. Lancet. 2013; 381: 763–73. doi: 10.1016/S0140-6736(12)62130-8 PMID: 23415015
16. Magnuson HJ, Thomas EW, Olansky S, Kaplan BI, De Mello L, Cutler JC. Inoculation syphilis in human volunteers. Medicine (Baltimore). 1956; 35: 33–82. PMID: 13296652
17. Peterman TA, Su J, Bernstein KT, Weinstock H. Syphilis in the United States: on the rise? Expert review of anti-infective therapy. 2015; 13: 161–8. doi: 10.1586/14787210.2015.990384 PMID: 25487961
18. Marks M, Mitja O, Solomon AW, Asiedu KB, Mabey DC. Yaws. Br Med Bull. 2015; 113: 91–100. doi: 10.1093/bmb/ldu037 PMID: 25525120
19. Hausfater G. Predatory behaviour of yellow baboons. Behaviour. 1976; 56:44–68. doi: 10.2307/4533713
20. Peeling RW, Ye H. Diagnostic tools for preventing and managing maternal and congenital syphilis: an overview. Bull World Health Organ. 2004; 82: 439–46. PMID: 15356937
21. Salazar JC, Hazlett KR, Radolf JD. The immune response to infection with *Treponema pallidum*, the stealth pathogen. Microbes Infect. 2002; 4: 1133–40. doi: 10.1016/S1286-4579(02)01638-6 PMID: 12361913
22. Teunis PF, van Eijkeren JC, Ang CW, van Duynhoven YT, Simonsen JB, Strid MA, et al. Biomarker dynamics: estimating infection rates from serological data. Statistics in medicine. 2012; 31: 2240–8. doi: 10.1002/sim.5322 PMID: 22419564
23. Delobel J, Rubin O, Prudent M, Crettaz D, Tissot JD, Lion N. Biomarker analysis of stored blood products: emphasis on pre-analytical issues. Int J Mol Sci. 2010; 11: 4601–17. doi: 10.3390/ijms11114601 PMID: 21151459
24. Carlson JA, Dabiri G, Cribier B, Sell S. The immunopathobiology of syphilis: the manifestations and course of syphilis are determined by the level of delayed-type hypersensitivity. Am J Dermatopathol. 2011; 33: 433–60. doi: 10.1097/DAD.0b013e3181e8b587 PMID: 21694502
25. Zinner D, Groeneveld LF, Keller C, Roos C. Mitochondrial phylogeography of baboons (*Papio* spp.): indication for introgressive hybridization? BMC Evol Biol. 2009; 9:83. doi: 10.1186/1471-2148-9-83 PMID: 19389236

26. Zinner D, Buba U, Nash S, Roos C. Pan-African voyagers. The phylogeography of baboons. In: Sommer V, Ross C, editors. Primates of Gashaka socioecology and conservation in Nigeria's biodiversity hotspot. New York: Springer; 2011. pp. 267–306.

27. Zinner D, Wertheimer J, Liedigk R, Groeneveld LF, Roos C. Baboon phylogeny as inferred from complete mitochondrial genomes. Am J Phys Anthropol. 2013; 150: 133–40. doi: 10.1002/ajpa.22185 PMID: 23180628

28. Marks M, Mitja O, Vestergaard LS, Pillay A, Knauf S, Chen CY, et al. Challenges and key research questions for yaws eradication. The Lancet infectious diseases. 2015; 31: 2240–8. doi: 10.1016/S1473-3099(15)00136-X PMID: 26362174

29. The World Health Organization. Eradication of yaws—the Morges strategy. Releve epidemiologique hebdomadaire/Section d'hygiene du Secretariat de la Societe des Nations = Weekly epidemiological record/Health Section of the Secretariat of the League of Nations. 2012; 87: 189–94. PMID: 24340400

7.3.2 Endemicity of Yaws and Seroprevalence of *Treponema pallidum* Antibodies in Nonhuman Primates, Kenya

prolonged CTLA4 therapy. Open Forum Infect Dis. 2017;4:ofx100. https://doi.org/10.1093/ofid/ofx100
5. Muftuoglu M, Olson A, Marin D, Ahmed S, Mulanovich V, Tummala S, et al. Allogeneic BK virus-specific t cells for progressive multifocal leukoencephalopathy. N Engl J Med. 2018;379:1443–51. https://doi.org/10.1056/NEJMoa1801540

Address for correspondence: Arnaud Del Bello, CHU Rangueil, Nephrology Dialysis and Organ Transplant, 1 Ave Jean Poulhès, Toulouse 31059, France; email: delbello.a@chu-toulouse.fr

Endemicity of Yaws and Seroprevalence of *Treponema pallidum* Antibodies in Nonhuman Primates, Kenya

Dawn M. Zimmerman, Emily H. Hardgrove, Michael E. von Fricken, Joseph Kamau, Daniel Chai, Samson Mutura, Velma Kivali, Fatima Hussein, Peris Ambala, Andrea Surmat, Joseph G. Maina, Sascha Knauf

Author affiliations: Smithsonian Conservation Biology Institute, Washington DC, USA (D.M. Zimmerman, E.H. Hardgrove, M.E. von Fricken); George Mason University, Fairfax, Virginia, USA (D.M. Zimmerman, M.E. von Fricken); Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA (E.H. Hardgrove); Institute of Primate Research, Nairobi, Kenya (J. Kamau, D. Chai, S. Mutura, F. Hussein, P. Ambala); International Livestock Research Institute, Nairobi (V. Kivali); Mpala Research Centre and Wildlife Foundation, Nanyuki, Kenya (A. Surmat); Kenya Wildlife Service, Nairobi (J.G. Maina); German Primate Center, Goettingen, Germany (S. Knauf)

DOI: https://doi.org/10.3201/eid2511.190716

Human yaws has historically been endemic to Kenya, but current epidemiologic data are lacking. We report seroprevalence for *Treponema pallidum* antibodies in olive baboons (*Papio anubis*) and vervet monkeys (*Chlorocebus pygerythrus*) in Laikipia County, Kenya. Our results suggest endemicity of the yaws bacterium in monkeys, posing a possible zoonotic threat to humans.

Yaws is a disease caused by the bacterium *Treponema pallidum* subsp. *pertenue*, which is believed to be an exclusively human pathogen (*1*). However, this bacterium has recently been identified in African nonhuman primates (NHPs) (*2*), raising concerns about a possible zoonotic reservoir for human infection. Kenya is 1 of 76 countries that the World Health Organization categorizes as previously endemic for yaws, but no current data support its presence or absence (http://apps.who.int/gho/data/node.main.NTDYAWSEND). However, sustainable yaws eradication will rely on information about transmission dynamics and potential links between human and NHP *T. pallidum* strains (*3*).

In the early 1960s, Fribourg-Blanc and Mollaret tested 150 serum samples from wild-caught baboons (*Papio* sp.) from Guinea and Kenya (*4*). Although 72 (65%) of 111 serum samples from Guinea were positive for *T. pallidum* antibodies, none of the samples from Kenya were positive. In subsequent years, an additional 276 serum samples from baboons in Kenya supported the absence of *T. pallidum* infection. However, a more recent study of baboon samples collected during 1977–1994 in Kenya reported serologic evidence of *T. pallidum* infection in Nanyuki, Laikipia County (prevalence 57.5%) (*5*). For our study, we hypothesized that 39 years after the first samples were positive for antibodies against *T. pallidum* in Nanyuki (*5*), infection is still present in the NHP population.

All animal protocols were approved by the Kenya Wildlife Service (permit #4004), the Institute of Primate Research Scientific and Ethics Review Committee, and the Smithsonian Institution Animal Use and Care Committee. In October 2016, we sampled 65 olive baboons (*Papio anubis*) and 2 vervet monkeys (*Chlorocebus pygerythrus*) at sites surrounding the Mpala Research Centre in Laikipia County, Kenya. We performed a preliminary serologic screening by using the immunochromatographic Dual Path Platform (DPP) HIV-Syphilis Assay (Chembio Diagnostic Systems, Inc., http://chembio.com) according to the manufacturer guidelines. This syphilis (*T. pallidum*) assay is a useful screening tool because antibodies against *Treponema* subspecies are cross-reactive (*6*). We tested 67 samples with the DPP assay; 49 were positive and 18 negative.

However, because this test is not certified for use with NHPs, we subsequently confirmed results by using the *T. pallidum* Particle Agglutination Assay (TPPA) (SERODIA TPPA, https://www.fujirebio-us.com), which has been validated for use in baboons (*7*). Of the 52 samples tested with the TPPA assay, there were 33 positive, 6 negative, and 13 inconclusive results. Inconclusive TPPA results indicate nonspecific antibodies reacting with nonsensitized particles. Because of limited sample material, we were unable to perform repeated testing with a preabsorption step to remove all nonspecific binding antibodies (as described in the assay manual) and therefore excluded the inconclusive TPPA results from our analysis.

If we defined seropositive monkeys as those with positive results for the TPPA or DPP, 1 of 2 vervet monkeys and 53 (85.5%) of 62 baboons were seropositive. Male baboons (90.4%, 38/42) had a relative seropositivity risk ratio of 1.3 (95% CI 0.984–1.858) when compared with female baboons (72.2%, 13/18); however, this difference was not significant (p = 0.111 by Fisher exact test). If we included age, in addition to sex, in the analysis, adult male and female baboons both showed 100% seropositivity (21/21 and 10/10, respectively). Subadult males and females also showed seropositivity of 100% (6/6 and 1/1, respectively). Juveniles had a combined seropositivity of 61.1%: a total of 81.8% (9/11) of males and 28.6% (2/7) of females were seropositive. Infants had the lowest seroprevalence rate (50%, 2/4) (Table).

None of the tested NHPs had overt clinical signs of infection, such as skin lesions, which might have contained *T. pallidum* DNA. However, several other studies found that NHPs are frequently seropositive for *T. pallidum* antibodies without clinical lesions (*5,8,9*). Because wild NHPs are not treated and bacterial clearance is unlikely, the absence of lesions presumably corresponds to the latency stage of infection, which is also a key characteristic of human treponematoses (*10*). Future molecular investigations should include nontreponemal tests to further support the assumption that animals are in the latency stage and should target the DNA of the pathogen, which would enable comparison of *T. pallidum* strains of NHP origin from Kenya with those infecting NHPs in neighboring countries and possibly humans. In Tanzania, a country that has a similar history of previous yaws endemicity in humans and lacks current prevalence data, clinical lesions have been documented in olive baboons, vervet monkeys, yellow baboons, and blue monkeys, in addition to widespread seroprevalence in NHPs closely matching previous human infection geographic distribution (*9*).

Our results suggest that evidence of *Treponema* exposure in NHPs continues to be present in Laikipia County almost 4 decades after it was first detected. Our data provide further evidence that, in East Africa, *T. pallidum* infection is endemic to NHPs and that multiple NHP taxa contain antibodies indicating latent infection. Providing reliable information on the epidemiology of treponematoses in humans and NHPs has major programmatic implications for yaws eradication. Under a One Health approach, we call for additional yaws surveillance in communities in Kenya, especially in regions where NHPs and humans coexist.

Table. Demographic data and serologic results for nonhuman primates sampled for *Treponema pallidum* antibodies, Laikipia County, Kenya, October 2016*

Species, age group†	No. positive/no. tested (%)	
	Male	Female
Olive baboon (*Papio anubis*)		
Adult	21/21 (100)	10/10 (100)
Subadult	6/6 (100)	1/1 (100)
Juvenile	9/11 (82)	2/7 (29)
Infant	2/4 (50)	ND
Subtotal	38/42 (90)	13/18 (72)
Vervet monkey (*Chlorocebus pygerythrus*)		
Adult	0/1 (0)	ND
Juvenile	1/1 (100)	ND
Subtotal	1/2 (50)	ND
Total	39/44 (89)	13/18 (72)

*Samples were tested by using the Dual Path Platform Assay or the *Treponema pallidum* Particle Agglutination Assay. ND, not done.
†Age ranges for *P. anubis* baboons, infant, <1.3 y; male juvenile, 1.3–6 y; female juvenile, 1.3–5 y; male subadult, 6–9 y; female subadult 5–6 y; male adult, >10 y; female adult, >6 y (Appendix reference *1*, https://wwwnc.cdc.gov/EID/article/25/11/19-0716-App1.pdf). Age ranges for *C. pygerythrus* monkeys: juvenile, 22–40 mo; adult, ≥40 months (Appendix reference *2*).

Acknowledgments

We thank the Kenya Wildlife Service and Mpala Research Centre for assistance during this study.

This study was supported by the US Agency for International Development Emerging Pandemic Threats PREDICT Project (cooperative agreement mo. GHN-A-OO-09-00010-00) and the German Research Foundation (grant DFG KN 1097/3-1 to S.K.).

About the Author

Dr. Zimmerman is director of wildlife health and associate program director for the Smithsonian Conservation Biology Institute's Global Health Program, Washington, DC, and country lead for the US Agency for International Development Emerging Pandemic Threats PREDICT program in Kenya. Her primary research interests include applying a One Health approach to the conservation of critically endangered wildlife species and the mitigation of emerging infectious diseases at the wildlife–human interface.

References

1. Marks M, Lebari D, Solomon AW, Higgins SP. Yaws. Int J STD AIDS. 2015;26:696–703. https://doi.org/10.1177/0956462414549036
2. Knauf S, Gogarten JF, Schuenemann VJ, De Nys HM, Düx A, Strouhal M, et al. Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. Emerg Microbes Infect. 2018;7:157. https://doi.org/10.1038/s41426-018-0156-4
3. Bodimeade C, Marks M, Mabey D. Neglected tropical diseases: elimination and eradication. Clin Med (Lond). 2019;19:157–60. https://doi.org/10.7861/clinmedicine.19-2-157
4. Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. Primates Med. 1969;3:113–21.
5. Harper KN, Fyumagwa RD, Hoare R, Wambura PN, Coppenhaver DH, Sapolsky RM, et al. *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS One. 2012;7:e50882. https://doi.org/10.1371/journal.pone.0050882
6. Chi KH, Danavall D, Taleo F, Pillay A, Ye T, Nachamkin E, et al. Molecular differentiation of *Treponema pallidum* subspecies in skin ulceration clinically suspected as yaws in Vanuatu using real-time multiplex PCR and serological methods. Am J Trop Med Hyg. 2015;92:134–8. https://doi.org/10.4269/ajtmh.14-0459

7. Knauf S, Dahlmann F, Batamuzi EK, Frischmann S, Liu H. Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Negl Trop Dis. 2015;9:e0003637. https://doi.org/10.1371/journal.pntd.0003637
8. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IA, Nordhoff M, et al. *Treponema* infection associated with genital ulceration in wild baboons. Vet Pathol. 2012;49:292–303. https://doi.org/10.1177/0300985811402839
9. Chuma IS, Batamuzi EK, Collins DA, Fyumagwa RD, Hallmaier-Wacker LK, Kazwala RR, et al. Widespread *Treponema pallidum* infection in nonhuman primates, Tanzania. Emerg Infect Dis. 2018;24:1002–9. https://doi.org/10.3201/eid2406.180037
10. Marks M, Mitjà O, Solomon AW, Asiedu KB, Mabey DC. Yaws. Br Med Bull. 2015;113:91–100. https://doi.org/10.1093/bmb/ldu037

Address for correspondence: Dawn M. Zimmerman, Smithsonian Conservation Biology Institute, National Zoological Park, 3001 Connecticut Ave NW, Washington DC 20008, USA: email: zimmermand@si.edu

Middle East Respiratory Syndrome Coronavirus, Saudi Arabia, 2017–2018

Ahmed Hakawi, Erica Billig Rose, Holly M. Biggs, Xiaoyan Lu, Mutaz Mohammed, Osman Abdalla, Glen R. Abedi, Ali A. Alsharef, Aref Ali Alamri, Samar Ahmad Bereagesh, Kamel M. Al Dosari, Saad Abdullah Ashehri, Waad Ghassan Fakhouri, Saleh Zaid Alzaid, Stephen Lindstrom, Susan I. Gerber, Abdullah Asiri, Hani Jokhdar, John T. Watson

Author affiliations: Ministry of Health, Riyadh, Saudi Arabia (A. Hakawi, M. Mohammed, O. Abdalla, A.A. Alsharef, A.A. Alamri, S.A. Bereagesh, K.M. Al Dosari, S.A. Ashehri, W.G. Fakhouri, S.Z. Alzaid, A. Asiri, H. Jokhdar); Centers for Disease Control and Prevention, Atlanta, Georgia, USA (E.B. Rose, H.M. Biggs, X. Lu, G.R. Abedi, S. Lindstrom, S.I. Gerber, J.T. Watson)

DOI: https://doi.org/10.3201/eid2511.190726

We characterized exposures and demographics of Middle East respiratory syndrome coronavirus cases reported to the Saudi Arabia Ministry of Health during July 1–October 31, 2017, and June 1–September 16, 2018. Molecular characterization of available specimens showed that circulating viruses during these periods continued to cluster within lineage 5.

Middle East respiratory syndrome (MERS) coronavirus (MERS-CoV) epidemiology in Saudi Arabia is characterized by healthcare-associated outbreaks (*1*,*2*), occasional household-contact transmission (*3*), and sporadic cases without apparent links to other known cases (*4*,*5*). Since 2015, healthcare-associated transmission has decreased as infection prevention and control practices have improved (*6*); however, sporadic cases continue to be reported, often associated with contact with dromedaries (*4*,*7*). Surveillance and routine investigation of recent MERS cases are critical to monitor the epidemiology of this emerging pathogen. We characterized exposures and demographics of MERS cases reported to the Saudi Arabia Ministry of Health during July 1–October 31, 2017, and June 1–September 16, 2018, and performed molecular characterization of available specimens to describe circulating viruses during these periods.

We summarized demographics and exposures using Ministry of Health investigation data. To further characterize exposures among sporadic cases (no known epidemiologic link to a hospital outbreak or known case) reported during July 1–October 31, 2017, we conducted telephone interviews using a standardized questionnaire addressing demographics and activities during the 14 days before symptom onset (exposure period). For deceased or unavailable patients, we interviewed proxies. We did not conduct interviews for cases reported during June 1–September 16, 2018; this period was selected because of specimen availability. Cases were confirmed by testing respiratory specimens with MERS-CoV real-time reverse transcription PCR assays (*8*). We shipped 20 specimens to the US Centers for Disease Control and Prevention for genome sequence analysis (*9*).

During July 1–October 31, 2017, a total of 61 MERS cases were reported from 12 of 13 administrative regions. Median patient age was 50 (range 10–89) years; 43 (70%) were male, and 23 (38%) died. Nine (15%) cases were associated with a hospital outbreak, 10 (16%) were household contacts of known cases, and 42 (69%) were classified as sporadic and further investigated. During November 2017, we interviewed 35 (83%) sporadic case-patients, 9 directly and 26 by proxy; 7 were unavailable. Among the 42 sporadic case-patients, median age was 57 (range 25–89) years; 35 (83%) were male, and 33 (79%) reported underlying conditions, most commonly diabetes (n = 24) and hypertension (n = 23). All were symptomatic and hospitalized; 22 (52%) died. During the exposure period, 21 (50%) sporadic case-patients reported camel contact: 12 had direct contact (touching), 5 indirect contact (visiting a setting with camels or exposure to others with direct camel contact), and 4 contact that could not be further classified. Among patients with camel contact, 6 also reported visiting a healthcare facility without a known

7.3.3 Widespread *Treponema pallidum* Infection in Nonhuman Primates, Tanzania

Widespread *Treponema pallidum* Infection in Nonhuman Primates, Tanzania

Idrissa S. Chuma, Emmanuel K. Batamuzi,[1] D. Anthony Collins, Robert D. Fyumagwa, Luisa K. Hallmaier-Wacker, Rudovick R. Kazwala, Julius D. Keyyu, Inyasi A. Lejora, Iddi F. Lipende, Simone Lüert, Filipa M.D. Paciência, Alexander Piel, Fiona A. Stewart, Dietmar Zinner, Christian Roos, Sascha Knauf

We investigated *Treponema pallidum* infection in 8 nonhuman primate species (289 animals) in Tanzania during 2015–2017. We used a serologic treponemal test to detect antibodies against the bacterium. Infection was further confirmed from tissue samples of skin-ulcerated animals by 3 independent PCRs (*polA*, *tp47*, and *TP_0619*). Our findings indicate that *T. pallidum* infection is geographically widespread in Tanzania and occurs in several species (olive baboons, yellow baboons, vervet monkeys, and blue monkeys). We found the bacterium at 11 of 14 investigated geographic locations. Anogenital ulceration was the most common clinical manifestation; orofacial lesions also were observed. Molecular data show that nonhuman primates in Tanzania are most likely infected with *T. pallidum* subsp. *pertenue*–like strains, which could have implications for human yaws eradication.

The geographic distribution of infection with the bacterium *Treponema pallidum* in nonhuman primates (NHPs) in Africa has been reported to closely match the one seen in human yaws in Africa before the first yaws eradication campaign (*1*). Some Africa countries, such as Tanzania, have a history of human yaws but lack recent epidemiologic data that support elimination (*2*). At the same time, many of these countries report NHP infection with *T. pallidum* strains that are highly similar to the human yaws–causing *T. pallidum* subsp. *pertenue* (TPE) (*3*,*4*; S. Knauf et al., unpub. data, https://www.biorxiv.org/content/early/2017/05/10/135491) and thus make NHP infection an important issue for a One Health approach.

The first published report of *T. pallidum* infection in Tanzanian NHPs came from anogenital ulcerated olive baboons (*Papio anubis*) at Gombe National Park (GNP) in the late 1980s (*5*), followed by cases reported from olive baboons at Lake Manyara National Park (LMNP) (*3*,*6*,*7*) and Serengeti National Park (SNP) (*3*). Clinical manifestations of *T. pallidum* infection in NHPs ranged from asymptomatic to severe skin ulceration mainly affecting the face or genitalia (*8*). Although early serologic investigations conducted by Fribourg-Blanc in West Africa confirmed widespread infection in several NHP species (e.g., baboons [*Papio* sp.], guenons [*Cercopithecus* sp.], red colobus [*Piliocolobus badius*], and chimpanzees [*Pan trogoldytes*]) (*9*), the infection in Tanzania was exclusively reported from olive baboons in northern parts of the country. Despite the close genetic relationship to human yaws–causing TPE strains (*3*,*4*; S. Knauf et al., unpub. data, https://www.biorxiv.org/content/early/2017/05/10/135491), and in the absence of recent reports of human yaws in Tanzania (*10*), it is currently unclear whether NHP strains naturally infect humans.

As a starting point and basis for advanced epidemiologic investigations, our main objective was to investigate the geographic distribution and host species composition of *T. pallidum* infection in free-ranging NHPs in Tanzania. We hypothesized that, in Tanzania, A) NHPs other than olive baboons are infected with the *T. pallidum* bacterium and B) that infection is not restricted to northern parts of the country.

Materials and Methods

Study Design, Sampling Sites, and Animals

We applied a cross-sectional study design using semirandom selection of free-ranging NHPs in selected areas in

Author affiliations: Sokoine University of Agriculture, Morogoro, Tanzania (I.S. Chuma, E.K. Batamuzi, R.R. Kazwala, I.F. Lipende); German Primate Center, Leibniz-Institute for Primate Research, Göttingen, Germany (I.S. Chuma, L.K. Hallmaier-Wacker, S. Lüert, F.M.D. Paciência, D. Zinner, C. Roos, S. Knauf); Jane Goodall Institute, Kigoma, Tanzania (D.A. Collins); Tanzania Wildlife Research Institute, Arusha, Tanzania (R.D. Fyumagwa, J.D. Keyyu); Tanzania National Parks, Arusha (I.A. Lejora); Liverpool John Moores University, Liverpool, UK, and Greater Mahale Ecosystem Research and Conservation Project, Kigoma (A. Piel, F.A. Stewart)

DOI: https://doi.org/10.3201/eid2406.180037

[1]Deceased.

Tanzania. Selection of NHPs was biased toward animals with visible skin ulcers. Sampling took place at Arusha National Park (ANP), GNP, Katavi NP (KNP), LMNP, Mahale NP (MNP), Mikumi NP (MKNP), Ngorongoro Conservation Area (NCA), Ruaha NP (RNP), Selous Game Reserve (SGR), SNP, Tarangire NP (TNP), Udzungwa NP (UNP), and Issa Valley (Issa), as well as Jozani-Chwaka Bay NP–Masingini Forest (JCBNP) on Unguja Island, Zanzibar (Figure 1). We investigated the following species: olive baboon, yellow baboon (*Papio cynocephalus*), blue monkey (*Cercopithecus mitis*), red-tailed monkey (*Cercopithecus ascanius*), vervet monkey (*Chlorocebus pygerythrus*), Udzungwa red colobus (*Piliocolobus gordonorum*), Zanzibar red colobus (*Piliocolobus kirkii*), and Ugandan red colobus (*Piliocolobus tephrosceles*). Using FreeCalc (http://epitools.ausvet.com.au/content.php?page=FreeCalc2), and based on our previous study at LMNP (*6*) that showed a disease prevalence of 85%, we calculated a sample size of ≥4 (expected disease prevalence 85%) to 21 (expected disease prevalence 25%) per sample site as statistically sufficient to demonstrate freedom from *T. pallidum* infection using imperfect tests and allowing for small populations (online Technical Appendix 1, https://wwwnc.cdc.gov/EID/article/24/6/18-0037-Techapp1.pdf).

Anesthesia and Sampling

We studied the animals in accordance with applicable regulations and guidelines (online Technical Appendix 1). The sampling of blood and skin tissue followed a standardized protocol that we previously applied for baboons (*6*,*11*). In brief, the NHPs were chemically immobilized by remote distance injection of 10.0 mg ketamine/kg body mass (Kyron Laboratories, Johannesburg, South Africa) in combination with 0.2 mg/kg medetomidine (Domitor; Pfizer, Berlin, Germany). Anesthetics were intramuscularly injected using a cold-gas immobilization rifle (MOD JM; Dan-Inject ApS, Børkop,

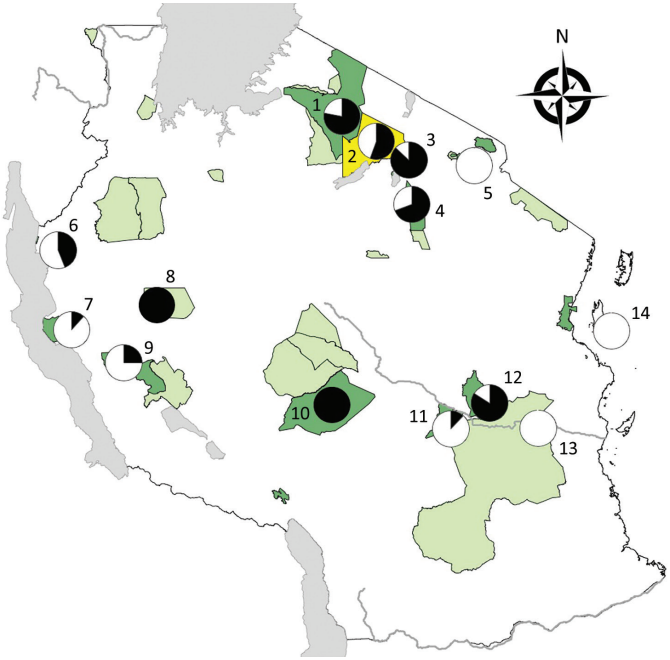


Figure 1. Protected areas and sites where free-ranging nonhuman primates (NHPs) were sampled in a study of *Treponema pallidum* infection, Tanzania. 1, Serengeti National Park (41 NHPs); 2, Ngorongoro Conservation Area (18 NHPs); 3, Lake Manyara National Park (38 NHPs); 4, Tarangire National Park (26 NHPs); 5, Arusha National Park (14 NHPs); 6, Gombe National Park (32 NHPs); 7, Mahale National Park (17 NHPs); 8, Issa Valley (2 NHPs); 9, Katavi National Park (12 NHPs); 10, Ruaha National Park (18 NHPs); 11, Udzungwa National Park (25 NHPs); 12, Mikumi National Park (25 NHPs); 13, Selous Game Reserve (8 NHPs); 14, Jozani-Chwaka Bay National Park–Masingini Forest on Unguja Island, Zanzibar (13 NHPs). Dark green indicates national parks; light green indicates game reserves; yellow indicates conservation area. Circle graphs: black, NHPs *T. pallidum*–positive (serology and/or PCR); white, NHPs T. pallidum–negative (serology and PCR). The map was produced with ArcMap version 10.0 (ESRI, Redlands, CA, USA) by using shape files available from ESRI (national boundary of Tanzania, water bodies of Africa, main rivers of Africa). The shape files of the conservation areas of Tanzania were provided by the Tanzania National Park Authority and are available free from http://www.arcgis.com/home/item.html?id = 9b06fe723ad14991b30b1b85953224c1. Prevalence circles were generated using Excel version 15.38 (Microsoft, Redmond, WA, USA).

Denmark) and appropriate projectiles. Immobilized NHPs were continuously observed for vital parameters such as respirations, pulse frequency, and internal body temperature. We monitored pulse frequency and blood oxygen saturation using a Nellcor OxiMax N65 Pulse Oximeter (Tyco Healthcare Deutschland GmbH, Neustadt, Germany). Anesthetized animals underwent a standardized health check with special focus on skin lesions. We collected whole blood from the femoral vein using an S-Monovette closed blood collection system (Sarstedt, Nümbrecht, Germany) mounted with a 20G needle. We collected two 9-mL serum tubes under aseptic conditions. We then centrifuged serum tubes at 55,000 relative centrifugation force for 15 min, transferred serum into cryovials, and stored the vials in liquid nitrogen. In animals with skin lesions, we took a 6-mm biopsy from the skin ulcer using a sterile dermal biopsy punch. From each animal (and ulcer), we preserved tissue samples in lysis buffer (10 mmol/L Tris [pH 8.0], 0.1 EDTA, and 0.5% sodium dodecyl sulfate).

We treated animal wounds with Silverspray (Silver Aluminum Aerosol; Henry Schein, Langen, Germany) and allowed animals to recover under close supervision. Samples were temporarily stored at –80°C at the Tanzania Wildlife Research Institute headquarters (Arusha, Tanzania). Aliquots were exported to the German Primate Center (Göttingen, Germany) for further analysis and additional confirmation.

Serologic Testing

We used a commercially available treponemal test (ESPLINE TP; Fujirebio Diagnostics, Hannover, Germany) to check all serum samples for *T. pallidum* antibodies. The assay has been validated for use in baboons (*12*), where it performed with 97.7% (95% CI 87.7%–99.9%) sensitivity and 96.0% (95% CI 79.7%–99.9%) specificity. We tested serum samples on the day of sampling and operated and interpreted test cassettes according to the manufacturer's guidance.

DNA Extraction from Skin Tissue

We extracted DNA following the standard protocol of the QIAmp DNA Mini Kit (QIAGEN, Hilden, Germany), with some minor modifications. In brief, we cut ≈25 mg tissue into small pieces and incubated in 180 μL lysis buffer, in which the sample had been stored since collection. After adding 20 μL proteinase K, samples were digested overnight at 56°C and 900 rpm (Thermomixer Comfort; Eppendorf, Hamburg, Germany). We added an additional washing step using 300 μL AW1 buffer and eluted the DNA twice with 100 μL AE buffer. We further purified extracted DNA using glycogen precipitation according to the protocol published in Knauf et al. (*13*).

TP_0105 (*polA*) Amplification and Sequencing

We performed PCR targeting the polymerase 1 gene (*TP_0105*, *polA*) of *T. pallidum* by using primers designed by Liu et al. (*14*). This assay has a reported sensitivity of 95.8% and a specificity of 95.7% and has been demonstrated to segregate pathogenic *T. pallidum* subspecies from nonpathogenic treponemes, other spirochetes, and 59 species of bacteria and viruses including those causing genital ulcers in humans (*14*). The 50-μL reaction volume comprised 25 μL 2 × Universe High-Fidelity Hot Start DNA Polymerase Master Mix (Biotool, Munich, Germany), 17 μL RNAase free water, 2 μL of each 10 μmol/L primer, 1 μL DNA polymerase (1 U/μL), 1 μL of 10 mmol/L each dNTP, and 2 μL template DNA, independent of DNA concentration. We conducted amplification in a SensoQuest Labcycler using the following thermocycling conditions: predenaturation at 95°C for 3 min, followed by 50 cycles each with 95°C for 15 s, 60°C for 15 s, and 72°C for 30 s. The profile was completed with a postextension step at 72°C for 5 min and indefinite cooling of the PCR product at 8°C. All *polA* PCR products were run on a 1% agarose gel to check for PCR performance and correct amplicon size. We gel extracted a representative subset of the PCR products (n = 19), purified with the QIAGEN Gel Extraction Kit (QIAGEN), and Sanger sequenced using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and the amplification primers. Sequencing was performed by Seqlab Sequence Laboratories (Microsynth, Göttingen, Germany).

TP_0574 (*tp47*) Quantitative PCR

We performed TaqMan real-time PCR targeting a 132-bp fragment of the *TP_0574* gene. Primers and probe used were published elsewhere (*15*). The reaction encompassed 10 μL TaqMan Universal MasterMix II (no Uracil-N glycosylase; Applied Biosystems) and 1.8 μL of each 10 μmol/L primer and the probe. Total genomic DNA concentration added to each reaction was normalized to 100 ng. Molecular-grade water was used to adjust the reaction volume to 20 μL. Cycling conditions were as follows: 50°C for 2 min, 95°C for 10 min, followed by 50 cycles each at 95°C for 15 s and 60°C for 60 s. Reactions were run on a StepOnePlus Real-Time PCR System (Applied Biosystems). We measured all samples as triplicates and analyzed data using StepOne version 2.3 software (Applied Biosystems).

TP_0619 Amplification and Sequencing

We performed PCR targeting the *TP_0619* gene of *T. pallidum* to distinguish infection with TPE or *T. pallidum* subsp. *endemicum* (TEN) strains from infection with *T. pallidum* subsp. *pallidum* (TPA) strains. At this locus, TPA differs from TPE and TEN in ≥73 positions (online Technical

Appendix 1 Figure). We used primers 5′-TTACCCAGACATTTTTCTCCACATA-3′ and 5′-TACAAGCTCCCACAATGCCA-3′ to amplify a 608-bp fragment. The PCR conditions and working steps were identical to the PCR targeting the *polA* gene, except that the annealing temperature was adjusted to 55°C.

Data Analysis

We performed statistical analyses using GraphPad Prism version 7.0c (GraphPad Software, La Jolla, CA, USA), and R version 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria). We compared variables such as the presence of *T. pallidum* antibodies and clinical manifestations per species by using $2 \times 2 \times n$ contingency tables and a 2-tailed Fisher exact test. We used a χ^2 test to compare the outcome of >2 sampling sites using $n \times 2$ contingency tables. Proportions were tested at a critical probability of 0.05 and 95% CI. We considered $p \leq 0.05$ as statistically significant.

We analyzed and edited retrieved sequence data using 4Peaks 1.8 (http://www.nucleobytes.com) and SeaView 4.5.4 software (*16*). We compared sequences with respective orthologs available in GenBank using a standard nucleotide BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Results

NHP Species

We sampled 289 NHPs (Table) and confirmed previously reported *T. pallidum* infection in olive baboons at GNP (*5*,*17*,*18*), SNP (*3*), NCA (*3*), and LMNP (*3*,*6*,*7*). In addition, we report *T. pallidum* infection in yellow baboons, vervet monkeys, and blue monkeys in different regions of Tanzania (Table; Figure 1; online Technical Appendix 2, https://wwwnc.cdc.gov/EID/article/24/6/18-0037-Techapp2.xlsx).

The overall mean seropositivity of *T. pallidum* infection in the NHP samples was 53.3% (154/289). More female (82/135 [60.7%]) than male (72/154 [46.8%]) NHPs had *T. pallidum* antibodies. Overall, 35/45 (77.8%) vervet monkeys, 85/137 (62.0%) olive baboons, 33/75 (44.0%) yellow baboons and 1/15 (6.7%) blue monkeys had antibodies against the bacterium. Most (94 [61.0%]) of the 154 seropositive NHPs appeared healthy without any clinical skin lesions. The association between *T. pallidum* antibodies and skin ulceration was tested using 2-tailed Fisher exact test and was significant in olive baboons (n = 137; odds ratio [OR] 15.95 [95% CI 4.7–51.1]; p<0.0001) and yellow baboons (n = 75; OR 11.04 [95% CI 1.7–126.8]; p = 0.0185), but not in vervet (n = 45; OR ∞ [95% CI 0.0–1.0]; p = 0.0888) and blue monkeys (n = 15; OR 0.00 [95% CI 0.0–126.0]; p>0.9999 [dataset is provided in online Technical Appendix 2]). No *T. pallidum* antibodies were detected in the 10 Zanzibar red colobus, 3 Udzungwa red colobus, 2 Ugandan red colobus, and 2 red-tailed monkeys sampled. Moreover, none of these 4 species showed any kind of skin ulceration (Table).

Clinical Manifestations

Among the 156 *T. pallidum*–seropositive and/or PCR-positive NHPs (including 2 serologically negative but PCR-positive animals) and across the different sampling sites, we found anogenital ulcers associated with the infection (Figure 2, panel A) in 59.8% ± 23.9% of the yellow baboons (mean ± SEM, 6 investigated sites; data were analyzed as fraction of *T. pallidum*–infected animals with anogenital lesions per sampling site); 45.6% ± 16.2% of the olive baboons (mean ± SEM, 6 investigated sites); and 31.6% ± 9.4% of the infected vervet monkeys (mean ± SEM, 9 investigated sites). One of the 2 *T. pallidum*–infected blue monkeys showed anogenital skin ulceration; the second animal was clinically healthy. Orofacial lesions (Figure 2, panel B) were exclusively observed in

Table. Test results of *Treponema pallidum* infection in samples of free-ranging nonhuman primate species, Tanzania*

		Total/seropositive/skin lesion/PCR positive†					
		Positive			Negative		
Species	No. (%)	Total	Male	Female	Total	Male	Female
Olive baboon (*Papio anubis*)	137 (47.4)	86	34/34/12/12	52/51/31/30‡	51	29/0/1‡/NA	22/0/1‡/NA
Yellow baboon (*Papio cynocephalus*)	75 (26.0)	33	17/17/2/2	16/16/5/5	42	27/0/1§/NA	15/0/0/NA
Vervet monkey (*Chlorocebus pygerythrus*)	45 (15.6)	35	21/21/10/8‡	14/14/1/1	10	7/0/0/NA	3/0/0/NA
Blue monkey (*Cercopithecus mitis*)	15 (5.2)	2	1/0/1/1	1/1/0/0	13	8/0/0/NA	5/0/0/NA
Red-tailed monkey (*Cercopithecus ascanius*)	2 (0.7)	0	NA	NA	2	2/0/0/NA	NA
Zanzibar red colobus (*Piliocolobus kirkii*)	10 (3.5)	0	NA	NA	10	4/0/0/NA	6/0/0/NA
Udzungwa red colobus (*Piliocolobus gordonorum*)	3 (1.0)	0	NA	NA	3	2/0/0/NA	1/0/0/NA
Ugandan red colobus (*Piliocolobus tephrosceles*)	2 (0.7)	0	0/0/0/0	0/0/0/0	2	2/0/0/NA	NA
Total	289 (100.0)	156	73	83	133	81	52

*Results are based on the consensus of detected *T. pallidum* antibodies (ESPLINE TP) and PCR results of 3 independent gene targets (*polA*, *tp47*, and *TP_0619*). NA, not applicable.
†PCR was conducted only on animals with skin lesions.
‡No skin sample was available for some positive animals.
§Skin lesion at the genitalia most likely from fight; no tissue sample available.

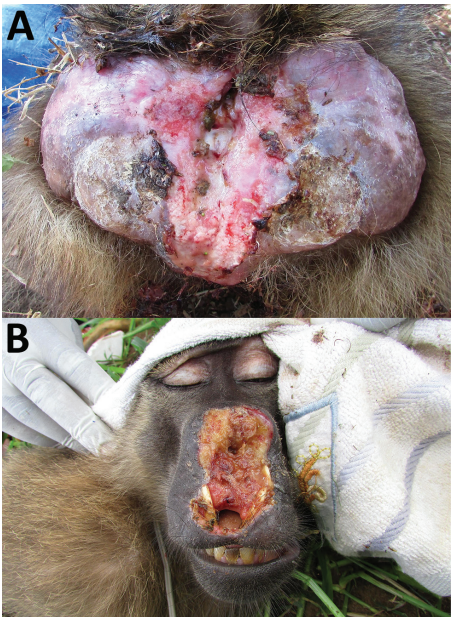


Figure 2. *Treponema pallidum*–induced clinical manifestations affecting olive baboons (*Papio anubis*), Tanzania. A). Lesions on the anogenital area of animal at Lake Manyara National Park. B) Facial lesions of animal at Tarangire National Park. Orofacial lesions were found only in olive baboonsA color version of this figure is available online (http://wwwnc.cdc.gov/EID/article/24/6/18-0037-F2.htm).

olive baboons at SNP, TNP, and LMNP, of which 2 olive baboons at TNP and 1 at SNP were included in our study. These animals represent 3.5% of the 86 *T. pallidum*–seropositive and/or PCR-positive sampled olive baboons. One animal from TNP had concurrent orofacial and anogenital skin ulcerations. We also observed these ulcerations in olive baboons at LMNP, although capture and sampling of these animals was not possible.

Geographic Distribution

Our results provide evidence for *T. pallidum* infection in NHPs at 11 of the 14 sites investigated (Figure 1; online Technical Appendix 2). The only sites where infection was not detected were ANP (14 NHPs), SGR (9 NHPs), and JCBNP (13 NHPs). We found *T. pallidum*–positive vervet monkeys in all areas where the species was examined (GNP, KNP, LMNP, MKNP, MNP, RNP, SNP, TNP, UNP) except for the 1 animal from Zanzibar (JCBNP). One PCR-positive and anogenital ulcerated blue monkey from LMNP had reproducibly negative serologic results. Because sampling was biased toward animals with skin lesions, we more objectively compared field sites by analyzing data from animals that appeared to be clinically unaffected. Healthy-looking olive baboons were significantly more often *T. pallidum*–positive at LMNP (n = 6/6) than at any other sampling area in Tanzania where the species is present (ANP [n = 0/12], GNP [n = 8/23], NCA [1/9], SNP [n = 16/25], TNP [n = 12/17]; 6 × 2 contingency table: $\chi^2 = 30.15$, df = 5; p<0.0001). Likewise, clinically unaffected yellow baboons were significantly more often *T. pallidum*–infected at MKNP (n = 16/19) than at any of the other sampling areas in Tanzania where the species is present (KNP [n = 0/6], MNP [n = 0/10], RNP [n = 8/16], SGR [n = 0/7], UNP [n = 2/17]; 6 × 2 contingency table: $\chi^2 = 38.39$, df = 5; p<0.0001). In the vervet monkeys, we found no differences among sampling sites (GNP [n = 3/3], KNP [n = 2/5], LMNP [n = 1/2], MKNP [n = 2/3], MNP [n = 1/2], RNP [n = 4/4], SNP [n = 8/8], TNP [n = 3/6], JCBNP [n = 0/1]; 9 × 2 contingency table: $\chi^2 = 12.97$, df = 8; p = 0.1130), but sample size per site was low (online Technical Appendix 2).

Molecular Characterization of *T. pallidum* Samples

In the 65 animals with skin ulcers, we confirmed *T. pallidum* by amplification of a part of the *polA* gene (classic PCR) and/or the *tp47* locus (quantitative PCR; 59/60 animals tested positive; online Technical Appendix 2). For 5 animals, we did not perform PCR because of limited quantities of samples. All obtained sequences were identical. We deposited a representative sequence of the *polA* gene from a yellow baboon (16RUF8140716) in GenBank (accession no. MF627733). Of 58 tested animals, 56 were positive in the PCR targeting the *TP_0619* locus. For 7 NHPs, no PCR was performed because of sample limitations. Again, all 41 sequences obtained were identical. We deposited a representative sequence from a vervet monkey (4KNF2121016) in GenBank (accession no. MF754122). The haplotype was identical to those derived from TPE and TEN strains but different from TPA strains in ≥73 positions (online Technical Appendix 1 Figure).

Discussion

We confirmed *T. pallidum* infection in 4 free-ranging NHP species at 11 of 14 investigated sites in Tanzania. Our data for GNP must be interpreted with caution. GNP has a history of treating infected baboons with antimicrobial drugs (*17*), which might have affected prevalence rates and clinical manifestations. The finding that clinically unaffected olive baboons at LMNP, but also many animals at SNP and TNP, were infected with the bacterium (as indicated by serology; Table) shows that clinical manifestations are not representative of the actual prevalence of the disease. This finding is consistent with reports from an earlier

investigation of olive baboons at LMNP in 2007 (*6*) and in Guinea baboons (*Papio papio*) in the Niokolo Koba National Park, Senegal (*11*). In the context of human *T. pallidum* infection, where a latent stage is a key feature of infection (*19*) and which equally features positive serology in the absence of active skin lesions (*20*), this finding could argue for a similarity of disease progression in the NHP host. However, in the absence of long-term monitoring data for infected NHPs, relapsing cases, which would indicate the latent stage, cannot be identified, and standardized laboratory infection might be needed to obtain those data.

Although reduced susceptibility for *T. pallidum* infection is possible in some of the investigated species (colobines), it is likely that infection is not yet present because of behavioral and ecologic constraints between the infected and noninfected species. At least in a recent publication, a Ugandan red colobus was described with suspected active yaws-like lesions in Uganda (*21*). Consequently, we note that our sample size for colobines and red-tailed monkeys was insufficient. As a result, a conclusive evaluation on possible *T. pallidum* infection in these species was not possible. The same applied for sites where the number of infected NHPs was critically low, for example, UNP and MNP or the negative tested areas at ANP (14 animals), SGR (9 animals), and JCBNP (13 animals), as well as the NCA crater region where all 8 olive baboons were tested negative. We found *T. pallidum*–infected vervet monkeys with and without skin ulcers in 9 of the 10 sites where the species has been investigated. This finding and the larger number and geographic extent of *T. pallidum* infection in *Chlorocebus* sp. (*4*,*11*,*22*–*24*) deserve further attention in prospective studies, especially in areas where the species is present but has not yet been tested.

All *T. pallidum*–positive NHPs in this study revealed a *TP_0619* sequence that points toward infection with either TPE or TEN strains (online Technical Appendix 1 Figure). In the context of the geographic distribution of TEN strains (dry areas in Sahelian Africa and western Asia) (*25*) as well as the information obtained from the whole-genome sequences of the Tanzanian simian strains LMNP1 and 2 (S. Knauf et al., unpub. data, https://www.biorxiv.org/content/early/2017/05/10/135491), which are considered TPE strains, we assume that TPE is the dominant, if not exclusive, *T. pallidum* subspecies infecting Tanzanian NHPs. Further clarification will be achieved when multilocus strain typing data and whole-genome sequence data of the NHP samples become available.

In humans, TPE is mainly transmitted by direct skin-to-skin contact (*26*). A possible important alternate route of infection has been discussed through the involvement of flies as a vector (*27*,*28*). Although both options are at least theoretically possible for NHPs (*13*), direct contact should be considered the most likely way of intraspecies and interspecies transmission. Such transmission is further supported by reports of the close association and interaction (play, fight, or hunt) among different NHP species (*29*–*31*). Again, multilocus strain typing and whole-genome sequence data of the strains infecting NHPs in Tanzania are likely to contribute to a better understanding of host–pathogen coevolution and will provide details of the relatedness of the *T. pallidum* subspecies that infect the different NHP taxa.

Human yaws is known to be endemic to 13 countries, but Tanzania is among the 76 countries with a known history of the disease that lack recent epidemiologic data (*2*). More precisely, the disease was reported to be endemic in humans in the western areas along Lake Tanganyika and in southern Tanzania (*32*). Extensive elimination efforts decreased the reported incidence of human yaws in Tanzania from 120,000 cases in 1927 to 52,000 in 1950 (*33*) and 71 in 1978 (https://web.gideononline.com). At the same time, the wide distribution of *T. pallidum* infection in NHP on Tanzania's mainland (*7*) and the chronic infection with locally high prevalence rates (e.g., LMNP [*6*]) suggest the pathogen has been present in the respective NHP populations for at least several decades. However, current data are insufficient to develop a conclusive biogeographic scenario about the origin and spread of the infection. The first published report on *T. pallidum* infection in NHPs in 1989 (*5*) involved olive baboons at GNP. Although this is no evidence for the origin of *T. pallidum* infection in NHPs in Tanzania, it is interesting in the context of a possible anthropozoonotic introduction of the disease. GNP is in the region that has been historically classified as an area to which human yaws in Tanzania is endemic (*33*). Furthermore, GNP is close to the Democratic Republic of the Congo, a country that still reports cases of human yaws (*34*). However, all of this is speculative, and whole-genome data are needed from NHPs and human strains from the same area to provide a deeper understanding on the origin and transmission of *T. pallidum* in NHPs in Tanzania.

In a larger context, neighboring countries currently do not report NHPs with *T. pallidum*–confirmed skin lesions, although animals from East Africa (not further classified) (*22*) and Kenya (*3*) have tested serologically positive. Because *T. pallidum* infection in NHPs in Africa is widespread (*1*), further investigations should specifically include more East Africa countries, particularly those that share their borders with Tanzania.

We showed that *T. pallidum* infection in NHPs in Tanzania is geographically widespread and present in several Old World monkey species, namely olive and yellow baboons, vervet monkeys, and blue monkeys (hypothesis A). We identified the pathogen in almost all investigated sites

covering large parts of Tanzania's mainland (hypothesis B) and showed that NHPs in Tanzania are most likely infected by TPE strains. Nevertheless, our overall sample size does not permit a conclusive statement on *T. pallidum* prevalence in NHPs at any of the sampled sites. Further studies on the spatial distribution of NHP infection with *T. pallidum* and advanced genetic characterization of simian strains are crucial for identifying NHPs as a possible reservoir for human infection (*35*). In light of the data and for a sustainable eradication of human yaws, a One Health approach in which animal and human health is investigated (*36*) is needed.

Acknowledgments

We thank all people and staff members working with government and partner institutions in Germany and Tanzania who tirelessly strived to facilitate the smooth undertaking of this study at all levels from planning and execution of fieldwork to laboratory analyses. We thank Sheila A. Lukehart and Charmie Godornes for technical support for the *TP_0619* PCR and scientific advice. We thank the Government of the United Republic of Tanzania and the Revolutionary Government of Zanzibar for permission and logistical support needed in undertaking this study through responsible ministries, institutions, and authorities. These include: Ministry for Education and Vocational Training, the Commission for Science and Technology, Ministry for Natural Resources and Tourism, Ministry for Agriculture, Natural Resources, Livestock and Fisheries (SMZ), Department of Forestry and Non-renewable Natural Resources, Tanzania Wildlife Authority, Tanzania Wildlife Research Institute, Tanzania National Parks, Ngorongoro Conservation Area Authority, Sokoine University of Agriculture, and the Department of Livestock Development. We are also grateful for the support received from the Jane Goodall Institute, National Institute for Medical Research, Ugalla Primate Project, Jozani-Chwaka Bay National Park, and Masingini Forest authorities.

The study was funded by the German Research Foundation (KN1097/3-1 and KN1097/4-1 [to S.K.], RO3055/2-1 [to C.R.], and ZI548/5-1 [to D.Z.]).

About the Author

Mr. Chuma is a veterinarian and PhD student at the Sokoine University of Agriculture in Tanzania. His research focuses on the epidemiology and characterization of nonhuman primate infection with *T. pallidum* under the aspect of One Health.

References

1. Knauf S, Liu H, Harper KN. Treponemal infection in nonhuman primates as possible reservoir for human yaws. Emerg Infect Dis. 2013;19:2058–60. http://dx.doi.org/10.3201/eid1912.130863
2. Marks M. Yaws: towards the WHO eradication target. Trans R Soc Trop Med Hyg. 2016;110:319–20. http://dx.doi.org/10.1093/trstmh/trw032
3. Harper KN, Fyumagwa RD, Hoare R, Wambura PN, Coppenhaver DH, Sapolsky RM, et al. *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS One. 2012;7:e50882. Erratum in: PLoS One. 2014;9:e92489. http://dx.doi.org/10.1371/journal.pone.0050882
4. Zobaníková M, Strouhal M, Mikalová L, Cejková D, Ambrožová L, Pospíšilová P, et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl Trop Dis. 2013;7:e2172. http://dx.doi.org/10.1371/journal.pntd.0002172
5. Wallis J, Lee DR. Primate conservation: the prevention of disease transmission. Int J Primatol. 1999;20:803–26. http://dx.doi.org/10.1023/A:1020879700286
6. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IA, Nordhoff M, et al. *Treponema* infection associated with genital ulceration in wild baboons. Vet Pathol. 2012;49:292–303. http://dx.doi.org/10.1177/0300985811402839
7. Mlengeya TDK. Distribution pattern of a sexually transmitted disease (STD) of olive baboon in Lake Manyara National Park, Tanzania [dissertation]. Moshi (Tanzania): College of African Wildlife Management (MWEKA); 2004. p. 1–40.
8. Harper KN, Knauf S. *Treponema pallidum* infection in primates: Clinical manifestations, epidemiology, and evolution of a stealthy pathogen. In: Brinkworth JF, Pechenkina K, editors. Primates, pathogens, and evolution. Vol. 138. New York: Springer; 2013. p. 189–219
9. Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. Primates Med. 1969;3:113–21.
10. Narain JP, Jain SK, Bora D, Venkatesh S. Eradicating successfully yaws from India: the strategy & global lessons. Indian J Med Res. 2015;141:608–13.
11. Knauf S, Barnett U, Maciej P, Klapproth M, Ndao I, Frischmann S, et al. High prevalence of antibodies against the bacterium *Treponema pallidum* in Senegalese Guinea baboons (*Papio papio*). PLoS One. 2015;10:e0143100. http://dx.doi.org/ 10.1371/journal.pone.0143100
12. Knauf S, Dahlmann F, Batamuzi EK, Frischmann S, Liu H. Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Negl Trop Dis. 2015;9:e0003637. Erratum in: PLoS Negl Trop Dis. 2015;9:e0003757. http://dx.doi.org/10.1371/journal.pntd.0003637
13. Knauf S, Raphael J, Mitjà O, Lejora IAV, Chuma IS, Batamuzi EK, et al. Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. EBioMedicine. 2016;11:85–90. http://dx.doi.org/10.1016/j.ebiom.2016.07.033
14. Liu H, Rodes B, Chen CY, Steiner B. New tests for syphilis: rational design of a PCR method for detection of *Treponema pallidum* in clinical specimens using unique regions of the DNA polymerase I gene. J Clin Microbiol. 2001;39:1941–6. http://dx.doi.org/10.1128/JCM.39.5.1941-1946.2001
15. Marra CM, Tantalo LC, Sahi SK, Dunaway SB, Lukehart SA. Reduced *Treponema pallidum*–specific opsonic antibody activity in HIV-infected patients with syphilis. J Infect Dis. 2016;213:1348–54. http://dx.doi.org/10.1093/infdis/jiv591
16. Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 2010;27:221–4. http://dx.doi.org/10.1093/molbev/msp259
17. Collins D, Sindimwo A, Wallis P. Reproductive disease of olive baboons (*Papio anubis*) of Gombe National Park: outbreak, time-course, and attempts to limit recurrence. In: Msuha M, Nindi S, Kakengi V, Ntalwila J, Kohi E, editors. Proceedings of the eighth TAWIRI Scientific Conference. Climate change: challenges and opportunities towards sustainable wildlife conservation and development. Arusha (Tanzania): Tanzania Wildlife Research Institute; 2011. p. 236–48.

18. Wallis J. Prevention of disease transmission in primate conservation. Ann N Y Acad Sci. 2000;916:691–3. http://dx.doi.org/10.1111/j.1749-6632.2000.tb05364.x
19. Marks M, Solomon AW, Mabey DC. Endemic treponemal diseases. Trans R Soc Trop Med Hyg. 2014;108:601–7. http://dx.doi.org/10.1093/trstmh/tru128
20. Marks M, Mitjà O, Solomon AW, Asiedu KB, Mabey DC. Yaws. Br Med Bull. 2015;113:91–100. http://dx.doi.org/10.1093/bmb/ldu037
21. Krief S, Berny P, Gumisiriza F, Gross R, Demeneix B, Fini JB, et al. Agricultural expansion as risk to endangered wildlife: pesticide exposure in wild chimpanzees and baboons displaying facial dysplasia. Sci Total Environ. 2017;598:647–56. http://dx.doi.org/10.1016/j.scitotenv.2017.04.113
22. Felsenfeld O, Wolf RH. Serological reactions with treponemal antigens in nonhuman primates and the natural history of treponematosis in man. Folia Primatol (Basel). 1971;16:294–305. http://dx.doi.org/10.1159/000155411
23. Fribourg-Blanc A, Mollaret HH, Niel G. Serologic and microscopic confirmation of treponemosis in Guinea baboons [in French]. Bull Soc Pathol Exot Filiales. 1966;59:54–9.
24. Baylet R, Thivolet J, Sepetjian M, Bert J. Seroepidemiological studies on primate treponematosis in Senegal [in French]. Bull Soc Pathol Exot Filiales. 1971;64:836–41.
25. Giacani L, Lukehart SA. The endemic treponematoses. Clin Microbiol Rev. 2014;27:89–115.
26. Richard M, Knauf S, Lawrence P, Mather AE, Munster VJ, Müller MA, et al. Factors determining human-to-human transmissibility of zoonotic pathogens via contact. Curr Opin Virol. 2017;22:7–12. http://dx.doi.org/10.1016/j.coviro.2016.11.004
27. Kumm HW, Turner TB. The transmission of yaws from man to rabbits by an insect vector, *Hippelates pallipes* (Loew). Am J Trop Med Hyg. 1936;s1–16:245–71.
28. Satchell GH, Harrison RA. Experimental observations on the possibility of transmission of yaws by wound-feeding Diptera, in western Samoa. Trans R Soc Trop Med Hyg. 1953;47:148–53. http://dx.doi.org/10.1016/0035-9203(53)90068-6
29. Morris K, Goodall J. Competition for meat between chimpanzees and baboons of the Gombe National Park. Folia Primatol (Basel). 1977;28:109–21. http://dx.doi.org/10.1159/000155801
30. Hausfater G. Predatory behavior of yellow baboons. Behaviour. 1976;56:44–68.
31. Rose M. Interspecific play between free ranging guerezas (*Colobus guereza*) and vervet monkeys (*Cercopithecus aethiops*). Primates. 1977;18:957–64. http://dx.doi.org/10.1007/BF02382946
32. Hackett CJ. Extent and nature of the yaws problem in Africa. Bull World Health Organ. 1953;8:129–82, discussion 205–10.
33. Hackett CJ. Consolidation phase of yaws control; experiences in Africa. Bull World Health Organ. 1953;8:299–343, discussion 365–8.
34. Gerstl S, Kiwila G, Dhorda M, Lonlas S, Myatt M, Ilunga BK, et al. Prevalence study of yaws in the Democratic Republic of Congo using the lot quality assurance sampling method. PLoS One. 2009;4:e6338. http://dx.doi.org/10.1371/journal.pone.0006338
35. Hallmaier-Wacker LK, Munster VJ, Knauf S. Disease reservoirs: from conceptual frameworks to applicable criteria. Emerg Microbes Infect. 2017;6:e79. http://dx.doi.org/10.1038/emi.2017.65
36. Webster JP, Gower CM, Knowles SC, Molyneux DH, Fenton A. One Health—an ecological and evolutionary framework for tackling neglected zoonotic diseases. Evol Appl. 2016;9:313–33. http://dx.doi.org/10.1111/eva.12341

Address for correspondence: Sascha Knauf, Work Group Neglected Tropical Diseases, Infection Biology, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany; email: sknauf@dpz.eu

7.3.4 Nonhuman Primates Across Sub-Saharan Africa are Infected with the Yaws Bacterium *Treponema pallidum* subsp. *pertenue*

Knauf et al. *Emerging Microbes & Infections* (2018)7:157
DOI 10.1038/s41426-018-0156-4

Emerging Microbes & Infections
www.nature.com/emi

CORRESPONDENCE Open Access

Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*

Sascha Knauf[1], Jan F. Gogarten[2,3,4], Verena J. Schuenemann[5,20], Hélène M. De Nys[2,21], Ariane Düx[2,6], Michal Strouhal[7], Lenka Mikalová[7], Kirsten I. Bos[8], Roy Armstrong[9], Emmanuel K. Batamuzi[10], Idrissa S. Chuma[11], Bernard Davoust[12], Georges Diatta[13], Robert D. Fyumagwa[14], Reuben R. Kazwala[15], Julius D. Keyyu[14], Inyasi A. V. Lejora[11], Anthony Levasseur[12], Hsi Liu[16], Michael A. Mayhew[9], Oleg Mediannikov[12], Didier Raoult[12,13], Roman M. Wittig[3,17], Christian Roos[18], Fabian H. Leendertz[2], David Šmajs[7], Kay Nieselt[19], Johannes Krause[5,8] and Sébastien Calvignac-Spencer[2,6]

Dear Editor,

The bacterium *Treponema pallidum* (*TP*) causes human syphilis (subsp. *pallidum*; *TPA*), bejel (subsp. *endemicum*; *TEN*), and yaws (subsp. *pertenue*; *TPE*)[1]. Although syphilis has reached a worldwide distribution[2], bejel and yaws have remained endemic diseases. Bejel affects individuals in dry areas of Sahelian Africa and Saudi Arabia, whereas yaws affects those living in the humid tropics[1]. Yaws is currently reported as endemic in 14 countries, and an additional 84 countries have a known history of yaws but lack recent epidemiological data[3,4]. Although this disease was subject to global eradication efforts in the mid-20th century, it later reemerged in West Africa, Southern Asia, and the Pacific region[5]. New large-scale treatment options triggered the ongoing second eradication campaign, the goal of which is to eradicate yaws globally by 2020[5].

TPE is typically considered to be a strictly human pathogen, a perception that may partially have arisen from a lack of detailed data on nonhuman primate (NHP)-infecting treponemes. Indeed, a number of African NHPs show skin ulcerations that are suggestive of treponemal infections, and antibodies against *TP* have been detected in wild NHP populations[6,7]. Although genetic studies confirmed that monkeys and great apes are infected with *TP* strains[8–10], most of these analyses only used short DNA sequences. Thus, the small number of examined polymorphic sites largely precluded assignment of these strains to a particular *TP* subspecies[9], especially considering that sporadic recombination events between subspecies have been reported[11]. The only simian strain whose whole genome has been sequenced (Fribourg-Blanc, isolated from a Guinea baboon (*Papio papio*) in 1966[7]) unambiguously clustered with human-infecting *TPE* strains[12].

A fundamental question with regard to yaws evolution, and possibly yaws eradication, is whether humans and NHPs are commonly infected with the same pathogen (*TPE*) and whether transmission between NHPs and humans occurs. To determine which pathogen causes treponematoses in NHPs across sub-Saharan Africa, we collected samples from symptomatic wild individuals belonging to three NHP species (*Cercocebus atys*, *Chlorocebus sabaeus*, and *Papio anubis*) from four independent populations in West and East Africa (Fig. 1, Supplementary Table S1, Supplementary Materials). Samples were

Correspondence: David Šmajs (dsmajs@med.muni.cz) or Johannes Krause (krause@shh.mpg.de) or Sébastien Calvignac-Spencer (calvignacs@rki.de)
[1]Work Group Neglected Tropical Diseases, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany
[2]Epidemiology of Highly Pathogenic Microorganisms, Robert Koch Institute, Seestraße 10, 13353 Berlin, Germany
Full list of author information is available at the end of the article.
These authors contributed equally: Sascha Knauf, Jan F. Gogarten, Verena J. Schuenemann and Hélène M. De Nys.
Deceased: Emmanuel K. Batamuzi.
These authors also contributed equally: David Šmajs, Kay Nieselt, Johannes Krause and Sébastien Calvignac-Spencer.

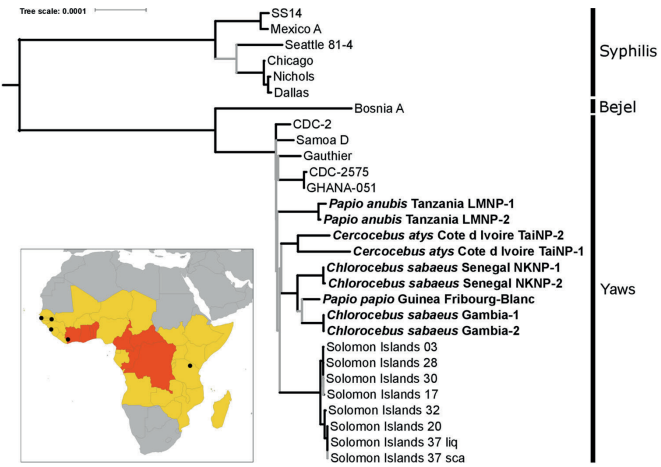


Fig. 1 Phylogenomic analysis of NHP- and human-infecting *Treponema pallidum* strains. NHP-infecting *Treponema pallidum* strains are indicated in bold. In this maximum likelihood tree, nodes with less than 95% bootstrap approximation support are indicated with gray lines. Tip labels indicate the NHP species sampled, the country of origin, and the sample ID. The scale is in nucleotide substitutions per site. The inset is a map of Africa showing the sites of origin of NHP samples from which a *TP* genome was determined (indicated with black circles). The 2013 yaws status of countries, based on the World Health Organization's Global Health Observatory (http://www.who.int/gho/en/), are indicated by color: gray indicates no previous history of yaws infections in humans, yellow indicates a country previously endemic for yaws for which the current status is unknown, and red indicates countries that are currently considered endemic for yaws

collected from NHPs at Taï National Park (TaïNP; Côte d'Ivoire), Bijilo Forest Park (BFP, the Gambia), Niokolo-Koba National Park (NKNP, Senegal), and Lake Manyara National Park (LMNP, Tanzania). Monkeys presented yaws-like orofacial and limb lesions (TaïNP and BFP) or ulcerative anogenital skin lesions (BFP, NKNP, and LMNP)[9].

Using a PCR-based assay, we demonstrated the presence of *TP* in skin lesion biopsies or swabs from NHPs inhabiting TaïNP (*C. atys*), BFP, and NKNP (*C. sabaeus*). *TP* infection in olive baboons (*P. anubis*) from LMNP had previously been confirmed[6]. Two samples per NHP population were selected for whole-genome sequencing based on a high *TP* copy number or the ability to amplify long PCR fragments (Supplementary Table S2). To overcome the presence of background host genomic DNA, we used targeted DNA capture coupled with next generation sequencing to reconstruct whole *TP* genomes[2,8]. Following quality filtering, removal of PCR duplicates, merging of different sequencing runs from the same sample, and mapping against the *TPE* strain Fribourg-Blanc reference genome, we obtained a range of 22,886–470,303 DNA sequencing reads per sample. All samples showed at least an 80% coverage of the reference genome with a depth coverage of three or higher; the average genome coverage depth was between 6.1-fold and 121.0-fold (Supplementary Table S3).

We generated maximum likelihood, Bayesian and maximum parsimony trees based on the genomes reconstructed in our study and all available reference genomes (total sequence length: 1,133,379 nucleotides). In all trees, the *TPE* and *TPA* strains formed reciprocally monophyletic groups, with a mean *TPE*/*TPA* strain divergence of 0.099%. NHP-infecting *TP* strains all clustered with human-infecting *TPE* strains (Fig. 1; Supplementary Figure S1). The *TPE* clade exhibited a star-like branching pattern with basal branches that were very short and received low statistical support. Importantly, this pattern does not support a clear reciprocal monophyly of the *TPE* strains infecting humans and NHPs. In line with this result, the minimum divergence between strains infecting humans and NHPs was lower than the maximum divergence among human or NHP-infecting strains (0.011% versus 0.015% and 0.024%). The human-infecting *TPE* strains Samoa D, CDC-2, CDC-2575, Ghana-051, and Gauthier, which span a broad geographic and temporal range (at least four decades), were less divergent from each other than the two strains infecting

sooty mangabeys from a single social group at TaïNP, which were collected in the same week (0.011% versus 0.017% sequence divergence, respectively). While intra-group strain divergence was low for the two African green monkey populations and the olive baboons (0.0003% and 0.0017%, respectively), intra-species strain divergence among African green monkeys was relatively high compared to the divergence observed between the two most divergent human strains (0.0094% versus 0.015%).

We determined the complete genome sequence and structure for the *TPE* strain from sample LMNP-1 (average depth of coverage: ×169; GenBank: CP021113; Supplementary Table S5-6)[12]. The genome structure of the LMNP-1 strain was the same as those of published complete genomes of human-infecting *TPE* strains and that of the simian strain Fribourg-Blanc. Furthermore, the genome of the LMNP-1 strain was more similar to that of the human-infecting *TPE* Gauthier strain than the simian isolate Fribourg-Blanc, showing differences at 266 and 325 chromosomal positions, respectively. Most differences were single-nucleotide substitutions or small indels (Supplementary Table S7). The LMNP-1 and Gauthier strains exhibited the same number of 24-bp repeats in the *TP_0470* gene ($n = 25$), and the Gauthier strain had only one 60-bp repeat more than the LMNP-1 strain in the *arp* gene (LMNP-1 $n = 9$ vs. Gauthier $n = 10$). All 60-bp repeats in the *arp* gene of the LMNP-1 strain were of Type II and were identical to other *TPE* strains[13]. The *tprK* gene of the LMNP-1 strain had only three variable regions, V5–V7, compared to other *TPE* strains. In addition to differences in the *TP_0433*, *TP_0470*, and *tprK* genes, relatively large indels were identified in *TPEGAU_0136* (33-nt long deletion; specific for the strains Gauthier and Samoa D), *TPFB_0548* (42-nt long deletion; specific for strain Fribourg-Blanc), and *TPEGAU_0858* (79-nt long deletion; specific for strain Gauthier), and in the intergenic regions (IGRs) between *TPEGAU_0628* and *TPEGAU_0629* (302-nt long deletion; specific for strain Gauthier) and *TPFB_0696* and *TPFB_0697* (430-nt long insertion; specific for strain Fribourg-Blanc); the lengths of the other sequence differences ranged between 1 and 15 nt. The structures of the rRNA operons in the LMNP-1 genome (coordinates 231,180–236,139; 279,584–284,533; according to *TPE* strain Gauthier: NC_016843.1) were similar to those in strains Gauthier, CDC-2, and Fribourg-Blanc, but were different than those in strains Samoa D, Samoa F, and CDC-1. The LMNP-1 16S–5S–23S region was identical in both operons, and the 23S rRNA sequences were identical to those in other *TPE* strains except for strain Fribourg-Blanc (having a single-nucleotide difference at position 458). We did not observe any mutations associated with macrolide resistance (e.g., A2058G, A2059G)[14]. When the two NHP-infecting TPE strains (Fribourg-Blanc and LMNP-1) were compared to the closest human-pathogenic TPE strains (CDC-2 and Gauthier) only 7.2 and 9.1% of all coding sequences (77 and 97 coding sequences out of 1065) contained amino acid substitutions, respectively, suggesting limited functional divergence among these strains (Supplementary Table S7-9).

Our findings unambiguously indicate that at least three African NHP species (representing four populations) from West and East Africa currently suffer from treponematosis caused by *TPE*. Taking into account the isolation of the Fribourg-Blanc strain from Guinea baboons in 1966 and its recent sequencing and identification as a member of the *TPE* clade[12], there are currently four African NHP species and five populations whose symptoms can be explained by *TPE* infections. Coupled with a growing number of clinical and serological observations[6,7,9,10], these findings suggest that infection of NHPs with *TPE* is common throughout sub-Saharan Africa. Thus, humans are not the exclusive host for the yaws bacterium, as NHPs are infected with the same bacterial agent.

TPE strains in NHPs exhibit considerable genetic diversity, which at least equals that found among published human-infecting *TPE* strains. Importantly, we found no evidence for a clear sub-differentiation of NHP-infecting and human-infecting *TPE* strains, i.e., these strains did not form well-supported reciprocally monophyletic groups. Rather, the star-like topology of our phylogenomic tree suggests a rapid initial radiation of the ancestor of *TPE*, which may have involved transmission across primate species barriers in the relatively distant past (with respect to the *TPE* clade depth). These results neither support nor allow us to exclude a possible recent transmission of *TPE* between NHPs and humans, especially due to the large geographic and temporal separation between the two groups of samples compared in this study. A major hurdle in identifying such potential transmission events is the availability of bacterial genomes. Despite large numbers of human cases, very few genomes have been determined from human-infecting *TPE* strains and only from a very limited geographic range. Generating additional human-infecting *TPE* genomes represents an important area of research, the results of which, when coupled with the genomes of the NHP-infecting *TPE* strains presented here, could enable the detection of recent zoonotic transmission events, should any exist.

Since yaws has not been reported for several decades in humans in countries where we observed NHPs to be infected with *TPE*, we expect that if transmission of *TPE* between NHPs and humans occurs, it does so at a very low frequency (as is the case for many zoonotic diseases). Of course, such a low frequency of zoonotic transmission

would not alone explain the reemergence of yaws, which is largely (or entirely) the consequence of continued human-to-human transmission. However, now that eradication of yaws appears within reach[15], the finding that *TPE* strains circulate in NHPs certainly supports the call for more research into their diversity and zoonotic potential.

Data availability
All raw sequence read files have been deposited in NCBI as part of the BioProject PRJNA343706.

Acknowledgements
This article represents a chapter in the doctoral dissertation of J.F.G. and has benefited greatly from the input of his supervisor committee and advisors Jonathan Davies, David Marcogliese, Charles Nunn, and Louis Lefebvre. Simone Lüert and Ulla Thiesen are thanked for discussions and technical support. We thank the Ivorian Ministry of Environment and Forests, the Ministry of Research, the directorship of the TaïNP, the Office Ivoirien des Parcs et Réserves, the Centre Suisse de Recherche Scientifique, the Taï Chimpanzee Project and the Taï Monkey project and their teams of field assistants for their support. We thank Jonathan Müller-Tiburtius, Therese Löhrich, Sylvain Lemoine, Simon T. Kannieu, Daniel Gnimion, Richard Peho, and Martina Magris for field assistance at TNP. For assistance and support in BFP, we thank the Department of Parks and Wildlife Management and the Forestry Department. For assistance and support with sequencing and read processing at the RKI, we are grateful to Wojciech Dabrowski, Julia Hinzmann, Andreas Nitsche, and Julia Tesch. We would like to thank the Senegalese authorities, in particular the officials of NKNP, who helped our work in Senegal immensely. We also thank Annick Abeille, who performed the initial PCR screening and sequencing of the Senegalese green monkey samples. For the work on East African simian strains, we would like to thank the TAWIRI, TANAPA, LMNP staff, especially the chief park warden, the ecology monitoring unit and park rangers, and the Tanzania Commission for Science and Technology (COSTECH). Details about the funding that supported this project are presented in the Supplementary Materials.

Author details
[1]Work Group Neglected Tropical Diseases, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany. [2]Epidemiology of Highly Pathogenic Microorganisms, Robert Koch Institute, Seestraße 10, 13353 Berlin, Germany. [3]Primatology Department, Max Planck Institute for Evolutionary Anthropology, Deutscher Pl. 6, 04103 Leipzig, Germany. [4]Department of Biology, McGill University, 1205 Avenue du Docteur-Penfield, Montréal QC H3A 1B1 Quebec, Canada. [5]Institute for Archaeological Sciences, University of Tübingen, Ruemelinstrasse 23, 72070 Tübingen, Germany. [6]Viral Evolution, Robert Koch Institute, Seestraße 10, 13353 Berlin, Germany. [7]Department of Biology, Faculty of Medicine, Masaryk University, Kamenice 5/A6, 625 00 Brno, Czech Republic. [8]Max Planck Institute for the Science of Human History, Kahlaische Strasse 10, 7745 Jena, Germany. [9]Department of Science, Natural Resources and Outdoor Studies, University of Cumbria, Fusehill Street, Carlisle, Cumbria CA12HH, UK. [10]Department of Veterinary Surgery and Theriogenology, College of Veterinary and Medical Sciences, Sokoine University of Agriculture, Morogoro, Tanzania. [11]Ecological Monitoring Department, Tanzania National Parks (TANAPA), P.O.Box 3134 Arusha, Tanzania. [12]Research Unit of Emerging Infectious and Tropical Diseases (URMITE) UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille University, 27, Bd Jean Moulin, 13385 Marseille cedex 05, Marseille, France. [13]Research Unit of Emerging Infectious and Tropical Diseases (URMITE) IRD 198, Campus IRD/UCAD, Hann Les Maristes, Dakar, Senegal. [14]Tanzania Wildlife Research Institute (TAWIRI), P.O. Box 661 Arusha, Tanzania. [15]Department of Veterinary Medicine and Public Health, College of Veterinary and Medical Sciences, Sokoine University of Agriculture, Morogoro, Tanzania. [16]National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Diseases Control and Prevention, Atlanta, GA, USA. [17]Taï Chimpanzee Project, Centre Suisse de Recherches Scientifiques, BP 1303, Abidjan 01, Abidjan, Côte d'Ivoire. [18]Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany. [19]Center for Bioinformatics, University of Tübingen, Sand 14, 72076 Tübingen, Germany. [20]Institute of Evolutionary Medicine, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. [21]Present address: Unité Mixte Internationale 233, Institut de Recherche pour le Développement, INSERM U1175, and University of Montpellier, Montpellier, France

Conflict of interest
The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (https://doi.org/10.1038/s41426-018-0156-4).

Received: 14 March 2018 Revised: 4 July 2018 Accepted: 7 August 2018
Published online: 19 September 2018

References

1. Giacani, L. & Lukehart, S. A. The endemic treponematoses. *Clin. Microbiol. Rev.* **27**, 89–115 (2014).
2. Arora, N. et al. Origin of modern syphilis and emergence of a pandemic *Treponema pallidum* cluster. *Nat. Microbiol.* **2**, 16245 (2016).
3. Marks, M. Yaws: towards the WHO eradication target. *Trans. R Soc. Trop. Med. Hyg.* **110**, 319–320 (2016).
4. World Health Organization. Eradication of yaws: procedures for verification and certification of interruption of transmission (World Health Organization, Geneva, 2018).
5. Asiedu, K., Fitzpatrick, C. & Jannin, J. Eradication of yaws: historical efforts and achieving WHO's 2020 target. *PLoS Negl. Trop. Dis.* **8**, e3016 (2014).
6. Felsenfeld, O. & Wolf, R. H. Serological reactions with treponemal antigens in nonhuman primates and the natural history of treponematosis in man. *Folia Primatol.* **16**, 294–305 (1971).
7. Fribourg-Blanc, A. & Mollaret, H. H. Natural treponematosis of the African primate. *Primates Med.* **3**, 113–121 (1969).
8. Gogarten, J. F. et al. Tools for opening new chapters in the book of *Treponema pallidum* evolutionary history. *Clin. Microbiol. Infect.* **22**, 916–921 (2016).
9. Knauf, S. et al. Treponema infection associated with genital ulceration in wild baboons. *Vet. Pathol.* **49**, 292–303 (2012).
10. Chuma, I. S. et al. Widespread *Treponema pallidum* infection in nonhuman primates, Tanzania. *Emerg. Infect. Dis.* **24**, 1002–1009 (2018).
11. Petrosova, H. et al. Whole genome sequence of *Treponema pallidum* ssp. pallidum, strain Mexico A, suggests recombination between yaws and syphilis strains. *PLoS Negl. Trop. Dis.* **6**, e1832 (2012).
12. Zobanikova, M. et al. Whole genome sequence of the Treponema Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. *PLoS Negl. Trop. Dis.* **7**, e2172 (2013).
13. Harper, K. N. et al. The sequence of the acidic repeat protein (arp) gene differentiates venereal from nonvenereal *Treponema pallidum* subspecies, and the gene has evolved under strong positive selection in the subspecies that causes syphilis. *FEMS Immunol. Med. Microbiol.* **53**, 322–332 (2008).
14. Lukehart, S. A. et al. Macrolide resistance in *Treponema pallidum* in the United States and Ireland. *N. Engl. J. Med.* **351**, 154–158 (2004).
15. Marks, M. et al. Mathematical modeling of programmatic requirements for yaws eradication. *Emerg. Infect. Dis.* **23**, 22–28 (2017).

7.3.5 Complete Genome Sequences of Two Strains of *Treponema pallidum* subsp. *pertenue* from Indonesia: Modular Structure of Several Treponemal Genes

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Indonesia: Modular structure of several treponemal genes

Michal Strouhal[1], Lenka Mikalová[1], Jan Haviernik[1], Sascha Knauf[2], Sylvia Bruisten[3], Gerda T. Noordhoek[4], Jan Oppelt[5,6], Darina Čejková[7], David Šmajs[1]*

1 Department of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic, **2** Work Group Neglected Tropical Diseases, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany, **3** Public Health Laboratory, Department of Infectious Diseases GGD Amsterdam, WT Amsterdam, the Netherlands, **4** Izore, Centrum Infectieziekten Friesland, EN Leeuwarden, the Netherlands, **5** CEITEC-Central European Institute of Technology, Masaryk University, Brno, Czech Republic, **6** National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czech Republic, **7** Department of Immunology, Veterinary Research Institute, Brno, Czech Republic

* dsmajs@med.muni.cz

OPEN ACCESS

Citation: Strouhal M, Mikalová L, Haviernik J, Knauf S, Bruisten S, Noordhoek GT, et al. (2018) Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Indonesia: Modular structure of several treponemal genes. PLoS Negl Trop Dis 12(10): e0006867. https://doi.org/10.1371/journal.pntd.0006867

Editor: Melissa J. Caimano, University of Connecticut Health Center, UNITED STATES

Received: June 19, 2018

Accepted: September 24, 2018

Published: October 10, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. The complete genome sequences from TPE Kampung Dalan K363 and TPE Sei Geringging K403 are available from the GenBank database (accession numbers CP024088.1 and CP024089.1, respectively).

Funding: This work was supported by the Grant Agency of the Czech Republic (GA17-25455S; gacr.cz) to DS and (GJ17-25589Y; gacr.cz) to MS.

Abstract

Background

Treponema pallidum subsp. *pertenue* (TPE) is the causative agent of yaws, a multistage disease endemic in tropical regions in Africa, Asia, Oceania, and South America. To date, seven TPE strains have been completely sequenced and analyzed including five TPE strains of human origin (CDC-2, CDC 2575, Gauthier, Ghana-051, and Samoa D) and two TPE strains isolated from the baboons (Fribourg-Blanc and LMNP-1). This study revealed the complete genome sequences of two TPE strains, Kampung Dalan K363 and Sei Geringging K403, isolated in 1990 from villages in the Pariaman region of Sumatra, Indonesia and compared these genome sequences with other known TPE genomes.

Methodology/principal findings

The genomes were determined using the pooled segment genome sequencing method combined with the Illumina sequencing platform resulting in an average coverage depth of 1,021x and 644x for the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes, respectively. Both Indonesian TPE strains were genetically related to each other and were more distantly related to other, previously characterized TPE strains. The modular character of several genes, including TP0136 and TP0858 gene orthologs, was identified by analysis of the corresponding sequences. To systematically detect genes potentially having a modular genetic structure, we performed a whole genome analysis-of-occurrence of direct or inverted repeats of 17 or more nucleotides in length. Besides in *tpr* genes, a frequent presence of repeats was found in the genetic regions spanning TP0126–TP0136, TP0856–TP0858, and TP0896 genes.

This work was also supported by funds from the Faculty of Medicine, Masaryk University (www.med.muni.cz), provided to junior researchers LM and MS. SK's work received partial support from the German Research Foundation (KN 1097/3-1; www.dfg.de). We acknowledge CF New Generation Sequencing Bioinformatics supported by the CIISB research infrastructure (LM2015043 funded by Ministry of Education, Youth and Sports of the Czech Republic; www.msmt.cz) for their support with obtaining scientific data presented in this paper. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the program "Projects of Large Research, Development, and Innovations Infrastructures" (LM2015042 CESNET; www.cesnet.cz), is greatly appreciated. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Conclusions/significance

Comparisons of genome sequences of TPE Kampung Dalan K363 and Sei Geringging K403 with other TPE strains revealed a modular structure of several genomic loci including the TP0136, TP0856, and TP0858 genes. Diversification of TPE genomes appears to be facilitated by intra-strain genome recombination events.

Author summary

Treponema pallidum subsp. *pertenue* (TPE) is the causative agent of yaws, a multi-stage disease that is endemic in tropical regions of Africa, Asia, Oceania, and South America. TPE belongs to the pathogenic treponemes and causes several human and animal infections. Whole genome sequences of two TPE strains isolated from patients in Indonesia were determined in this study. While both strains were highly related to other TPE strains isolated from humans and baboons, detailed genetic analyses revealed a modular character of several genes and genomic regions. While TPE genomes appear to be the most conserved genomes of uncultivable pathogenic treponemes, diversification of TPE genomes appears to be facilitated by intra-strain genome recombination events. In addition to genes with an identified modular structure, we identified additional genes that have direct or inverted repeats and thus have the potential for genetic reshuffling.

Introduction

The infectious agent of yaws, *Treponema pallidum* subsp. *pertenue* (TPE), causes chronic infections in children and young adults, which is characterized by skin lesions including nodules and ulcerations of the skin, which is later accompanied by joint, soft tissue, and bone manifestations (reviewed in [1]). Unlike the syphilis treponemes, *Treponema pallidum* subsp. *pallidum* (TPA), TPE and *Treponema pallidum* subsp. *endemicum* (TEN, the causative agent of endemic syphilis) are transmitted between individuals mostly through direct skin contact. However, possible sexual transmission has been reported for TEN, a treponeme highly related to TPE [2–4].

Only a limited number of TPE strains/isolates have been characterized to date, mainly as a result of the uncultivable character of TPE, low number of available laboratory strains, and a limited number of clinical isolates with sufficient numbers of treponemal DNA copies per sample. However, the recent study by Edmonson et al. [5] showed a successful long-time *in vitro* cultivation of syphilis treponemes that could be also potentially applied to TPE strains. This could result in an increase in the number of characterized TPE strains. So far, seven TPE strains have been completely sequenced, including five strains of human origin (CDC-2, CDC 2575, Gauthier, Ghana-051, and Samoa D) [6,7] and two TPE strains (Fribourg-Blanc and LMNP-1) isolated from a Guinea baboon (*Papio papio*) in West Africa [8] and an olive baboon (*Papio anubis*) from Tanzania [9], respectively. In addition to these complete genomes, genomes of 6 other TPE isolates of human origin [10] and 7 from nonhuman primates have been sequenced to draft genome quality [9]. TPE strains have been shown to be highly similar to syphilis-causing strains of *T. pallidum* subsp. *pallidum* (TPA) [6] and to the TEN strain Bosnia A [11].

While there is an increasing understanding of genome structure and plasticity in TPA and TEN [12,13], the genome characteristics of TPE remain largely unexplored. As a result, little is

known about intra-strain recombinations that occur in TPE strains and their role in genome evolution and diversification.

In this communication, we compared the complete genome sequences of two strains of TPE isolated in Indonesia to other available TPE whole genome sequences and identified regions resulting from intra-strain genome recombinations. While TPE genomes appear to be relatively conserved compared to the genomes of other uncultivable pathogenic treponemes, including TPA and TEN strains, genetic diversification of TPE genomes appears to be facilitated by intra-strain genome rearrangements.

Material and methods

Ethics statement

TPE strains Kampung Dalan K363 and Sei Geringging K403 originated from the study of Noordhoek *et al.* [14], where involved persons or parents of involved children gave informed consent for sample collection. No vertebrate animals were used in the study.

Strains used in this study

Two TPE strains, Kampung Dalan K363 and Sei Geringging K403, were used in this study. TPE Kampung Dalan K363 was isolated on January 5, 1990 and TPE Sei Geringging K403 on May 14, 1990 in villages in the Pariaman region of Sumatra, Indonesia. Both TPE strains were isolated from patients having skin lesions. Skin biopsies were first homogenized in PBS and intra-dermally inoculated into the shaved inguinal areas of Syrian Golden hamsters, which were later transported to the Netherlands [14]. Hamsters that developed skin lesions were sacrificed and the inguinal lymph nodes were homogenized in PBS and inoculated into the testes of New Zealand White rabbits. Several serial passages in rabbits were performed before samples were taken for isolation of treponemal DNA. TPE strains Kampung Dalan K363 and Sei Geringging K403 were provided as DNA samples by Dr. S. Bruisten (Public Health Laboratory, Department of Infectious Diseases GGD Amsterdam) who derived the DNA samples from crude treponemal lysates which were kindly donated by Dr. G. Noordhoek who collected these samples in Indonesia and processed them in the Netherlands [14]. The determination of the number of treponemal DNA copies per μl of samples was not performed.

Amplification of TPE genomic DNA

Total DNA of both samples was first amplified using multiple displacement amplification (REPLI-g kit, QIAGEN, Valencia, CA, USA) according to the manufacturer's instructions. The amplified DNA was then diluted 50-times and used as a template for TPE whole genome amplification, which was performed with treponemal specific primers as described previously [6–8,11,15]. For amplification of individual amplicons (n = 278, S1 Table), PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Otsu, Japan) was used. PCR products were generated using touchdown PCR under the following cycling conditions: initial denaturation at 94˚C for 1 min; 8 cycles: 98˚C for 10 s, 68˚C for 15 s (annealing temperature gradually reduced by 1˚C/every cycle), and 68˚C for 6 min; 35 cycles: 98˚C for 10 s, 61˚C for 15 s, and 68˚C for 6 min (43 cycles in total); followed by final extension at 68˚C for 7 min. All overlapping PCR products were purified using a QIAquick PCR Purification Kit (QIAGEN, Valencia, CA, USA) and mixed in equimolar amounts. Each pool of PCR products (n = 4 for each of TPE strains, S1 Table) was then used for whole genome sequencing.

Whole genome sequencing and *de novo* assembly of the TPE genomes

Individual pools of both TPE samples were used for Illumina Nextera XT library preparation and subsequently sequenced on a MiSeq platform (2x300 bp) at the sequencing facility of CEITEC (Brno, Czech Republic). Resultant sequencing data were quality pre-processed using Trimmomatic (v0.32) [16] with a sliding window length of 4 bp and a Phred quality threshold value equal to 17. After pre-processing, sequencing reads shorter than 50 bp were removed. Sequencing results for individual pools are summarized in S2 Table.

The Illumina sequencing reads corresponding to individual pools of both TPE samples were handled separately and *de novo* assembled using SeqMan NGen v4.1.0 software (DNASTAR, Madison, WI, USA) using default parameters. A total of 222, 249, 158, and 265 contigs from the TPE Kampung Dalan K363 strain and 124, 92, 93, and 87 contigs from the TPE Sei Geringging K403 strain were obtained for Pools 1–4, respectively (S2 Table). The resulting contigs obtained from both strains were then separately aligned to the TPE Samoa D genome (CP002374.1 [6]) using Lasergene software (DNASTAR, Madison, WI, USA). In parallel, the Illumina sequencing reads corresponding to individual pools were mapped to the corresponding pool sequences (S2 Table) of the TPE Samoa D genome (CP002374.1 [6]) and both the *de novo* and reference-guided approaches were compared. All genome gaps and discrepancies were resolved using Sanger sequencing. Altogether, 11 and 6 genomic regions of the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains were Sanger sequenced, respectively. The consensual sequences for individual pools were then used to compile the complete genome sequences of both TPE strains.

To determine the number of repetitions within the *arp* (TP0433) gene, the repetitive sequences (between coordinates 462430–463157 in the TPE Samoa D genome [6]) were amplified and Sanger sequenced using the primers 32BrepF1 (5′-CGTTTGGTTTCCCCTTTGTC-3′) and 32BrepR1 (5′-GTGGGATGGCTGCTTCGTATG-3′) as described elsewhere [17]. Similarly, the repetitive sequences within the TP0470 gene (Samoa D coordinates 498895–499200) were amplified and sequenced using the primers TPI34F4 (5′-GTCTTGTGCACATTATTCA AG-3′) and TPI34R5 (5′-CTTCGTGCAACATCGCTACG-3′). The intra-strain variability, relative to the length of several G/C-homopolymeric tracts, was identified in both genomes and the prevailing length of G/C regions was used in the final genome sequences.

Gene identification, annotation, and classification

Genes were annotated using Geneious software (v5.6.5) [18] as described previously [11] and were tagged with TPEKDK363_ and TPESGK403_ prefixes. Locus tag numbering corresponded to tag numbering for the orthologous genes annotated in the TPE Samoa D genome (CP002374.1 [6]). The TPE Samoa D genome contains one (TPESAMD_0005a) additionally annotated gene compared to the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes. This gene is fused to TP0006 and is designated as TPEKDK363_0006 and TPESGK403_0006, respectively. Four genes (TPEKDK363_0146, TPEKDK363_0520, TPEKDK363_0812, and TPEKDK363_0856a) and eight genes (TPESGK403_0126, TPESGK403_0146, TPESGK403_0312a, TPESGK403_0435a, TPESGK403_0520, TPESGK403_0812, TPESGK403_0865, and TPESGK403_0924a) were annotated as pseudogenes in the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes, respectively. Since the *tprK* gene showed intra-strain variability, the corresponding nucleotide positions were denoted with an "N" in the complete genome sequences. For proteins with unpredicted functions, a 150 bp-gene size limit was applied.

Identification of genetic heterogeneity in the sequenced genomes

The identification of genetic heterogeneity was carried out as described by Strouhal et al. [7]. Briefly, individual Illumina reads were mapped to the final version of the genome sequence using SeqMan NGen (v4.1.0) software with default parameters and requiring at least a 93% read identity relative to the reference genome. For the determination of the frequency of each nucleotide in every single genome position, the haploid Bayesian method was used for SNP calculation using the same software. Individual reads supporting a less frequent allele located at the 3'-terminus (i.e., five or less nucleotides) were omitted. At least thirty independent reads from both directions were required. Nucleotide positions located within homopolymeric tracts (defined as a stretch of six or more identical nucleotides) were excluded from analysis. Chromosomal loci showing genetic heterogeneity within TPE genomes were defined as those containing more than 8% alternative reads in regions having a coverage depth greater than 100x. Candidate sites were then visually inspected using SeqMan NGen (v4.1.0) software; the *tprK* (TP0897) gene, showing intra-strain variability, was excluded from the analysis.

Analysis of whole genome sequences

Phylogenetic trees of the TPE strains were constructed from available whole genome sequences (S3 Table) including the Samoa D (CP002374.1 [6]), CDC-2 (CP002375.1 [6]), Gauthier (CP002376.1 [6]), Fribourg-Blanc (CP003902.1 [8]), Ghana-051 (CP020365.1 [7]), CDC 2575 (CP020366.1 [7]), and LMNP-1 (CP021113.1 [9]) strains. Moreover, 6 additional draft genomes of TPE strains isolated on Solomon Islands [10] were used to determine the phylogenetic relatedness of TPE Kampung Dalan K363 and Sei Geringging K403 strains. The genome of TEN Bosnia A strain (CP007548.1 [11]) was used as an outgroup.

Whole genome alignment was constructed using SeqMan software (DNASTAR, Madison, WI, USA) and phylogenetic trees were constructed using the Maximum Likelihood method based on Tamura-Nei model [19] and with MEGA software [20]. Since there were chromosomal regions that included: (1) the *tprD* and *tprK* genes, (2) intergenic regions within both *rrn* operons, and (3) sequences in the *arp* and in the TP0470 genes, which are recombinant or repetitive in TPE strains (S3 Table), these regions were excluded from the phylogenetic analyses.

For analysis of the modular structure of the TP0136, TP0856, and TP0858 genes, additional available treponemal whole genome sequences were used including: TPA strains Nichols (CP004010.2 [21]), SS14 (CP004011.1 [21]), DAL-1 (CP003115.1 [22], Mexico A (CP003064.1 [23]), Chicago (CP001752.1 [24]), and Sea81-4 (CP003679.1 [25]) and *T. paraluisleporidarum* ecovar Cuniculus strain Cuniculi A (CP002103.1 [26]).

Identification of sequentially unique *k*-mers

For each TPE genome sequence (n = 9; S3 Table), the number of canonical *k*-mers of length 9–33 nt were determined using Jellyfish software (v2.0.0) [27]. The number of unique *k*-mers saturated at a length of 17 nts and the 17-mers and longer *k*-mers were used for further evaluation. In order to determine their exact locations and their exact numbers, the detected *k*-mers were mapped to the TPE genomes using EMBOSS fuzznuc (v6.6.0) [28]. Subsequently, *k*-mers were divided into two groups with the first group comprised of *k*-mers with exactly the same numbers in all tested TPE genomes and the second group comprised of *k*-mers with different numbers in at least one TPE genome. Localization of *k*-mers in the annotated genes of each TPE genome was carried out using BEDTools intersect (v2.26.0) [29]. For each gene, the number and type of overlapping *k*-mers was determined using R (v3.4.1, packages rio v0.5.5, dplyr v0.7.3) [30]. Similarly, for each *k*-mer, the number and type of overlapping genes was determined and *k*-mers with more than a single localization in at least one genome were extracted.

Nucleotide sequence accession numbers

The complete genome sequences from TPE Kampung Dalan K363 and TPE Sei Geringging K403 were deposited in the GenBank under accession number CP024088.1 and CP024089.1, respectively.

Results

Whole genome sequencing of the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains and *de novo* assembly of the genomes

Both TPE strains were sequenced using the pooled segment genomic sequencing (PSGS) protocol as previously described [6–8,11,15]. Illumina sequencing resulted in 7,545,122 paired reads and 1,241,564,236 total bases, with an average coverage depth of 1,021x for the TPE Kampung Dalan K363 genome, and 3,784,916 paired reads and 784,165,636 total bases, with an average coverage depth of 644x for the TPE Sei Geringging K403 genome. A total of 222, 249, 158, and 265 contigs for each of the pools 1–4 of the TPE Kampung Dalan K363 strain and 124, 92, 93, and 87 contigs for the 4 pools of the TPE Sei Geringging K403 strain were obtained by *de novo* assembly. Detailed characteristics of Illumina sequencing and *de novo* assembly are shown in S2 Table. The genome structures of both TPE strains were similar to other previously characterized TPE strains with no major chromosomal rearrangements. The summarized genomic features of TPE Kampung Dalan K363 and TPE Sei Geringging K403 were compared to the most closely related TPE Samoa D genome (CP002374.1 [6]). Details are shown in Table 1.

Analysis of whole genome sequences

The genomes of TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains differed in the number of repetitions within the *arp* (TP0433) and TP0470 genes (S3 Table). The TPE Kampung Dalan K363 strain contained 4 and 35 repetitions in the *arp* and TP0470 genes, respectively, while the TPE Sei Geringging K403 strain contained 2 and 28 repetitions within these genes, respectively. In both strains, the same repeat motif (Type II) within the *arp* gene was identified as previously shown in other TPE strains [31]. Both TPE genomes showed the same constitution of intergenic spacer regions within the *rrn* operons, i.e., tRNA-Ile/tRNA-Ala pattern [13,32] (S3 Table), and both contained the *tprD2* allele in the *tprD* locus [33] (S3 Table). Moreover, both TPE genomes differed in the sequences of the *tprK* gene variable regions [34–37].

The whole genome sequences of the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains were analyzed with respect to the occurrence of nucleotide diversity between both strains. As a result, the genomes differed in 38 single nucleotide positions (S4 Table). In addition, both genomes differed in 18 nucleotide positions within the TP0858 gene. Moreover, there were differences in both genomes in the number of 2 nt-long (TG) and 9 nt-long (TCCTCCCCC) repetitive sequences between coordinates 390964–390969 and 1051995–1052003 (according to the TPE Samoa D genome [6]), respectively. The genome of the TPE Kampung Dalan K363 strain contained 2 and 2 of these repetitive sequences, while the TPE Sei Geringging K403 genome contained 3 and 1 of these repetitions (S4 Table), respectively. Since both TPE strains Kampung Dalan K363 and Sei Geringging K403 underwent serial passages in hamsters and rabbits after their isolation from human patients, the identified genetic differences resulted either from the cultivation experiments in animals or were already present during infection of humans.

Table 1. Basic characteristics of the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes and their comparison to the published TPE Samoa D genome.

Genome parameter	TPE Kampung Dalan K363	TPE Sei Geringging K403	TPE Samoa D[a]
GenBank Accession No.	CP024088.1	CP024089.1	CP002374.1
Genome size	1,139,764 bp	1,139,464 bp	1,139,330 bp
G+C content	52.78%	52.77%	52.80%
No. of predicted genes	1124 including 54 untranslated genes[b]	1124 including 54 untranslated genes[b]	1125 including 54 untranslated genes[b]
Sum of the intergenic region length (% of the genome length)	53,026 bp (4.65%)	53,213 bp (4.67%)	52,844 bp (4.64%)
Average/median gene length	981.4/834.0 bp	981.5/834.0 bp	980.3/831.0 bp
Average/median gene length of genes with unknown function	819.1/636.0 bp	821.1/640.5 bp	843.4/657.0 bp
No. of genes encoded on plus/minus DNA strand	599/525	599/525	600/525
No. of annotated pseudogenes	4	8	6
No. of tRNA loci	45	45	45
No. of rRNA loci	6 (2 operons)	6 (2 operons)	6 (2 operons)
No. of ncRNAs	3	3	3

[a] [6]

[b] The TPE Samoa D genome contains one (TPESAMD_0005a) additionally annotated gene compared to the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes where this gene is fused to TP0006 and is designated as TPEKDK363_0006 and TPESGK403_0006, respectively.

https://doi.org/10.1371/journal.pntd.0006867.t001

As revealed by our phylogenetic analyses, the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains clustered together and were more distantly related to other complete genomes of TPE strains of human or baboon origin (Fig 1). In contrast to the Indonesian strains used in this study, all other TPE strains originated from Africa with the exception of the TPE Samoa D strain, which was isolated in Western Samoa in 1953 (S3 Table). Nevertheless, additional phylogenetic analysis including the recently published TPE draft genome sequences from 6 individuals from Solomon Islands [10] did not show clustering of TPE Kampung Dalan K363 and TPE Sei Geringging K403 with these strains (S1 File). While all Solomon Island isolates [10] clustered together and were more closely related to TPE Samoa D, TPE Kampung Dalan K363 and TPE Sei Geringging K403 belonged to distinct cluster.

Intra-strain heterogeneity in the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes

The TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes were inspected for the presence of genetic intra-strain heterogeneity [38]. While the genome of the TPE Kampung Dalan K363 strain contained 3 intra-strain heterogeneous sites, the genome of the TPE Sei Geringging K403 strain harbored only a single such site (S5 Table). The TPE Kampung Dalan K363 strain contained heterogeneous sites in genes TP0448 (encoding uracil phosphoribosyltransferase), TP0488 (coding for methyl-accepting chemotaxis protein), and TP1032 (encoding hypothetical protein). In the TPE Sei Geringging K403 strain, a single heterogeneous site was found in the TP0363 gene, encoding chemotaxis protein CheA, which is a sensor histidine kinase. All four heterogeneous sites resulted in amino acid replacements in the corresponding proteins (S5 Table).

In both TPE genomes, intra-strain variability in the length of G/C-homopolymeric tracts was identified as previously shown in other treponemal genomes [7,39,40]. Based on the prevailing length of G/C regions in the final genome sequences, 16 out of 44 such regions were

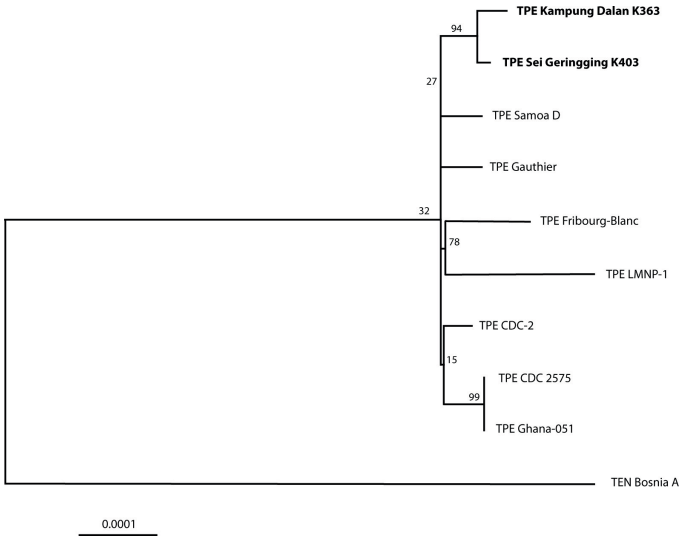


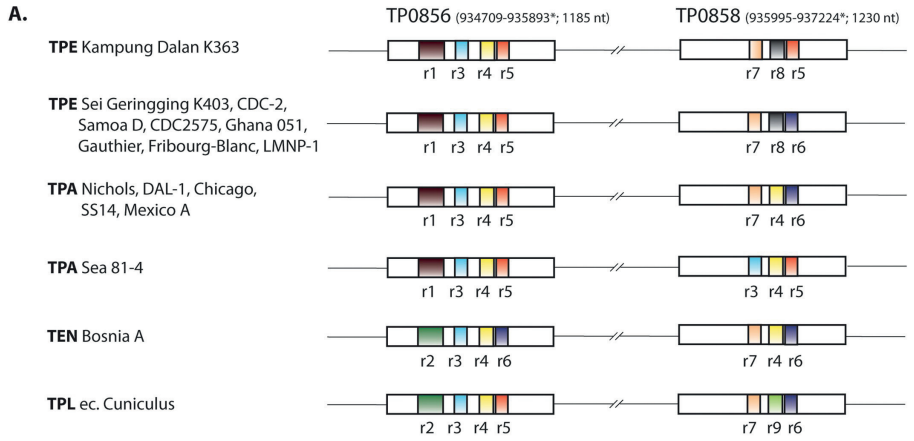
Fig 1. A phylogenetic tree constructed from the whole genome sequence alignment of available TPE and TEN complete genome sequences (S3 Table). The tree was constructed using the Maximum Likelihood method based on the Tamura-Nei model [19] and MEGA software [20]. The bar scale corresponds to a difference of 0.0001 nucleotides per site. Bootstrap values based on 1,000 replications are shown next to the branches. There were a total of 1,189 informative positions in the final dataset. Genome sequence of TEN strain Bosnia A was used as an outgroup. Both TPE strains of Indonesian origin were highly related to each other when compared to the genetic diversity detected among other, previously characterized TPE strains. Both intergenic spacer regions within the *rrn* operons, the *tprK* (TP0897) and *tprD* (TP0131) gene sequences and the repetitive sequences within the *arp* (TP0433) and TP0470 genes were omitted from the analysis. The TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains are shown in bold.

https://doi.org/10.1371/journal.pntd.0006867.g001

found to be different when comparing the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes (S6 Table). Therefore, four genes (TPESGK403_0126, TPESGK403_0312a, TPESGK403_0865, and TPESGK403_0924a) were annotated as pseudogenes in the TPE Sei Geringging K403 genome (S6 Table).

Modular structure of the TP0856 and TP0858 genes

Although both Indonesian TPE strains were highly related to each other, a relatively long stretch of nucleotide differences in the TP0858 gene sequence of both analyzed strains suggests a potential recombination event. Compared to the TP0858 sequence of the TPE Sei Geringging K403 strain (which was similar to the other TPE strains), the TP0858 sequence in the TPE Kampung Dalan K363 strain differed in 18 nucleotide positions (coordinates 819–853 in the TP0858 gene of the TPE Samoa D [6]). Moreover, the nucleotide sequence present in the TP0858 gene of the TPE Kampung Dalan K363 strain (i.e., r5 sequence; see Fig 2) was found to be identical with the one found between coordinates 798–832 in the TP0856 gene (TPE Samoa D gene coordinates). Interestingly, the same sequence (i.e., r5 sequence; see Fig 2) was detected also in the TP0858 gene of the TPA Sea 81–4 (coordinates 819–853 according to the



B.

r1	GCGTCGATGCGTATGTTTTTCCCTGAATCTGGGTTTGACTTTTCTCCCTCCACGGGGCCCGT (TPE Samoa D; TP0856, 349-410, 62 nt)
r2	AGTTCGATACGGGCGTTTCTCCCTGAGGGCTGGGTGGGCCTCAGTACCCTCGGTACGGTAATCGGACA (TPL ec. Cuniculus; TP0856, 349-416, 68 nt)
r3	AACGCACCTGTCGCTGGGTACCGACGTGGGTCTGCAA (TPA Sea 81-4; TP0856, 537-573, 37 nt; TP0858, 573-609, 37 nt)
r4	CTGCTGCAACGGCGAGAAGCCGTGCTGCAAGGACTGCGA (TPA Sea 81-4; TP0856, 750-788, 39 nt; TP0858, 765-803, 39 nt)
r5	TCCCTGCCAGGACGAAGCAACGCCGGGCTCACGGC (TPE Kampung Dalan K363; TP0856, 798-832, 35 nt; TP0858, 819-853, 35 nt)
r6	CCCCTGTCAGGACTGCAAAGACAAAGGTTCGGTGC (TEN Bosnia A; TP0856, 804-838, 35 nt; TP0858, 816-850, 35 nt)
r7	AGCACACCTCTGTCTCCAGTGACATCGGTTTGCAG (TPE Samoa D; TP0858, 578-612, 35 nt)
r8	CTGCTGCAACGGTAGTACGTGCTGCACGAAGAAGTGCGAAGG (TPE Samoa D; TP0858; 768-809; 42 nt)
r9	CTGTTGCAGCAAAGAGAAGTGCTGCACGAAGAAGTGCGAGGG (TPL ec. Cuniculus; TP0858; 768-809; 42 nt)

*according to TPE Samoa D (GenBank Acc. No. CP002374.1)

Fig 2. The modular structure of the TP0856 and TP0858 genes among completely sequenced treponemal strains. A. Please note the differences between sequence patterns of TPE Kampung Dalan K363 and the other TPE strains, the TPA Sea81-4 and the other TPA strains, and TPE and TPA/TEN strains in TP0858 gene. The r3 sequence (in TPA Sea81-4), r4 sequence (in TPA and TEN strains), r5 sequence (in Kampung Dalan K363 and TPA Sea81-4) and r6 sequence (in TEN Bosnia A) within TP0858 or TP0856 genes resulted probably from intra-strain recombination events. **B.** A list of repetitive (r3, r4, r5, r6) and non-repetitive (unique) sequences (r1, r2, r7, r8, r9).

https://doi.org/10.1371/journal.pntd.0006867.g002

TPE Samoa D TP0858 gene). Analysis of additional treponemal genomes revealed that the TEN Bosnia A strain contained identical sequences in the above described regions of both the TP0856 and TP0858 genes (i.e., r6 sequence; see Fig 2), even though these sequences in the TEN Bosnia A strain and the TPE Kampung Dalan K363 strain differed (Fig 2). Upstream of this sequence, between coordinates 768–809 in the TP0858 gene (TPE Samoa D gene coordinates), there was a 42 nt-long DNA region (i.e., r8 and r4 sequences; see Fig 2) showing identical sequences within TPE and TPA/TEN strains, respectively, but different in the TPE and TPA/TEN comparison (Fig 2). In addition, the TPA Sea81-4 strain contained identical

sequences (i.e., r3 sequence; see Fig 2) in both the TP0856 and TP0858 genes between coordinates 538–573 and 579–612 (TPE Samoa D gene coordinates), respectively. The modular structure of the TP0856 and TP0858 genes comprising all completely sequenced treponemal strains is depicted in more detail in Fig 2.

Protein sequence analyses revealed that the repetitive modules found within both TP0856 and TP0858 genes (e.g., r3 in TPA Sea81-4, r4 in TPA and TEN strains, r5 in TPE Kampung Dalan K363 and TPA Sea81-4, and r6 in TEN Bosnia A) used the same reading frame and therefore yielded the same amino acid sequence in both TP0856 and TP0858 proteins. Protein function analysis of TP0856 and TP0858 revealed presence of UPF0164, an uncharacterized protein family found only among *T. pallidum* strains. Members of this protein family belong to the membrane beta barrel superfamily. No motifs were found within these genes using the Motif search (https://www.genome.jp/tools/motif) and Pfam, NCBI-CDD and PROSITE Profile databases. However, as described recently [41], TP0856 and TP0858 proteins showed structural similarity to FadL, a long fatty acid transporter.

In addition, the sequences of TP0856 and TP0858 were analyzed by I-TASSER server [42] to predict the protein structure. The analyses revealed that most of the variable sites (i.e., r1, r2, r4, r5, r6, r8 and r9) of TP0856 and TP0858 represent coil sequences at the outer surface of β-barrels suggesting that these protein loci are exposed to the external milieu. The detailed overview of predicted structure for module sequences in TP0856 and TP0858 genes are shown in S7 Table.

Modular structure of the TP0136 gene

An alignment of both whole genome sequences of the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains revealed striking sequence differences in the TP0136 gene compared to other TPE strains. A detailed analysis of treponemal TP0136 gene orthologs identified a modular structure in the region between coordinates 158103–158250 (coordinates according to the TPE Samoa D genome [6]; see Fig 3). While in the TPE Samoa D and TPE Gauthier strains, the DNA region between coordinates 158196–158228 is represented by a 33-nt long sequence (i.e., r6 and r4 sequences; see Fig 3), in TPE strains CDC-2, CDC 2575, Fribourg-Blanc, and LMNP-1, the same region contains an additional copy of this 33-nt long sequence (Fig 3). In contrast, the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains contain within this region a duplicated segment from coordinates 158229–158250 (i.e., r5 and r2 sequences; see Fig 3), followed by another duplicated sequence from positions 158128–158195 (i.e., r3 and r4 sequences; see Fig 3). The sequence of the TP0136 gene from the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains thus resembles the TP0136 gene sequences found in TPA strains (Fig 3), where two 96 nt-long repetitions are present (i.e., comprising r1-r2-r3-r4 sequences). The modular structure of the TP0136 gene comprising all completely sequenced treponemal strains is depicted in more detail in Fig 3.

Prediction of treponemal genes with a modular structure in TPE genomes

To systematically detect genes showing a modular genetic structure, a whole genome analysis of the presence of direct or inverted repeats of 17 or more nucleotides in length was performed. The length of 17 or more nucleotides was based on an analysis of identified sequentially unique *k*-mers present in TPE genomes. Starting with *k*-mers 9 nts in length, the number of different *k*-mers increases with the length of *k*-mers until it reaches a maximum at 11 nts and then decreasing (Fig 4). In *k*-mers 17 nt in length, the number of detected different *k*-mers remains stable and therefore this length of *k*-mers was selected for identification of positions and multiplicity of these *k*-mers in the TPE strain genomes. The results of position- and

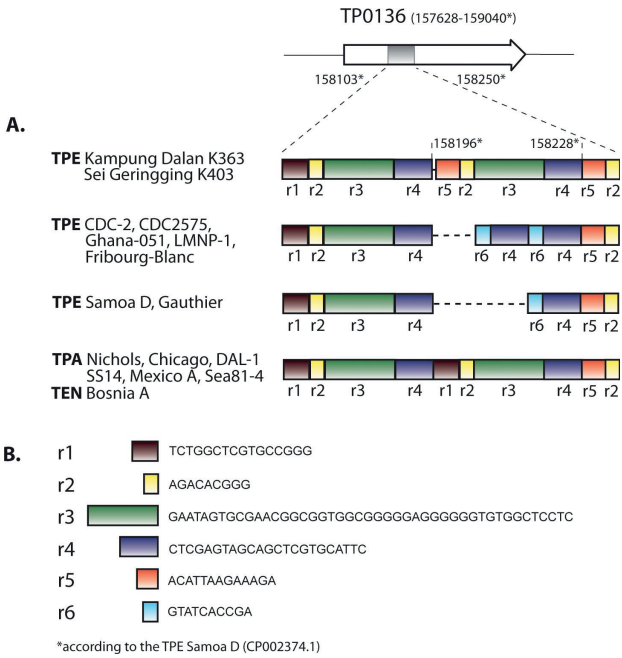


Fig 3. **A.** Modular structure of the TP0136 gene of the TPE Kampung Dalan K363, TPE Sei Geringging K403 and additional TPE, TEN, and TPA strains. While the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains resemble TPA strains, other TPE strains showed a deleted version of the TP0136 gene (Samoa D and Gauthier) or a deleted version with a duplicated subregion r6 and r4 (CDC-2, CDC 2575, Ghana-051, Fribourg-Blanc, and LMNP-1). The sequence of the TP0136 gene from the *Treponema paraluisleporidarum* ecovar Cuniculus strain Cuniculi A was not included in this analysis since this gene locus contains a genetically different sequence, which is identical (99.7% at the DNA level) to the TP0133 gene sequence. **B.** A list of repetitive sequences (r1, r2, r3, r4, r5, r6).

https://doi.org/10.1371/journal.pntd.0006867.g003

multiplicity-mapping of *k*-mers are summarized in Table 2. Besides *tpr* genes (*tprCDEFGIJK*), a frequent presence of repeats was found in the region spanning the TP0126–TP0136 genes, and in TP0856, TP0858, and TP0896 genes. Examples of other treponemal genes showing a different modular structure are presented in Fig 5.

Discussion

Two TPE strains isolated in 1990 from villages in the Pariaman region of Sumatra, Indonesia, were completely sequenced in this study using the pooled segment genomic sequencing (PSGS) approach. This approach allowed assembly and compilation of complete genome sequences without gaps or ambiguous nucleotide positions. The only exception was the variable regions within the *tprK* gene where consensus sequences were not determined due to intra-strain nucleotide sequence variability [34–37]. Both Indonesian TPE strains, i.e.,

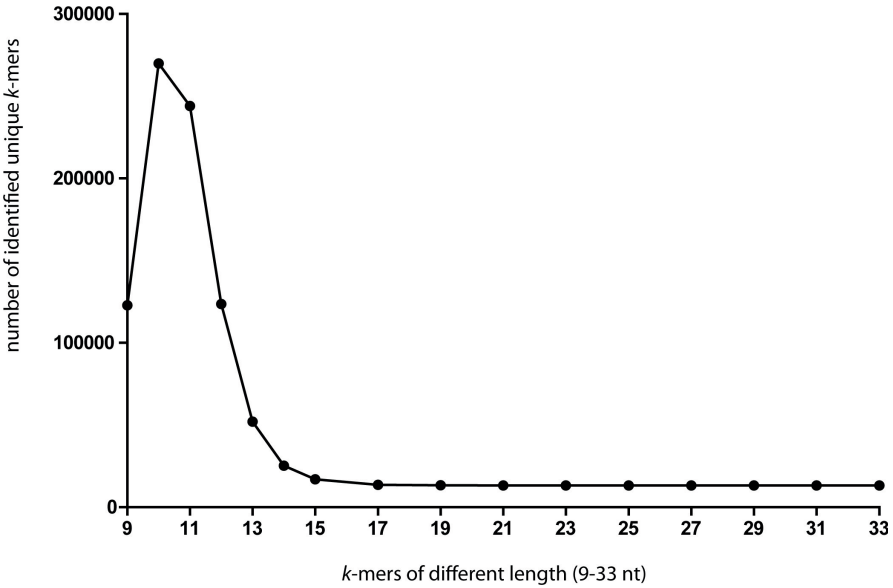


Fig 4. Number of identified *k*-mers of different length (9–33 nt) derived from TPE genome sequences. The number of unique *k*-mers saturated at a length of 17 nts, which was subsequently used as the cutoff length for further evaluations.

https://doi.org/10.1371/journal.pntd.0006867.g004

Table 2. The list of detected genes with direct or inverted repeats of 17 or more nucleotides in length.

Gene	Gene name	Identical repeat found in	Prediction of protein function
TP0117	*tprC*	*tprDFI*	Tpr protein C
TP0126abc		TP0129, TP0129ab	hypothetical protein
TP0129		TP0126abc, TP0129ab	glutamate 5-kinase
TP0130		*tprK*	repeat protein K[a]
TP0131	*tprD*	*tprCFI*	Tpr protein D
TP0136		TP0133, TP0134	fibronectin binding protein, outer membrane protein[a]
TP0136a		*tprK*	hypothetical protein
TP0313	*tprE*	*tprGJ*	Tpr protein E
TP0316	*tprF*	*tprCDI*	Tpr protein F
TP0317	*tprG*	*tprEJ*	Tpr protein G
TP0620	*tprI*	*tprCDF*	Tpr protein I
TP0621	*tprJ*	*tprGE*	Tpr protein J
TP0856		TP0858	lipoprotein[a]
TP0896		TP0126a, *tprK*	ATP synthase CF1 alpha subunit[a]

[a]protein predictions by Naqvi et al. [43]

https://doi.org/10.1371/journal.pntd.0006867.t002

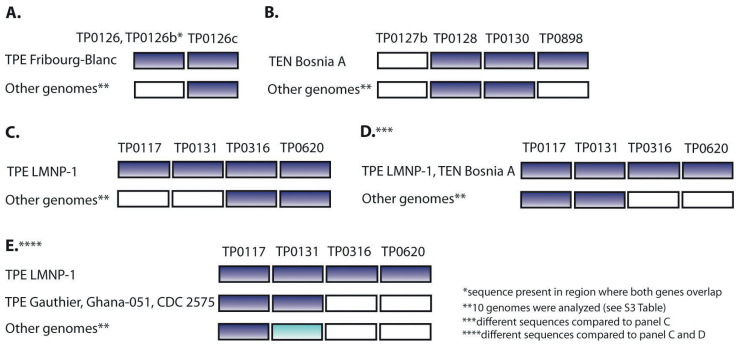


Fig 5. A schematic representation of examples of detected modular structure in several treponemal genes. A. The modular structure of TP0126/TP0126b and TP0126c genes (coordinates 148661–148691 and 149217–149247, respectively, according to the TPE Samoa D genome [6]). The TPE Fribourg-Blanc strain has identical sequences between regions where the TP0126 and TP0126b genes overlap and in TP0126c, while other analyzed genomes have different sequences in the TP0126/TP0126b. **B.** The modular structure of the TP0127b, TP0128, TP0130, and TP0898 genes (coordinates 150108–150120, 150605–150619, 151643–151659, and 977776–977789, respectively, according to the TPE Samoa D genome [6]). The TEN Bosnia A strain has identical sequences in regions of TP0128, TP0130, and TP0898 genes, while other genomes have the same sequences between regions of the TP0128 and TP0130 genes and between the TP0127b and TP0898 genes. **C.** The modular structure of the TP0117, TP0131, TP0316, and TP0620 genes (coordinates 134610–134655, 152056–152101, 331882–331927, and 672661–672706, respectively, according to the TPE Samoa D genome [6]). The TPE LMNP-1 strain has identical sequences in all of these genes while other genomes have the same sequences between regions of the TP0117 and TP0131 genes and between the TP0316 and TP0620 genes. **D.** The modular structure of the TP0117, TP0131, TP0316, and TP0620 genes (coordinates 134676–134672, 152122–152138, 331948–331964, and 672727–672743, respectively, according to the TPE Samoa D genome [6]). The TPE LMNP-1 and TEN Bosnia A strains have identical sequences in all of these genes while other genomes have the same sequences between regions of the TP0117 and TP0131 genes and between the TP0316 and TP0620 genes. **E.** The modular structure of the TP0117, TP0131, TP0316, and TP0620 genes (coordinates 134780–134804, 152226–152250, 332052–332076, and 672831–672855, respectively, according to the TPE Samoa D genome [6]). The TPE LMNP-1 strain has identical sequences in all these genes while TPE strains Gauthier, CDC 2575, and Ghana-051 have identical sequences between regions of the TP0117 and TP0131 genes and between the TP0316 and TP0620 genes. The remaining genomes have the same sequences between the TP0316 and TP0620 genes.

https://doi.org/10.1371/journal.pntd.0006867.g005

Kampung Dalan K363 and Sei Geringging K403, were genetically related to each other and both strains were more distantly related to other, previously characterized TPE strains.

Most of the available complete genome sequences of TPE strains originated in Africa except for the TPE Kampung Dalan K363 and Sei Geringging K403 strains that were isolated in Indonesia [14] and the TPE Samoa D strain that was isolated from the Samoan Islands in the central South Pacific, forming part of Polynesia [44]. This opens the question of whether TPE strains differ with respect to their geographical origin as shown by molecular typing studies of TPA [13]. A recent paper on TPE isolates sequenced from the Solomon Islands revealed 8 draft genome TPE sequences from 6 patients [10] showing that the Solomon Islands genome sequences represented a discrete TPE clade that was distinct from all previously sequenced TPE strains. Nevertheless, the phylogenetic analysis including also TPE strains from Solomon Islands [10] did not show clustering of TPE Kampung Dalan K363 and TPE Sei Geringging K403 with these strains (S1 File) despite their close geographical origin. However, the draft genome status of TPE Solomon Islands strains needs to be taken into account in the interpretation of the phylogeny shown in S1 File. Interestingly, the genetic features within the TP0858 gene of the TPE Kampung Dalan K363 strain, presented in this study, was similar to those found in all of the Solomon Islands isolates. This comprised a short sequence within the TP0858 gene that was conserved in all the Solomon Islands isolates suggesting that this

sequence is representative of isolates from the South Pacific region [10]. Moreover, this sequence is a part of the reverse primer binding site within the TP0858 gene, which was the target of a PCR assay designed by Chi and colleagues [45], which leads to false-negative PCR results on samples with this recombination [10,46].

A limited amount of genetic diversity within individual TPE strains was found in this study. Although it has been shown that the number of identified intra-strain heterogeneous sites correlates positively with the average depth of sequencing coverage, these genomes revealed just one and three such sites, although the average depth of sequencing coverage was well above 600x. Čejková et al. [38] proposed that the number of heterogeneous sites also reflects *T. pallidum* subspecies classification, where the majority of heterogeneous sites were found among TPA strains and not among TPE strains. This work appears to be consistent with this prediction same as the recently sequenced genomes of TPE strains Ghana-051 and CDC 2575, which showed a relatively limited number of heterogeneous sites (n = 13, n = 5; respectively) [7]. As shown in previous studies, all the alternative alleles identified in this study encoded non-synonymous amino acid replacements, suggesting an adaptive character for this genetic variability [38].

In general, pathogenic treponemes comprising TPA, TPE, and TEN strains or isolates, lack mobile genetic elements including pathogenicity islands, prophages, and plasmids [12,13]. It was long believed that the lack of mobile genetic elements is related to the absence of genetic recombination both within and between treponemal strains. Yet, due to recent accumulation of genetic data, treponemes appear to recombine genetic material both within and between genomes. One of the first observations describing intra-strain genetic recombinations (recombinations within genomes) came from studies on the *tprK* gene, which shows increasing variability during the course of human infection [35,36]. The underlying mechanism here is gene conversion using sequences from the flanking regions of *tprD* [34]. Later, Gray et al. [47] demonstrated that intra-genomic recombination has played a significant role in the evolution of *tpr* genes (*tprCDIGJK*), which have evolved through gene duplication and gene conversion. As an example, the occurrence of *tprD* and *tprD2*, both found within TPA clusters (Nichols-like and SS14-like) and within TPE strains [33], suggests a gene conversion mechanism in copying the *tprC* allele (that is identical with the *tprD* allele) to the *tprD* locus [33,47]. Similarly, the TP0136 locus of the *Treponema paraluisleporidarum* ecovar Cuniculus strain Cuniculi A contains an almost identical copy of the TP0133 gene sequence, suggesting the same mechanism as for *tprD*/*tprD2* allele alternation [13,26]. A similar situation was recently found in TPE samples isolated on Lihir Island, Papua New Guinea [48], where the TP0136 allele also had an intriguing sequence identity to the TP0133 gene. The authors proposed a possible interstrain recombination between treponemal species, however, intra-strain recombination by copying the TP0133 allele to the TP0136 locus appears to be more plausible. As shown in Čejková et al. [32], two rRNA (*rrn*) operons occurred in two different *rrn* spacer patterns (i.e., tRNA-Ala/tRNA-Ile and tRNA-Ile/tRNA-Ala patterns) and these variants were found independently of species/subspecies classification, time, and geographical source of the treponemal strains, suggesting the existence of reciprocal recombination in treponemes. Besides intra-genomic recombination events, traces of interstrain (intergenomic) recombination between TPA, TPE, and TEN strains have been proposed for several genetic loci [3,11,23].

Comparisons of genome sequences of the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains as well as analysis of other TPE strains revealed a modular structure for at least three gene loci including TP0136, TP0856, and TP0858, suggesting that the recombination within treponemal genomes can result in substantial changes in gene and protein sequences. Further systematic analyses revealed additional gene loci with a modular genetic structure that differ in certain strain(s) compared to others, these genes included TP0126,

TP0126b, TP0126c, TP0127b, TP0128, TP0130, TP0898, and *tprCDFI* (TP0117, TP0131, TP0316, and TP0620), indicating that this mechanism, which enables genetic diversification, is quite common in treponemal genomes. Moreover, there were additional genes identified that have direct or inverted repeats (summarized in Table 2) and thus have the potential for genetic reshuffling. The analysis of these genes revealed that these loci were limited to specific and relatively short genomic regions. In addition, these regions were often found in paralogous gene families including *tpr* genes (*tprCDEFGIJK*), the paralogous family of TP0133, TP0134, TP0136, and TP0462 genes, and the paralogous family of TP0548, TP0856, TP0858, TP0859, and TP0865 genes.

As a consequence of the inherent variability of these paralogous families, restriction fragment length polymorphism (RFLP) analysis of the *tprE* (TP0313), *tprG* (TP0317), and *tprJ* (TP0621) genes became a part of the CDC-typing scheme that determines, in addition to the *tprEGJ* RFLP pattern, a number of 60-bp tandem repeats within *arp* (TP0433) [49]. Interestingly, members of two additional paralogous families including TP0136 and TP0548 are targets of sequencing-based molecular typing [50–55].

Similar to the TP0136 protein, TP0856 and TP0858 are predicted lipoproteins [43]. Structure similarity of TP0856 and TP0858 proteins to FadL, a long chain fatty acid transporter, was recently published [41] and both proteins are members of a FadL-like family (TP0548, TP0856, TP0858, TP0859, TP0865) found in *T. pallidum*. Moreover, most of the variable sites (i.e., r1, r2, r4, r5, r6, r8 and r9) of TP0856 and TP0858 were located in loops suggesting that these protein loci are exposed to the external milieu. The TP0136 gene has been shown to have heterogeneous sequences among *T. pallidum* strains [56,57]. Moreover, the TP0136 lipoprotein was demonstrated to be exposed on the surface of the bacterial outer membrane and was shown to bind to the extracellular matrix glycoproteins fibronectin and laminin [56]. Immunization with recombinant TP0136 delayed ulceration in experimentally infected rabbits but did not prevent infection or the formation of skin lesions [56]. The NH_2-terminus of the TP0136 protein comprises a region with a modular structure overlapping the major fibronectin binding activity domain [57]. The modular structure was identified within the two 96 nt-long repetitions that are present in TPA strains and in the TEN Bosnia A strain. Interestingly, the sequence of the TP0136 gene in the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains resembled the TP0136 gene sequences found in TPA strains representing a new molecular type in the yaws MLST typing scheme [48]. Moreover, the specific insertion in TP0136 in the DAL-1 genome [22] was predicted to contain donor sequences for the *tprK* gene of *T. pallidum* [58].

In the case of TprC protein, one of the predicted antigenic epitopes on the 3D predicted structure, E3 (residues 575–583), partially overlaps with a recombinant region [59,60]. Our findings are therefore consistent with relatively frequent genetic recombinations operating at certain treponemal loci and these recombinations likely result in novel amino acid sequences exposed to the external milieu.

In summary, although TPE genomes appear to be the most conserved genomes of the uncultivable pathogenic treponemes [12,13], diversification of TPE genomes appears to be facilitated by intra-strain genome recombination events and rearrangements. Analysis of additional genomes will likely reveal more potential recombinations in the future.

Supporting information

S1 Table. PSGS (Pool Segment Genome Sequencing) approach—list of treponeme-specific primers used for the amplification of TP intervals of the TPE Kampung Dalan K363 and TPE Sei Geringging K403 strains.
(XLSX)

S2 Table. Next-generation sequencing statistics for the TPE Sei Geringging K403 and TPE Kampung Dalan K363 genomes.
(XLSX)

S3 Table. *Treponema* strains/genomes used in analyses.
(XLSX)

S4 Table. Nucleotide differences between the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes. *tprK* (TP0897) and TP0858 genes were excluded from the analysis.
(XLSX)

S5 Table. Intra-strain heterogeneity found in the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes. Minor alleles with a frequency over 8% and coverage depth over 100x are shown. The *tprK* (TP0897) gene was excluded from the analysis.
(XLSX)

S6 Table. Differences in G/C-homopolymeric tracts when comparing the TPE Kampung Dalan K363 and TPE Sei Geringging K403 genomes.
(XLSX)

S7 Table. The overview of predicted structure for module sequences in TP0856 and TP0858 genes.
(PDF)

S1 File. A tree constructed from the whole genome sequence alignment of available TPE and TEN genome sequences including also the draft genomes of TPE strains isolated on Solomon Islands [10]. The tree was constructed using the Maximum Likelihood method based on the Tamura-Nei model [19] and MEGA software [20]. The bar scale corresponds to a difference of 0.0001 nucleotides per site. Bootstrap values based on 1,000 replications are shown next to the branches. There were a total of 1,104,654 positions in the final dataset. Both TPE strains of Indonesian origin were highly related to each other when compared to the genetic diversity detected among other, previously characterized TPE strains. TPE strains from Solomon Islands [10] did not show clustering of TPE Kampung Dalan K363 and TPE Sei Geringging K403 with these strains despite their close geographical origin. Draft genome sequences from study by Marks et al. [10] are marked by prefix ERR and can be accessible from GenBank database using the corresponding name e.g. ERR1470334.
(PDF)

Acknowledgments

We thank prof. dr. Ernst Stolz from the Netherlands and dr. Jubianto Judarnarso from Indonesia who enabled the original yaws survey in Indonesia in 1988 where both TPE strains Kampung Dalan K363 and TPE Sei Geringging K403 were isolated. We thank Thomas Secrest (Secrest Editing, Ltd.) for his assistance with the English revision of the manuscript.

Author Contributions

Conceptualization: Michal Strouhal, Gerda T. Noordhoek, David Šmajs.

Data curation: Lenka Mikalová, Jan Oppelt, Darina Čejková.

Formal analysis: Michal Strouhal, Jan Haviernik, Sascha Knauf, Jan Oppelt.

Funding acquisition: Michal Strouhal, Lenka Mikalová, Sascha Knauf, David Šmajs.

Investigation: Michal Strouhal, Lenka Mikalová, Jan Haviernik.

Methodology: Michal Strouhal, Lenka Mikalová, Gerda T. Noordhoek.

Project administration: Michal Strouhal, David Šmajs.

Resources: Michal Strouhal, Sylvia Bruisten, Gerda T. Noordhoek, Jan Oppelt, Darina Čejková, David Šmajs.

Software: Jan Oppelt, Darina Čejková.

Supervision: Michal Strouhal, David Šmajs.

Validation: Michal Strouhal, Lenka Mikalová, Jan Haviernik, Sylvia Bruisten, David Šmajs.

Visualization: David Šmajs.

Writing – original draft: Michal Strouhal, Lenka Mikalová, David Šmajs.

Writing – review & editing: Michal Strouhal, Lenka Mikalová, Jan Haviernik, Sascha Knauf, Sylvia Bruisten, Gerda T. Noordhoek, Jan Oppelt, Darina Čejková, David Šmajs.

References

1. Giacani L, Lukehart SA. The endemic treponematoses. Clin Microbiol Rev. 2014; 27: 89–115. https://doi.org/10.1128/CMR.00070-13 PMID: 24396138
2. Grange PA, Mikalová L, Gaudin C, Strouhal M, Janier M, Benhaddou N, et al. *Treponema pallidum* 11qj subtype may correspond to a *Treponema pallidum* subsp. *endemicum* strain. Sex Transm Dis. 2016; 43: 517–518. https://doi.org/10.1097/OLQ.0000000000000474 PMID: 27419817
3. Mikalová L, Strouhal M, Oppelt J, Grange PA, Janier M, Benhaddou N, et al. Human *Treponema pallidum* 11q/j isolate belongs to subsp. *endemicum* but contains two loci with a sequence in TP0548 and TP0488 similar to subsp. *pertenue* and subsp. *pallidum*, respectively. PLoS Negl Trop Dis. 2017; 11: e0005434. https://doi.org/10.1371/journal.pntd.0005434 PMID: 28263990
4. Noda AA, Grillová L, Lienhard R, Blanco O, Rodríguez I, Šmajs D. Bejel in Cuba: molecular identification of *Treponema pallidum* subsp. *endemicum* in patients diagnosed with venereal syphilis. Clin Microbiol Infect. 2018 Feb 15. pii: S1198-743X(18)30154-X. https://doi.org/10.1016/j.cmi.2018.02.006 PMID: 29454847
5. Edmondson DG, Hu B, Norris SJ. Long-term in vitro culture of the syphilis spirochete *Treponema pallidum* subsp. *pallidum*. MBio 2018; 9: e01153–18. https://doi.org/10.1128/mBio.01153-18 PMID: 29946052
6. Čejková D, Zobaníková M, Chen L, Pospíšilová P, Strouhal M, Qin X, et al. Whole genome sequences of three *Treponema pallidum* ssp. *pertenue* strains: yaws and syphilis treponemes differ in less than 0.2% of the genome sequence. PLoS Negl Trop Dis. 2012; 6: e1471. https://doi.org/10.1371/journal.pntd.0001471 PMID: 22292095
7. Strouhal M, Mikalová L, Havlíčková P, Tenti P, Čejková D, Rychlík I, et al. Complete genome sequence of two strains of *Treponema pallidum* subsp. *pertenue* from Ghana, Africa: identical genome sequences in samples isolated more than 7 years apart. PLoS Negl Trop Dis. 2017; 11: e0005894. https://doi.org/10.1371/journal.pntd.0005894 PMID: 28886021
8. Zobaníková M, Strouhal M, Mikalová L, Čejková D, Ambrožová L, Pospíšilová P, et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl Trop Dis. 2013; 7: e2172. https://doi.org/10.1371/journal.pntd.0002172 PMID: 23638193
9. Knauf S, Gogarten J, Schuenemann VJ, De Nys HM, Duex A, Strouhal M, et al. African nonhuman primates are infected with the yaws bacterium Treponema pallidum subsp. pertenue; 2017. Preprint. Available from: bioRxiv. doi.org/10.1101/135491.
10. Marks M, Fookes M, Wagner J, Butcher R, Ghinai R, Sokana O, et al. Diagnostics for yaws eradication: insights from direct next generation sequencing of cutaneous strains of *Treponema pallidum*. Clin Infect Dis. 2017 Oct 16. https://doi.org/10.1093/cid/cix892 PMID: 29045605
11. Štaudová B, Strouhal M, Zobaníková M, Čejková D, Fulton LL, Chen L, et al. Whole genome sequence of the *Treponema pallidum* subsp. *endemicum* strain Bosnia A: the genome is related to yaws treponemes but contains few loci similar to syphilis treponemes. PLoS Negl Trop Dis. 2014; 8: e3261. https://doi.org/10.1371/journal.pntd.0003261 PMID: 25375929

12. Šmajs D, Norris SJ, Weinstock GM. Genetic diversity in *Treponema pallidum*: implications for pathogenesis, evolution and molecular diagnostics of syphilis and yaws. Infect Genet Evol. 2012; 12: 191–202. https://doi.org/10.1016/j.meegid.2011.12.001 PMID: 22198325

13. Šmajs D, Strouhal M, Knauf S. Genetics of human and animal uncultivable treponemal pathogens. Infect Genet Evol. 2018; 61:92–107. https://doi.org/10.1016/j.meegid.2018.03.015 PMID: 29578082

14. Noordhoek GT, Engelkens HJ, Judanarso J, van der Stek J, Aelbers GN, van der Sluis JJ, et al. Yaws in West Sumatra, Indonesia: clinical manifestations, serological findings and characterisation of new *Treponema* isolates by DNA probes. Eur J Clin Microbiol Infect Dis. 1991; 10: 12–19. PMID: 2009873

15. Weinstock GM, Šmajs D, Hardham J, Norris SJ. From microbial genome sequence to applications. Res Microbiol. 2000; 151: 151–158. PMID: 10865961

16. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

17. Mikalová L, Strouhal M, Čejková D, Zobaníková M, Pospíšilová P, Norris SJ, et al. Genome analysis of *Treponema pallidum* subsp. *pallidum* and subsp. *pertenue* strains: most of the genetic differences are localized in six regions. PLoS One. 2010; 5: e15713. https://doi.org/10.1371/journal.pone.0015713 PMID: 21209953

18. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012; 28: 1647–1649. https://doi.org/10.1093/bioinformatics/bts199 PMID: 22543367

19. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993; 10: 512–526. https://doi.org/10.1093/oxfordjournals.molbev.a040023 PMID: 8336541

20. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016; 33: 1870–1874. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

21. Pětrošová H, Pospíšilová P, Strouhal M, Čejková D, Zobaníková M, Mikalová L, et al. Resequencing of Treponema pallidum ssp. pallidum strains Nichols and SS14: correction of sequencing errors resulted in increased separation of syphilis treponeme subclusters. PLoS One. 2013; 8: e74319. https://doi.org/10.1371/journal.pone.0074319 PMID: 24058545

22. Zobaníková M, Mikolka P, Čejková D, Pospíšilová P, Chen L, Strouhal M, et al. Complete genome sequence of Treponema pallidum strain DAL-1. Stand Genomic Sci. 2012; 7: 12–21. https://doi.org/10.4056/sigs.2615838 PMID: 23449808

23. Pětrošová H, Zobaníková M, Čejková D, Mikalová L, Pospíšilová P, Strouhal M, et al. Whole genome sequence of Treponema pallidum ssp. pallidum, strain Mexico A, suggests recombination between yaws and syphilis strains. PLoS Negl Trop Dis. 2012; 6: e1832. https://doi.org/10.1371/journal.pntd.0001832 PMID: 23029591

24. Giacani L, Jeffrey BM, Molini BJ, Le HT, Lukehart SA, Centurion-Lara A, et al. Complete genome sequence and annotation of the *Treponema pallidum* subsp. *pallidum* Chicago strain. J Bacteriol. 2010; 192: 2645–2646. https://doi.org/10.1128/JB.00159-10 PMID: 20348263

25. Giacani L, Iverson-Cabral SL, King JC, Molini BJ, Lukehart SA, Centurion-Lara A. Complete genome sequence of the Treponema pallidum subsp. pallidum Sea81-4 strain. Genome Announc. 2014; 2: e00333–14. https://doi.org/10.1128/genomeA.00333-14 PMID: 24744342

26. Šmajs D, Zobaníková M, Strouhal M, Čejková D, Dugan-Rocha S, Pospíšilová P, et al. Complete genome sequence of Treponema paraluiscuniculi, strain Cuniculi A: the loss of infectivity to humans is associated with genome decay. PLoS One. 2011; 6: e20415. https://doi.org/10.1371/journal.pone.0020415 PMID: 21655244

27. Marcais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. Bioinformatics. 2011; 27: 764–770. https://doi.org/10.1093/bioinformatics/btr011 PMID: 21217122

28. Rice P, Longden I, Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. 2000; 16: 276–277. PMID: 10827456

29. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26: 841–842. https://doi.org/10.1093/bioinformatics/btq033 PMID: 20110278

30. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2017. Available from: https://www.R-project.org/

31. Harper KN, Liu H, Ocampo PS, Steiner BM, Martin A, Levert K, et al. The sequence of the acidic repeat protein (*arp*) gene differentiates venereal from nonvenereal Treponema pallidum subspecies, and the gene has evolved under strong positive selection in the subspecies that causes syphilis. FEMS

Immunol Med Microbiol. 2008; 53: 322–332. https://doi.org/10.1111/j.1574-695X.2008.00427.x PMID: 18554302

32. Čejková D, Zobaníková M, Pospíšilová P, Strouhal M, Mikalová L, Weinstock GM, et al. Structure of *rrn* operons in pathogenic non-cultivable treponemes: sequence but not genomic position of intergenic spacers correlates with classification of *Treponema pallidum* and *T. paraluiscuniculi* strains. J Med Microbiol. 2013; 62: 196–207. https://doi.org/10.1099/jmm.0.050658-0 PMID: 23082031

33. Centurion-Lara A, Giacani L, Godornes C, Molini BJ, Brinck Reid T, Lukehart SA. Fine analysis of genetic diversity of the tpr gene family among treponemal species, subspecies and strains. PLoS Negl Trop Dis. 2013; 7: e2222. https://doi.org/10.1371/journal.pntd.0002222 PMID: 23696912

34. Centurion-Lara A, LaFond RE, Hevner K, Godornes C, Molini BJ, Van Voorhis WC, et al. Gene conversion: a mechanism for generation of heterogeneity in the *tprK* gene of *Treponema pallidum* during infection. Mol Microbiol. 2004; 52: 1579–1596. https://doi.org/10.1111/j.1365-2958.2004.04086.x PMID: 15186410

35. LaFond RE, Centurion-Lara A, Godornes C, Rompalo AM, Van Voorhis WC, Lukehart SA. Sequence diversity of *Treponema pallidum* subsp. *pallidum tprK* in human syphilis lesions and rabbit-propagated isolates. J Bacteriol. 2003; 185: 6262–6268. https://doi.org/10.1128/JB.185.21.6262-6268.2003 PMID: 14563860

36. LaFond RE, Centurion-Lara A, Godornes C, Van Voorhis WC, Lukehart SA. *TprK* sequence diversity accumulates during infection of rabbits with *Treponema pallidum* subsp. *pallidum* Nichols strain. Infect Immun. 2006; 74: 1896–1906. https://doi.org/10.1128/IAI.74.3.1896-1906.2006 PMID: 16495565

37. Heymans R, Kolader ME, van der Helm JJ, Coutinho RA, Bruisten SM. *TprK* gene regions are not suitable for epidemiological syphilis typing. Eur J Clin Microbiol Infect Dis. 2009; 28: 875–878. https://doi.org/10.1007/s10096-009-0717-5 PMID: 19229562

38. Čejková D, Strouhal M, Norris SJ, Weinstock GM, Šmajs D. A retrospective study on genetic heterogeneity within Treponema strains: subpopulations are genetically distinct in a limited number of positions. PLoS Negl Trop Dis. 2015; 9: e0004110. https://doi.org/10.1371/journal.pntd.0004110 PMID: 26436423

39. Giacani L, Lukehart S, Centurion-Lara A. Length of guanosine homopolymeric repeats modulates promoter activity of subfamily II *tpr* genes of *Treponema pallidum* ssp. *pallidum*. FEMS Immunol Med Microbiol. 2007; 51: 289–301. https://doi.org/10.1111/j.1574-695X.2007.00303.x PMID: 17683506

40. Pinto M, Borges V, Antelo M, Pinheiro M, Nunes A, Azevedo J, et al. Genome-scale analysis of the non-cultivable Treponema pallidum reveals extensive within-patient genetic variation. Nat Microbiol. 2016; 2: 16190. https://doi.org/10.1038/nmicrobiol.2016.190 PMID: 27748767

41. Radolf JD, Kumar S. The *Treponema pallidum* outer membrane. Curr Top Microbiol Immunol. 2018; 415: 1–38. https://doi.org/10.1007/82_2017_44 PMID: 28849315

42. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res. 2015; 43: W174–W181. https://doi.org/10.1093/nar/gkv342 PMID: 25883148

43. Naqvi AA, Shahbaaz M, Ahmad F, Hassan MI. Identification of functional candidates amongst hypothetical proteins of *Treponema pallidum* ssp. *pallidum*. PLoS One. 2015; 10: e0124177. https://doi.org/10.1371/journal.pone.0124177 PMID: 25894582

44. Turner TB, Hollander DH. Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. Monogr Ser World Health Organ. 1957; 35: 3–266.

45. Chi KH, Danavall D, Taleo F, Pillay A, Ye T, Nachamkin E, et al. Molecular differentiation of *Treponema pallidum* subspecies in skin ulceration clinically suspected as yaws in Vanuatu using real-time multiplex PCR and serological methods. Am J Trop Med Hyg. 2015; 92: 134–138. https://doi.org/10.4269/ajtmh.14-0459 PMID: 25404075

46. Mikalová L, Šmajs D. Low-dose versus standard-dose azithromycin for treatment of yaws. Lancet Glob Health. 2018; 4: e357–e358. https://doi.org/10.1016/S2214-109X(18)30067-6

47. Gray RR, Mulligan CJ, Molini BJ, Sun ES, Giacani L, Godornes C, et al. Molecular evolution of the *tprC*, *D*, *I*, *K*, *G*, and *J* genes in the pathogenic genus *Treponema*. Mol Biol Evol. 2006; 23: 2220–2233. https://doi.org/10.1093/molbev/msl092 PMID: 16926243

48. Godornes C, Giacani L, Barry AE, Mitja O, Lukehart SA. Development of a Multilocus Sequence Typing (MLST) scheme for Treponema pallidum subsp. pertenue: application to yaws in Lihir Island, Papua New Guinea. PLoS Negl Trop Dis. 2017; 11: e60006113. https://doi.org/10.1371/journal.pntd.0006113 PMID: 29281641

49. Pillay A, Liu H, Chen CY, Holloway B, Sturm AW, Steiner B, et al. Molecular subtyping of Treponema pallidum subspecies pallidum. Sex Transm Dis. 1998; 25: 408–414. PMID: 9773432

50. Flasarová M, Šmajs D, Matějková P, Woznicová V, Heroldová-Dvoráková M, Votava M. Molecular detection and subtyping of Treponema pallidum subsp. pallidum in clinical specimens. Epidemiol Mikrobiol Imunol. 2006; 55: 105–111. PMID: 16970074

51. Woznicová V, Šmajs D, Wechsler D, Matějková P, Flasarová M. Detection of *Treponema pallidum* subsp. *pallidum* from skin lesions, serum, and cerebrospinal fluid in an infant with congenital syphilis after clindamycin treatment of the mother during pregnancy. J Clin Microbiol. 2007; 45: 659–661. https://doi.org/10.1128/JCM.02209-06 PMID: 17151205

52. Flasarová M, Pospíšilová P, Mikalová L, Vališová Z, Dastychová E, Strnadel R, et al. Sequencing-based molecular typing of Treponema pallidum strains in the Czech Republic: all identified genotypes are related to the sequence of the SS14 strain. Acta Derm Venereol. 2012; 92: 669–674. https://doi.org/10.2340/00015555-1335 PMID: 22434073

53. Grillová L, Pětrošová H, Mikalová L, Strnadel R, Dastychová E, Kuklová I, et al. Molecular typing of *Treponema pallidum* in the Czech Republic during 2011 to 2013: increased prevalence of identified genotypes and of isolates with macrolide resistance. J Clin Microbiol. 2014; 52: 3693–3700. https://doi.org/10.1128/JCM.01292-14 PMID: 25100820

54. Mikalová L, Grillová L, Osbak K, Strouhal M, Kenyon C, Crucitti T, et al. Molecular typing of syphilis-causing strains among human immunodeficiency virus-positive patients in Antwerp, Belgium. Sex Transm Dis. 2017; 44: 376–379. https://doi.org/10.1097/OLQ.0000000000000600 PMID: 28499290

55. Gallo Vaulet L, Grillová L, Mikalová L, Casco R, Rodríguez Fermepin M, Pando MA, et al. Molecular typing of Treponema pallidum isolates from Buenos Aires, Argentina: frequent Nichols-like isolates and low levels of macrolide resistance. PLoS One. 2017; 12: e0172905. https://doi.org/10.1371/journal.pone.0172905 PMID: 28235102

56. Brinkman MB, McGill MA, Pettersson J, Rogers A, Matějková P, Šmajs D, et al. A novel *Treponema pallidum* antigen, TP0136, is an outer membrane protein that binds human fibronectin. Infect Immun. 2008; 76: 1848–1857. https://doi.org/10.1128/IAI.01424-07 PMID: 18332212

57. Ke W, Molini BJ, Lukehart SA, Giacani L. Treponema pallidum subsp. pallidum TP0136 protein is heterogeneous among isolates and binds cellular and plasma fibronectin via its NH2-terminal end. PLoS Negl Trop Dis. 2015; 9: e0003662. https://doi.org/10.1371/journal.pntd.0003662 PMID: 25793702

58. Giacani L, Brandt SL, Puray-Chavez M, Brinck Reid T, Godornes C, Molini BJ, et al. Comparative investigation of the genomic regions involved in antigenic variation of the TprK antigen among treponemal species, subspecies, and strains. J Bacteriol. 2012; 194: 4208–4225. https://doi.org/10.1128/JB.00863-12 PMID: 22661689

59. Anand A, Luthra A, Dunham-Ems S, Caimano MJ, Karanian C, LeDoyt M, et al. TprC/D (Tp0117/131), a trimeric, pore-forming rare outer membrane protein of Treponema pallidum, has a bipartite domain structure. J Bacteriol. 2012; 194: 2321–2333. https://doi.org/10.1128/JB.00101-12 PMID: 22389487

60. Anand A, LeDoyt M, Karanian C, Luthra A, Koszelak-Rosenblum M, Malkowski MG, et al. Bipartite topology of *Treponema pallidum* repeat proteins C/D and I: outer membrane insertion, trimerization, and porin function require a C-terminal β-barrel domain. J Biol Chem. 2015; 290: 12313–12331. https://doi.org/10.1074/jbc.M114.629188 PMID: 25805501

7.3.6 Survey of Treponemal Infections in Free-Ranging and Captive Macaques, 1999-2012

Survey of Treponemal Infections in Free-Ranging and Captive Macaques, 1999–2012

Amy R. Klegarth, Chigozie A. Ezeonwu, Aida Rompis, Benjamin P.Y.-H. Lee, Nantiya Aggimarangsee, Mukesh Chalise, John Cortes, M. Feeroz, Barbara J. Molini, Bess C. Godornes, Michael Marks, Michael Schillaci, Gregory Engel, Sascha Knauf, Sheila A. Lukehart, Lisa Jones-Engel

Survey results showed treponemal infection among pet macaques in Southeast Asia, a region with a high prevalence of human yaws. This finding, along with studies showing treponemal infection in nonhuman primates in Africa, should encourage a One Health approach to yaws eradication and surveillance activities, possibly including monitoring of nonhuman primates in yaws-endemic regions.

Yaws, an endemic tropical disease distinguished by bone and skin lesions, is caused by infection with *Treponema pallidum* subsp. *pertenue* treponemes. Successful yaws treatment campaigns during 1950–1965 were followed by a resurgence of disease, and the World Health Organization (WHO) consequently mounted a yaws eradication campaign (*1*). Although the agent of yaws is spread among humans via direct contact, research has shown that nonhuman primates (NHPs) may serve as mammalian host reservoirs with the potential for zoonotic transmission (*2*). Successful eradication campaigns depend on there being no reservoir shielding the agent from eradication efforts; thus, the role that NHPs play in yaws among humans must be determined (*3*).

Author affiliations: University of Washington, Seattle, Washington, USA (A.R. Klegarth, C.A. Ezeonwu, B.J. Molini, B.C. Godornes, S.A. Lukehart, L. Jones-Engel); Udayana University, Denpasar, Bali, Indonesia (A. Rompis); University of Kent, Canterbury, UK (B.P.Y-H. Lee); Chiang Mai University, Chiang Mai, Thailand (N. Aggimarangsee); Tribhuvan University, Kirtipu, Nepal (M. Chalise); HM Government of Gibraltar, Gibraltar, Gibraltar (J. Cortes); Jahangirnagar University, Dhaka, Bangladesh (M. Feeroz); London School of Hygiene and Tropical Medicine, London, UK (M. Marks); University of Toronto Scarborough, Toronto, Ontario, Canada (M. Schillaci); Samuel Simmonds Memorial Hospital, Barrow, Alaska, USA (G. Engel); German Primate Center, Gottingen, Germany (S. Knauf)

DOI: http://dx.doi.org/10.3201/eid2305.161838

African Old World primates (OWPs) can be infected by *T. pallidum* and exhibit symptoms of yaws (*2*). Of note, the *Treponema* Fribourg-Blanc strain (isolated from a baboon in western Africa in 1966) exhibits remarkable genetic similarity to strains that cause yaws in humans (*4*) and in experiments, was shown capable of infecting humans (*5*). More recently, studies focusing on treponemal infections among NHPs in eastern Africa and the Republic of Congo showed that the NHP geographic range overlaps considerably with areas having a formerly high prevalence of yaws in humans (*2*).

Macaques (*Macaca* spp.), OWPs native to Asia and northern Africa, are susceptible to and have been experimentally infected with *T. pallidum* (*6*). After the initial WHO eradication efforts, yaws was believed to be largely eliminated from countries of mainland Asia, although reporting and active case detection have not been uniform throughout the region (*7*). Several island nations in Asia, however, continue to report active human yaws cases (*8*,*9*).

Macaques, the most widely distributed and numerous NHPs in the world, are sympatric with humans throughout Asia, thriving in human-altered environments and commonly kept as pets. To further characterize the role NHPs might play in the maintenance of *T. pallidum* subspecies, we screened an extensive archive of serum samples collected from free-ranging and captive macaques.

The Study

As part of a project characterizing the pathogen landscape among macaques and humans, we collected blood samples from NHPs during 1999–2012 and stored them at –80°C (*10*). We retrospectively screened samples from 734 macaques representing 13 species distributed throughout the animal's natural geographic range (Table 1). Study protocols were approved by the University of Washington Institutional Animal Care and Use Committee (no. 4233–01) and adhered to the American Society of Primatologists Principles for the Ethical Treatment of NHPs (https://www.asp.org/society/resolutions/EthicalTreatmentOfNonHumanPrimates.cfm).

We used a Macro-Vue RPR Card Test Kit (BD, Franklin Lakes, NJ, USA) to screen the 734 blood samples; 11 (1.5%) were positive (Table 2). The RPR (rapid plasma reagin) test, a lipoidal test (nontreponemal) for IgG and IgM typically associated with treponemal infection, can

Table 1. Number and species of free-ranging and captive macaques tested for treponemal infection, by location, 2000–2014*

Country, species	Year(s) sampled	Total no. sampled	No. captive	No. free-ranging
Nepal	2003			
Macaca mulatta		28	0	28
Bangladesh	2008–2012			
M. mulatta		137	14	123
Thailand	2003			
M. arctoides		2	2	0
M. assamensis		5	5	0
M. fascicularis		2	2	0
M. mulatta		9	9	0
M. nemestrina		4	4	0
Cambodia	2011			
M. fascicularis		39	0	39
M. leonina		5	0	5
M. nemestrina		1	0	1
M. spp. (hybrid)		3	0	3
Singapore	2003, 2005–2006, 2009			
M. fascicularis		76	0	76
Gibraltar	2004, 2009, 2013–2014			
M. sylvanus		124	0	124
Indonesia				
Bali	2000–2003			
M. fascicularis		157	0	157
Java	2002			
M. fascicularis		25	25	0
Sulawesi	2000–2002			
M. balantak		5	5	0
M. fascicularis		5	5	0
M. hecki		7	7	0
M. maura		9	9	0
M. nemestrina		2	2	0
M. nigra		22	14	8
M. nigrescens		11	11	0
M. ochreata		1	1	0
M. tonkeana		40	40	0
Macaca spp. (hybrid)		15	15	0
Total	2003–2014	734	170	564

*The 734 tested macaques represented 13 species. Captive category included pets, macaques used in performances, and macaques in zoos; free-ranging included wild macaques, urban macaques, and those at temples, shrines, and reserve parks.

occasionally elicit nonspecific responses. To confirm RPR-positive samples, we used ESPLINE TP (Fujirebio, Tokyo, Japan), an enzyme immunoassay for measuring reactivity to 2 recombinant *T. pallidum* antigens, Tp47 and Tp17. ESPLINE TP and RPR tests have been validated for use in OWPs (*11*). Of the 11 RPR test–positive samples, 1 was from Singapore; 2 from Bali, Indonesia; and 8 from Sulawesi, Indonesia. Six samples (all from Sulawesi) yielded confirmatory positive results on the ESPLINE TP assay. Of note, in Sulawesi, the only positive macaques were pets sampled from South Sulawesi and West Sulawesi Provinces, which make up the island's southwestern peninsula (Figure). We also used ESPLINE TP to test the 28 RPR test–negative samples from Sulawesi's southwestern peninsula; none tested positive.

At the time of sampling, the macaques underwent a physical examination, including close inspection of head, trunk, extremities, oral cavity, and genitals. We conducted a retrospective review of the data and found that none of the macaques had lesions typical of treponemal infection (*4*). Of the 734 macaques, 13, including 2 seropositive macaques from Sulawesi's southwestern peninsula, had hypopigmentation on the palms of their hands, feet, or both. Hypopigmentation is rarely seen in yaws but is a common manifestation of pinta, which is caused by infection with *T. carateum*, a close relative of *T. p. pertenue*.

Conclusions

Our findings show that pet macaques in Southeast Asia can be infected with *Treponema* spp. related to those that infect humans. The overall prevalence of infection was low in our survey, but the pocket of infection detected among pets in Sulawesi's southwestern peninsula is noteworthy. The demonstration of reactivity in the serologic tests provides unequivocal evidence that the macaques had been infected with *T. pallidum* or a highly related pathogen. We had hoped to amplify a portion of *tp0548*, a locus in the *T. pallidum* genome used for molecular typing, but no amplifiable pathogen DNA was found in the whole-blood samples that had been held in storage for >10 years. Therefore, we could not determine whether the treponemal strains from NHPs in Sulawesi resembled strains that cause human yaws.

Table 2. Treponemal infections in blood samples from free-ranging and pet macaques, by geographic location, 1999–2012*

Location	No. macaques positive/no. negative (% reactive)†	No. macaques tested/no. positive‡
Indonesia		
Bali	2/155 (1.3)	2/0
Java	0/25 (0)	NA
Sulawesi	8/109 (7.3)	8/6
Nepal	0/28 (0)	NA
Singapore	1/75 (1.3)	1/0
Bangladesh	0/137 (0)	NA
Thailand	0/22 (0)	NA
Cambodia	0/48 (0)	NA
Gibraltar	0/124 (0)	NA
Total	11/734 (1.5)	11/0

*NA, indicates that samples in the region were not tested.
†Determined by using the Macro-Vue RPR (rapid plasma reagin) test (BD, Franklin Lakes, NJ, USA).
‡Determined by using ESPLINE TP (Fujirebio Inc., Tokyo, Japan), a reagent for the detection of *Treponema pallidum* antibodies.

Sulawesi, the third largest island in the Indonesian archipelago, has a population of ≈17 million persons and 7 endemic macaque species. The seropositive samples from South Sulawesi and West Sulawesi Provinces were collected in July and August of 2000, immediately pre-

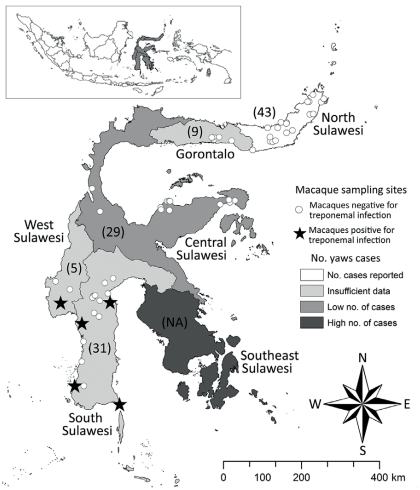


Figure. Individual sampling sites where macaques were tested for infection with *Treponema* spp. during 1999–2012 and the number of human yaws cases during 2001–2011, Sulawesi, Indonesia. Numbers in parentheses indicate number nonhuman primates sampled in each of the 6 provinces. ESPLINE TP (Fujirebio Inc., Tokyo, Japan) reagent for the detection of *T. pallidum* antibodies was used to determine whether macaque samples were positive for treponemal infection. The number of human yaws cases was determined by the World Health Organization (*1*). Inset map shows the location of Sulawesi in Indonesia (gray shading). NA, not available.

dating an active yaws outbreak among humans in the region that caused 241 documented cases in the neighboring southeastern peninsula during 2001–2011 (WHO, http://apps.who.int/iris/bitstream/10665/75528/1/WHO_HTM_NTD_IDM_2012.2_eng.pdf) (Figure). During that outbreak, WHO characterized the South Sulawesi and West Sulawesi Provinces as "data deficient" regions in regard to the status of yaws among the human population. Most macaques whose samples were used in this study were free-ranging, but all of the macaques sampled in South Sulawesi and West Sulawesi Provinces had been captured at a young age for use as pets. The association between humans and pet macaques is often intimate, with the sharing of food; space; and physical contact through grooming, play, or aggression (*12*). Two of the *Treponema* spp.–infected pets were owned by the same person and housed together. Studies of pet macaques in Sulawesi and their owners have indicated that infectious agents can move between these populations (*12*,*13*). Although the treponemal serologic status of the pet owners in this study is unavailable, the fact that seropositive pet NHPs from a region neighboring an area with a high number of human yaws cases suggests that the NHP cases resulted from treponeme transmission from humans to pets.

All macaques in this study, with the exception of *M. sylvanus* from Gibraltar, were from historically yaws-endemic areas where WHO conducted past yaws eradication campaigns. Much of Asia has a rich tradition of human–NHP commensalism, and macaques are common in villages, often as pets (*10*). Moreover, we previously showed that macaques can harbor an array of mammalian picornaviruses, astroviruses, and mycobacteria (*13*–*15*), underscoring the role of macaques in the ecology of these pathogens. However, as with our current study of treponemal infections, definitive evidence for transmission and the direction of transmission have not been established for these pathogens.

Our findings of treponemal infections among macaques in Southeast Asia, along with published work showing infection in NHPs in Africa (*4*), should encourage holistic and One-Health approaches to eradication and surveillance activities, including consideration of monitoring NHPs in yaws-endemic regions. Such approaches are particularly relevant for pet NHPs, which can easily be assessed and treated. The human–NHP interface is ancient and complex, and continued research, particularly in yaws-endemic regions, can help to ameliorate concerns as a second WHO yaws eradication campaign moves forward.

Acknowledgments

We are particularly grateful to all of the communities, temple committees, and government agencies in the areas where we

have been sampling monkeys for years: Lembaga Ilmu Pengetahuan, Indonesia; S. Chan and the staff of the Central Nature Reserve, National Parks Board, Singapore; M. Pizarro and the staff of the Gibraltar Ornithological and Natural History Society; S. Begum, M. K. Hasan, S. Akhtar, and the students and faculty of the Department of Zoology, Jahangirnagar University; and S. Son, C. Kimleng, S. Bunnary, T. Sothearos, M. Sisiket, H. Davun, and K. Pal in Cambodia and Nepal. We are also grateful to D. Cohn, A. Fuentes, J. Supriatna, R. Babo, Y. Paramastri, E. Iskandar, J. Froehlich, L. Engel, H. Engel, and L. Johnson for supporting and participating in this research.

This work was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH) (grant R01 AI078229 and R03 AI064865 to L.J.E.); National Center for Research Resources, NIH (grant P51 RR000166RR to L.J.E); US Public Health Services (grant R01 AI42143 to S.A.L.); Defense Advanced Research Projects Agency (grant N66001-02-C-8072 to L.J.E.); University of Toronto Connaught Fund (M.S.); the Chicago Zoological Society (to L.J.E.); and University of New Mexico Research Allocations Committee (L.J.E.).

Dr. Klegarth is a postdoctoral research associate in the Evolutionary Emergence of Infectious Diseases Laboratory at the University of Washington. Her research is focused on human–wildlife conflict, with an emphasis on urban nonhuman primates, and the infectious agents that are transmitted at this interface.

References

1. The World Health Organization. Eradication of yaws—the Morges strategy. Wkly Epidemiol Rec. 2012;87:189–94.
2. Knauf S, Liu H, Harper KN. Treponemal infection in nonhuman primates as possible reservoir for human yaws. Emerg Infect Dis. 2013;19:2058–60. http://dx.doi.org/10.3201/eid1912.130863
3. Marks M, Mitjà O, Vestergaard LS, Pillay A, Knauf S, Chen CY, et al. Challenges and key research questions for yaws eradication. Lancet Infect Dis. 2015;15:1220–5. http://dx.doi.org/10.1016/S1473-3099(15)00136-X
4. Zobaníková M, Strouhal M, Mikalová L, Cejková D, Ambrožová L, Pospíšilová P, et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl Trop Dis. 2013;7:e2172. http://dx.doi.org/10.1371/journal.pntd.0002172
5. Smith JL, David NJ, Indgin S, Israel CW, Levine BM, Justice J Jr, et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br J Vener Dis. 1971;47:226–51.
6. Nichols HJ. Experimental yaws in the monkey and rabbit. J Exp Med. 1910;12:616–22. http://dx.doi.org/10.1084/jem.12.5.616
7. Lahariya C, Pradhan SK. Can Southeast Asia eradicate yaws by 2010? Some lessons from the Yaws Eradication Programme of India. Natl Med J India. 2007;20:81–6.
8. Mitjà O, Marks M, Konan DJ, Ayelo G, Gonzalez-Beiras C, Boua B, et al. Global epidemiology of yaws: a systematic review. Lancet Glob Health. 2015;3:e324–31. http://dx.doi.org/10.1016/S2214-109X(15)00011-X
9. Hotez PJ, Bottazzi ME, Strych U, Chang LY, Lim YAL, Goodenow MM, et al. Neglected tropical diseases among the Association of Southeast Asian Nations (ASEAN): overview and update. PLoS Negl Trop Dis. 2015;9:e0003575. 10.1371/journal.pntd.0003575 http://dx.doi.org/10.1371/journal.pntd.0003575
10. Jones-Engel L, Steinkraus KA, Murray SM, Engel GA, Grant R, Aggimarangsee N, et al. Sensitive assays for simian foamy viruses reveal a high prevalence of infection in commensal, free-ranging Asian monkeys. J Virol. 2007;81:7330–7. 10.1128/JVI.00343-07 http://dx.doi.org/10.1128/JVI.00343-07
11. Knauf S, Dahlmann F, Batamuzi EK, Frischmann S, Liu H. Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Negl Trop Dis. 2015;9:e0003637. http://dx.doi.org/10.1371/journal.pntd.0003637
12. Jones-Engel L, Engel GA, Schillaci MA, Babo R, Froehlich J. Detection of antibodies to selected human pathogens among wild and pet macaques (*Macaca tonkeana*) in Sulawesi, Indonesia. Am J Primatol. 2001;54:171–8. http://dx.doi.org/10.1002/ajp.1021
13. Wilbur AK, Engel GA, Rompis A, A Putra IG, Lee BP, Aggimarangsee N, et al. From the mouths of monkeys: detection of *Mycobacterium tuberculosis* complex DNA from buccal swabs of synanthropic macaques. Am J Primatol. 2012;74:676–86. http://dx.doi.org/10.1002/ajp.22022
14. Oberste MS, Feeroz MM, Maher K, Nix WA, Engel GA, Hasan KM, et al. Characterizing the picornavirus landscape among synanthropic nonhuman primates in Bangladesh, 2007 to 2008. J Virol. 2013;87:558–71. http://dx.doi.org/10.1128/JVI.00837-12
15. Karlsson EA, Small CT, Freiden P, Feeroz MM, Matsen FA IV, San S, et al. Non-human primates harbor diverse mammalian and avian astroviruses including those associated with human infections. PLoS Pathog. 2015;11:e1005225. http://dx.doi.org/10.1371/journal.ppat.1005225

Address for correspondence: Lisa Jones-Engel, University of Washington, Department of Anthropology, 128 Denny Hall, 4218 Memorial Way NE, Seattle, WA 98105, USA; email: ljengel@uw.edu

7.4 Feasible Transmission Routes for *T. pallidum* Infection

A functional reservoir does not only require the maintenance of a pathogen but also feasible transmission routes (chapter 7.4) (Hallmaier-Wacker et al., 2017). This is in particular important for *T. pallidum* which is highly susceptible to drying and temperature changes (Willcox and Guthe, 1966a). The bacterium does not survive a long time outside of its host and therefore primary requires direct contact between a diseased and susceptible individual to get transmitted. However, other transmission pathways are thinkable as shown for insect transmission (Knauf et al., 2016). The following publications are themed to describe possible transmission routes and their investigation in NHPs.

7.4.1 Strain Diversity of *Treponema pallidum* subsp. *pertenue* suggests rare interspecies transmission in African nonhuman primates

SCIENTIFIC REPORTS
nature research

OPEN

Strain diversity of *Treponema pallidum* subsp. *pertenue* suggests rare interspecies transmission in African nonhuman primates

Received: 11 February 2019
Accepted: 19 September 2019
Published online: 02 October 2019

Idrissa S. Chuma[1,2,3], Christian Roos[2,4], Anagaw Atickem[5], Torsten Bohm[6], D. Anthony Collins[7], Linda Grillová[8,9], Luisa K. Hallmaier-Wacker[1,2], Rudovick R. Kazwala[3], Julius D. Keyyu[10], Simone Lüert[1,2], Ulrich Maloueki[6,11], Jan Oppelt[9,12], Klára J. Petrželková[13,14,15], Alexander Piel[16,17], Fiona A. Stewart[16,17], David Šmajs[9] & Sascha Knauf[1,18]

In our most recent study, we found that in Tanzania infection with *Treponema pallidum* (*TP*) subsp. *pertenue* (*TPE*) is present in four different monkey species. In order to gain information on the diversity and epidemiological spread of the infection in Tanzanian nonhuman primates (NHP), we identified two suitable candidate genes for multi-locus sequence typing (MLST). We demonstrate the functionality of the MLST system in invasively and non-invasively collected samples. While we were not able to demonstrate frequent interspecies transmission of *TPE* in Tanzanian monkeys, our results show a clustering of *TPE* strains according to geography and not host species, which is suggestive for rare transmission events between different NHP species. In addition to the geographic stability, we describe the relative temporal stability of the strains infecting NHPs and identified multi-strain infection. Differences between *TPE* strains of NHP and human origin are highlighted. Our results show that antibiotic resistance does not occur in Tanzanian *TPE* strains of NHP origin.

Nonhuman primates (NHPs) in Africa are naturally infected with *Treponema pallidum* subsp. *pertenue* (*TPE*)[1], the bacterium causing human yaws. In our most recent study, we found that in Tanzania infection with *TPE* is present in four different NHP species (olive baboon (*Papio anubis*), yellow baboon (*Papio cynocephalus*), vervet monkey (*Chlorocebus pygerythrus*), and blue monkey (*Cercopithecus mitis*))[2]. Moreover, we showed that infection is geographically widespread within Tanzania. Although we confirmed infection by serology and PCR, the data were insufficient to describe the epidemiology of the disease. Further insights into the inter- and intraspecies spread of the *TPE* bacterium in Tanzanian NHPs will contribute to our understanding of transmission pathways and pathogen maintenance, which are crucial elements for the identification of a functional disease reservoir[3].

[1]Work Group Neglected Tropical Diseases, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany. [2]Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany. [3]Department of Veterinary Medicine and Public Health, Sokoine University of Agriculture, Morogoro, Tanzania. [4]Gene Bank of Primates, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany. [5]Department of Zoological Sciences, Addis Ababa University, Addis Ababa, Ethiopia. [6]African Parks, Odzala-Kokoua National Park, Brazzaville, Republic of the Congo. [7]Jane Goodall Institute, Kigoma, Tanzania. [8]Biology of Spirochetes Unit, Department of Microbiology, Institut Pasteur, Paris, France. [9]Department of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic. [10]Tanzania Wildlife Research Institute, Arusha, Tanzania. [11]Department of Biology, Faculty of Sciences, Kinshasa University, Kinshasa, Democratic Republic of the Congo. [12]CEITEC-Central European Institute of Technology, Masaryk University, Brno, Czech Republic. [13]Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic. [14]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Ceske Budejovice, Czech Republic. [15]Zoo Liberec, Liberec, Czech Republic. [16]School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, UK. [17]Greater Mahale Ecosystem Research and Conservation Project, Dar es Salaam, Tanzania. [18]Division of Microbiology and Animal Hygiene, University of Goettingen, Goettingen, Germany. Idrissa S. Chuma and Christian Roos contributed equally. Correspondence and requests for materials should be addressed to S.K. (email: sknauf@dpz.eu)

Locus[$]	Length	Protein function	No. of variable sites	SNVs frequency/kbp
TPESAMD_0136	1,412	FIBRONECTIN-BINDING protein	41	29.04
TPESAMD_0548	1,298	FadL ortholog, outer membrane protein	36	27.73
TPESAMD_0858	1,229	FadL ortholog	33	26.85
TPESAMD_0488	2,537	METHYL-accepting chemotaxis protein	31	12.22
TPESAMD_0865	1,445	FadL ortholog	26	17.99
TPESAMD_0326	2,502	BamA	22	8.40

Table 1. Genes with the highest SNVs frequency per kbp containing 22 and more SNVs among samples listed in Table S1. [$]Anotation and length of the genes were identified according to the *TPE* reference genome Samoa D (GenBank accession number CP002374.1). Protein predictions by Brinkman *et al.*[35] and Radolf and Kumar[36].

The chance that NHPs infected with *TPE* are a potential source for human infection has been discussed for tropical Africa[4]. However, naturally occurring transmission from NHPs to humans and *vice versa* has not been confirmed by current data, although phylogenetic analyses of whole genome sequences from *TPE*s of human and NHP origin suggest a rapid initial radiation of the ancestor of *TPE* across the different primate taxa, including humans[1].

Molecular typing is used to accurately distinguish between different strains of *T. pallidum* (*TP*) for epidemiological and surveillance analysis. The method has been extensively applied to the syphilis–causing bacterium (subsp. *pallidum*, *TPA*)[5] where it is used to describe its spatial, e.g.[6–9], and temporal, e.g.[10], subtype composition. Two recent studies suggested new typing systems for human yaws[11,12]. It was unclear though, whether the existing molecular typing systems for *TPA* or *TPE* of human origin can be applied to *TP* strains originating from NHPs.

In the current study, we identified suitable candidate genes for multi-locus sequence typing (MLST) in *TP* samples of NHP origin and investigated strain diversity of the NHP infecting strains in Tanzania. We hypothesized that interspecies transmission in NHPs is ongoing. Moreover, we show that our typing system can be applied to samples from other regions of Africa and to analyze *TP* in non-invasively collected fecal samples.

Materials and Methods

Ethical statement. No animals were handled specifically for this study. The ethical statement for the Tanzanian NHP samples has been published elsewhere[2,13]. Lesion swabs from Ethiopian grivet monkeys (*Chlorocebus aethiops*) were taken as part of a research investigation conducted by AA under the Ethiopian Wildlife Conservation Authority reference number 15ET-0000-BS-01. Noninvasively collected fecal samples of western lowland gorillas (*Gorilla gorilla gorilla*) from the southern part of Odzala-Kokoua National Park (OKNP) originate from a collaboration signed under a MoU between the Foundation Odzala-Kokoua, the German Primate Center, and the Institute of Vertebrate Biology, Czech Academy of Sciences in November 2017.

Design of the multi-locus sequence typing system. In order to identify most suitable candidate genes for MLST in *TPE* strains of NHP origin, we used 23 available complete and draft genome sequences of *TPE* from both human and NHPs from Africa and the Pacific regions (Table S1). Several criteria were applied to obtain most suitable gene loci for *TPE* MLST. First, we identified the most variable genes with accumulated single nucleotide variants (SNVs) in short DNA fragments (genes containing the highest SNVs frequency per kbp) and, at the same time, with potential ability to distinguish all strains used for this analysis (containing 22 and more variable sites; Table S1). We identified six candidate genes (Table 1) and compared the resolution power of phylogenetic trees based on genome-wide data and phylogenetic trees based on sequences of individual genes.

With this approach, we propose a new MLST scheme for *TPE* strains of NHP origin, based on sequencing of two variable loci (*TP0548* and *TP0488*). The typing scheme is able to reveal 70% of whole genome resolution. Further details on the identification of most variable genes, resolution power, and the selection of most suitable typing loci are provided in the Technical Appendix.

Samples included into the study and DNA extraction. Our study used *TP* positive DNA samples from 85 NHPs of six different species and three African countries (Tables 2 and S3). The samples originated from previously published[2,13] and ongoing research investigations. The different methods of DNA extraction are presented in the Technical Appendix.

DNA target enrichment. Before MLST, DNA extracted from fecal samples was enriched for bacterial DNA using the Looxter Enrichment Kit (Analytik Jena, Jena, Germany) following the manufacturer's protocol.

Polymerase chain reactions. *Multi-Locus Sequence Typing system*. *TP0548*: Amplification of a fragment of the *TP0548* gene was achieved using a nested PCR. The two-step PCR amplified a 1,065-bp long fragment of the target gene. Amplification and sequencing primers were used as reported elsewhere[14]. Briefly, the 50-µl reaction volume comprised 25 µl of the 2x Universe buffer (Universe High Fidelity Hot Start DNA Polymerase Kit, Biotool, Munich, Germany), 17 µl RNAase free water, 2 µl of each 10 µM primer, 1 µl DNA polymerase (1 U/µl), 1 µl of the dNTP mix (10 mM each), and 2 µl template DNA, independent of DNA concentration. Amplification was performed in a SensoQuest Labcycler using the following thermocycler conditions: pre-denaturation at 95 °C for 3 min, followed by 40 and 30 cycles, respectively, each with 95 °C for 15 sec, 48 °C for 15 sec, and 72 °C for 30 sec. Each of the PCR runs ended with a post-extension step at 72 °C for 5 min.

NHP species	Sample type (n NHPs)[$]	Geographic origin	Reference
Olive baboon (*Papio anubis*)	Skin (61)[%], lesion swab (3)	Tanzania	(2)
Yellow baboon (*Papio cynocephalus*)	Skin (7)	Tanzania	(2)
Vervet monkey (*Chlorocebus pygerythrus*)	Skin (8)	Tanzania	(2)
Blue monkey (*Cercopithecus mitis*)	Skin (1)	Tanzania	(2)
Grivet monkey (*Chlorocebus aethiops*)	Lesion swabs (2)	Ethiopia	Ongoing research, unpublished
Western lowland gorilla (*Gorilla gorilla gorilla*)	Feces (4)	Republic of the Congo	Ongoing research, unpublished

Table 2. Overview of NHP species, sample types, and geographic origin. Only previously *TP* positive tested samples have been included. [$]The number of NHPs (n) which were sampled is not necessarily equal to the number of strain sequences. In a few cases (n = 3) multi-strain infection was present, which increased the sequence data output. [%]Including one lymph node aspiration sample.

TP0488: The PCR amplified a 837 bp-long fragment of the *TP0488* gene using 5′-CCC TGC GCA CCA AGC TC-3′ and 5′-ACA CAG GCC CCA TAA ACT-3′ primers. Briefly, the 51-µl reaction volume comprised 45 µl Platinum PCR Super Mix High Fidelity (ThermoFisher Scientific, Munich, Germany), 2 µl of each 10 µmol/L primer, and 2 µl template DNA, independent of DNA concentration. Amplification was performed in a SensoQuest Labcycler using the following thermocycler conditions: pre-denaturation at 94 °C for 2 min, followed by 80 cycles each with 94 °C for 15 sec, 59 °C for 15 sec, and 68 °C for 1 min.

Additional gene targets. *TP0619*: Although the *TP0619* locus is not part of the newly designed typing system, it was amplified to further discriminate between *TPA* and *Treponema pallidum* subsp. *endemicum* (*TEN*) strains. We performed the PCR as described previously[2]. PCR conditions included a pre-denaturation at 95 °C for 3 min, followed by 40 cycles each with 95 °C for 15 sec, 55 °C for 15 sec, and 72 °C for 30 sec. The PCR run ended with a post-extension step at 72 °C for 5 min.

Gel electrophoresis, purification, and DNA sequencing. All PCR products were run on 1% agarose gels to check for PCR performance and correct amplicon size. Products of the appropriate size were gel extracted and purified with the Qiagen Gel Extraction Kit (Qiagen, Hilden, Germany) and subsequently Sanger sequenced using the Microsynth SeqLab Laboratory service (Microsynth, Göttingen, Germany).

23S ribosomal RNA gene restriction enzyme analysis. Additional analyses of the 23 S ribosomal RNA gene was conducted to identify the two point-mutations including A2058G and A2059G that encode for macrolide resistance[14–16]. We note here that this locus is not part of the MLST, but it provides an essential information on macrolide resistance in NHP infecting *TP* strains. Methods were performed according to the procedures described by Lukehart *et al.*[17] with minor modifications. Amplification was done using a semi-nested PCR. The primers of the first PCR were S-primer 5′-GTA CCG CAA ACC GAC ACA G-3′ and AS-primer 5′-GCG CGA ACA CCT CTT TTT AC-3′ using an annealing temperature of 62 °C and 35 cycles. All other PCR conditions were identical with those for *TP0619*. The second PCR was performed with the S-primer from the first PCR and AS-primer 5′-AGT CAA ACC GCC CAC CTA C-3′. Amplification was conducted with an annealing temperature of 63 °C and 30 cycles. Again, all other PCR conditions were the same as for *TP0619*. PCR products were run on 1% agarose gels to check for the expected 628 bp-sized PCR product. Subsequently, each sample was digested overnight with restriction enzyme *Mbo* II (R0148, BioLabs New England, Beverly, MA, USA) at 37 °C and in a second reaction with *Bsa* I (R0535, BioLabs New England, Beverly, MA, USA) at 50 °C. Interpretation of results was performed as published previously[17]. *TPA* strains Street 14 (known mutation at A2058G, digested with *Mbo* II) and UW330B (known mutation at A2059G, digested with *Bsa* I) were included as positive controls.

Data analysis. Sequence data were analyzed, edited, and aligned using Geneious v11.1.4 (Biomatters Ltd., Auckland, New Zealand). We compared sequences with respective orthologs available in GenBank using BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) with megablast or blastn tools. The Free Phylogenetic Network Software (www.fluxus-eneineering.com) was used to create a median-joining network as described elsewhere[18]. The respective concatenated sequence alignment (*TP0548* and *TP0488*) was used in combination with the assigned traits 'NHP species' and 'sample location trait'. Maximum-parsimony (MP) trees, with gaps coded as fifth character, were constructed with SeaView 4.5.6[19]. Maximum-likelihood (ML) trees were calculated in IQ-TREE 1.6.1[20] with the respective best-fit models. To obtain node support, all trees were constructed with 1,000 bootstrap replicates.

Results

PCR results obtained from 85 NHP samples, of which 79 originated from Tanzanian monkeys (olive baboon, yellow baboon, vervet monkey, and blue monkey), two from Ethiopian grivet monkeys (*Chlorocebus aethiops*), and four from western lowland gorillas (*Gorilla gorilla gorilla*) from the Republic of the Congo, are summarized in Table S3. In general, PCR performances of the two gene targets used for the strain typing were equivalent with 78 (*TP0548*) and 67 (*TP0488*) resulting sequence data. This created a total of 59 concatenated sequences (the two gene targets combined) that were used for further analysis. We excluded the grivet monkey and the gorilla sample from the network analysis (Fig. 1) in order to maintain the highest possible resolution for the Tanzanian

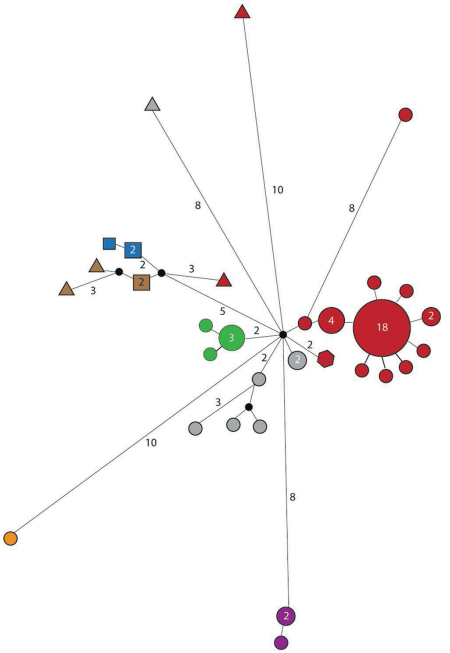


Figure 1. Median-joining network using 1,773 bp – long concatemer of *TP0488* and *TP0548* loci from 57 Tanzanian NHPs samples. The number of mutations, when >1, is given close to branches. Inferred allelic variants (median vectors) are shown as small black connecting circles. If contiguous, indels were considered as a single event only. The number of individuals, when >1, is shown inside the circles and are dependent on circle size. Species trait is given in the geometric form: circle = *Papio anubis* (n = 46); squares = *Papio cynocephalus* (n = 5); triangles = *Chlorocebus pygerythrus* (n = 5); hexagon = *Cercopithecus mitis* (n = 1). Sample location trait is given by the color code: blue = UG (n = 3); orange = TN (n = 1); brown = Ruaha National Park (n = 4); red = Lake Manyara National Park (n = 34); grey = Serengeti National Park (n = 7); green = Ngorongoro Conservation Area (n = 5); violet = Gombe National Park (n = 3).

NHP samples where a high number of samples came from one geographic location. However, we used the complete concatenated sequence alignment, including the data obtained from the grivet monkey and gorilla samples, to construct ML (Fig. 2) and MP (Fig. S6) trees. For the latter, gaps were coded as fifth character. NHP strains included in this study clustered with human yaws-causing strains and were clearly separated from the *TEN* strain Bosnia A as well as the human syphilis-causing *TPA* strains (Figs 2 and S7). Within the *TPE* clade, bootstrap support was mostly weak (<80%) with some exceptions. One of these notable exceptions was the separation of the grivet monkey infecting strain from Ethiopia and the strain that was generated from a gorilla fecal sample. Sequences of both strains (14AWM2051017 and 3DZAKM13280917) differed in only one nucleotide position and always clustered together (Figs 2 and S6). Overall, we found a geographic clustering of *TPE* strains, instead of clustering by host species. None of the samples from the Tanzanian NHPs (n = 72/76; four samples did not generate a PCR product) has been tested positive for the mutations in the 23 S ribosomal RNA genes that code for macrolide resistance. The grivet monkey and the gorilla samples were not tested for microbial resistance.

Sequences obtained from the *TP0619* gene were identical across all NHP species and sampling sites, including the grivet monkey and the gorilla samples. A representative sequence obtained from a vervet monkey sampled at Katavi National Park (4KNF2121016) was published previously under the GenBank accession number MF754122[2].

In general, all NHP-derived sequences of the *TP0548* locus including the grivet monkey and the gorilla sequence, contained one section of the sequence where most of the nucleotide variation is found (Fig. S1). This distinguishes them not only from human syphilis- (*TPA*) and bejel-causing (*TEN*) strains but also from *TPE* strains of human origin, where there are three and two variable regions, respectively. Sequences of the *TP0548*

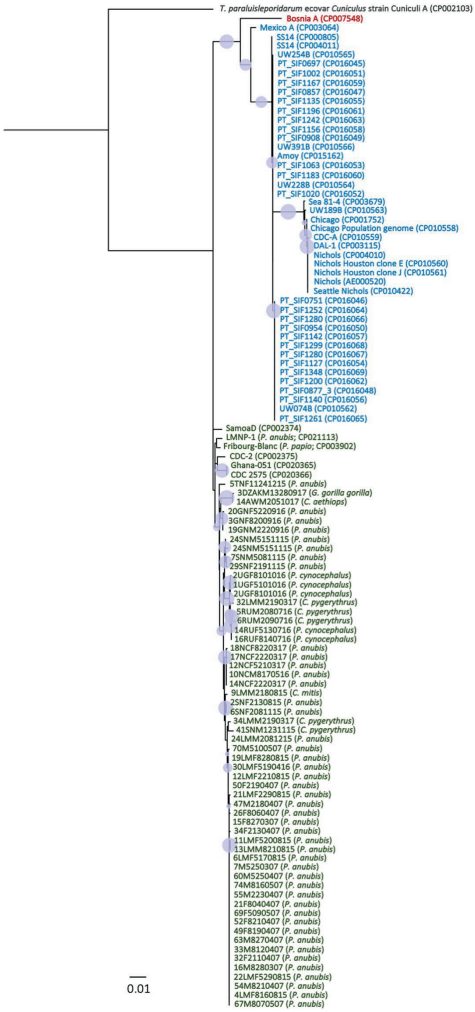


Figure 2. Rooted ML tree based on the concatenated sequences used for MLST (*TP0548* and *TP0488*). The tree is based on 1,773 nts and 1,000 bootstrap replicates. Bootstrap values from 80–100% are highlighted as light blue circles of respective size. NHP species and/or GenBank accession numbers of published strains are provided in parentheses following the name of the strain. In all cases were the species is not mentioned, sequences are from *TP* of human origin. Blue = subsp. *pallidum*, green = subsp. *pertenue*, red = subsp. *endemicum*. The pathogen causing rabbit syphilis, *Treponema paraluisleporidarum* ecovar *Cuniculus* strain Cuniculi A, is used as an outgroup. The bar refers to substitutions per site.

and *TP0488* loci showed comprehensive variability within and across the different sampling locations as well as between the different NHP species in Tanzania. Corresponding ML and MP trees each constructed for *TP0548* (Figs S2 and S3) and *TP0488* (Figs S4 and S5) were similar in topology. We note here that sample size was low

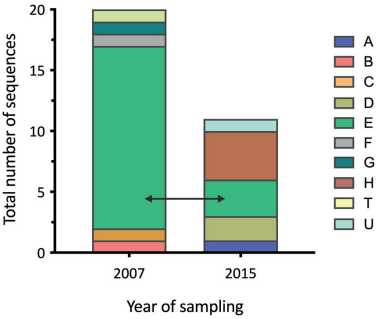


Figure 3. Temporal strain composition for samples collected from olive baboons at Lake Manyara National Park. The different colors indicate different genotypes. Sampling of olive baboons at LMNP was conducted in 2007 and 2015. The strain "E" (light green; e.g., 4LMF8160815) was the dominating strain in 2007 and was still present in baboons sampled eight years later (2015).

for some of the species (e.g., blue monkey) and that we included here only the Tanzanian samples to match the network analysis (Fig. 1).

Since the largest number of samples was taken at Lake Manyara National Park (LMNP) and samples were collected in 2007, 2015, and 2017, respectively, we were able to plot the temporal strain composition for this specific sampling location. Figure 3 illustrates the temporal strain composition for all samples that originate from olive baboons. Using a sequence alignment and base-by-base comparison, we found that in 2007 there were six strains, with a dominating strain "E" (e.g., 4LMF8160815). Samples from olive baboons at LMNP taken eight years later reconfirmed the existence of this strain. The temporal stability of strains was further supported by strain "P" (e.g., 10NCM8170516) that was collected from two olive baboons at the Ngorongoro Conservation Area in 2015 and one olive baboon in 2016 (Table S4).

Discussion

Frequent transmission of *TPE* strains across the different NHP species would likely result in a higher number of shared and identical sequences in different species at the same geographic site. Although we found no identical strain sequences in different primate species at one site (Table S4), we still observed a geographic clustering of (closely related) *TPE* strain sequences and not a clustering according to NHP host species (Fig. 1). This suggests that interspecies transmission occurs, albeit rarely. That we found no identical strain sequences in different primate species could be due to the small sample size for some of the investigated NHP species, but generally argues against frequent transmission. This gains further support by our finding that the strains of NHP origin are relatively stable over time as we show for LMNP (Fig. 3, Table S4). This is consistent with what we know from human infecting *TPA* strains[21–23]. The 2007 dominating strain "E" (e.g., 4LMF8160815), which was isolated from olive baboons was still present in the infected baboon population in 2015. Although we do not see a strict geographic pattern when the two target genes were analyzed individually (Figs S2–S5), when examining the concatenated alignment (median-joining network (Fig. 1) and the ML tree (Fig. 2)), geographic strain clustering can be observed. This is another indicator for the relative temporal and geographic stability of the strains that infect NHPs. The recently estimated low mutation rate in human *TPE* strains[22] coincides with the genetic stability observed among NHP *TPE* strains analyzed in this study. Yet, feasible interspecies transmission routes for *TP* exist[3] and have been discussed for flies[24] and were proven for sexual intercourse between different NHP species based on host genetic data[25].

Despite the differences in the number of variable sites at the *TP0548* locus (Fig. S1), we see a close association of NHP and human infecting *TPE* strains. This was expected, since the genome of *TPE* LMNP-1 strain (GenBank accession number CP021113), which was obtained from an olive baboon at LMNP in 2007, was found to be closely related to human yaws causing strains similar to all other *TPE* strains of NHP origin[1].

The absence of antibiotic resistance to azithromycin in all tested *TP* strains from NHPs in Tanzania is a positive sign and is probably related to the absence of treatment of infected monkeys. Currently only Gombe National Park has a history of treating infected baboons with antimicrobials[26]. In human yaws, it has been shown that after a single treatment round with antibiotic macrolides, resistance emerges[27] even though the *de novo* emergence of such mutations is lower than 10^{-3} per treated patient[28]. The risk of emerging antimicrobial resistance is of major concern for human infection and would also draw major implications for the conservation of endangered NHP species such as gorillas (*Gorilla gorilla*). Similar to human yaws elimination[27], responsible treatment of infected NHP populations requires resistance monitoring and possible ring-fencing with effective alternative antimicrobials.

The identification of two different strains obtained from independent PCRs in three different NHPs (*TP0548*: 11LMF8190815 (*P. anubis*; 15-bp indel); *TP0488*: 2UGF8101016 (*P. cynocephalus*; one SNV) and 24SNM5151115

(*P. anubis*; three SNVs); Table S4) supports the concept of absence of cross immunity between different *TP* strains[29] and boosts evidence for recombination events found in the *TP* bacterium[30–32] under natural conditions.

The initial CDC typing system for *TPA* made use of the number of 60-bp repeats found in the acidic repeat protein (*arp* (*TP0433*)) gene in combination with differences found in the *tpr* subfamily II genes (*tprE* (*TP0313*), *tprG* (*TP0317*), and *tprJ* (*TP0621*)[9]. A subsequently introduced enhanced typing system included a portion of the *TP0548* locus[33]. While the enhanced typing system has also been used for the typing of human *TPE* strains[12], it did not overcome the difficulties associated with amplification of 60-bp repeats of the *arp* gene or the uncertainties associated with amplification of three different *tpr*-subfamily II genes in one single assay, followed by subsequent restriction enzyme analysis[11]. A recently published alternative method for MLST included the widely used *TP0548* locus but also two additional loci located in the *TP0136* and *TP0326* genes[11]. We took these loci into account, but in our analysis, it became evident that the *TP0136* locus in the Tanzanian *TP* strains of NHP origin was highly conserved. A representative sequence is accessible under GenBank accession number CP021113.1 (nt158,275-159,195). Whether this is a characteristic of NHP infecting strains or a spatial property of strains originating from Tanzania is unknown. It underlines, however, that a globally applied strain typing system for *TPE* requires a comprehensive database of high-quality genomes obtained from larger numbers of clinical samples from yaws endemic areas.

We identified several gene loci as suitable candidates to be used in *TPE* strains of human and NHP origin. The two gene targets that we selected for MLST (*TP0548* and *TP0488*) originated from a number of suitable candidates (Table 2) and selection was based on best PCR performance in the clinical NHP samples that were included in this study. We neither aimed for the design of a *TPE* typing system that can be used in a clinical environment for human infection, nor did we anticipate a typing system that is suitable for a global approach. Amplification of relatively long sequence parts (e.g., *TP0548* enhanced typing system determines 84 bp[33] vs. 1,065 bp determined in this study) was therefore not considered an issue. Both loci that were used for the typing, *TP0548* and *TP0488*, are reported to show signs of recombination in human syphilis causing *TPA* strains[30–32]. This, however, was not considered a limitation since our MLST typing system was designed to describe strain variability within a given population of NHPs and is not used to describe a detailed geophylogeny of *TP*. While the *TP0619* sequence, which codes for a protein family of Domains of Unknown Function (DUF)2715 and which appear to be restricted to *TP*, is not part of our typing system, it is useful to support the difference of the *TP* strains of NHP origin analyzed in this study from human syphilis-causing *TPA* strains. It is currently unclear if wild NHPs are also naturally infected with *TPA* strains.

Invasive sampling of NHPs in the wild is generally associated with challenges in terms of ethics and logistics. As a consequence, it is easier to obtain a greater number of non-invasively collected fecal samples than samples that originate from invasively sampled individuals. Moreover, prospective epidemiological studies would benefit from sampling regimes that allow the screening and subsequent typing of *TP* strains in non-invasively collected samples from a greater geographic area. For this reason, we tested our newly established MLST system in non-invasively collected fecal samples from gorillas at Odzala-Kokoua National Park in the Republic of the Congo, a place where gorillas with ulcerative skin lesions have been frequently sighted[34] (Fig. S7). We were able to successfully strain type the gorilla samples and could show that the strain from which we obtained all three target sequences, clusters with a strain isolated from a grivet monkey in Ethiopia. The reason for the close association of the strain of the grivet monkey and the gorilla origin is currently unclear and answering this question requires intensified sampling in the respective geographic areas and in-between to obtain more sequence data.

Conclusion

The high number of *TP* infected NHPs in Africa, the different species that have been confirmed as hosts[2,4], as well as the recently documented close genetic and functional similarity of NHP and human infecting *TPE*[1] requires epidemiological data for a better understanding on how the infection is maintained in primate populations and whether or not it is transmitted to humans. Our study provides an important contribution to answer the question on interspecies transmission in primate infecting *TP*, although further sampling is needed to increase confidence in the results. However, with our data we were able to show that interspecies transmission in Tanzanian monkeys is likely although rare. As humans are primates, the most important question to answer in future studies is, whether or not *TPE* strains of NHP origin transmit to humans and/or *vice versa*. The interspecies transmission of *TPE* in nonhuman primates, is not necessarily predictive for spillovers to humans.

Data Availability

GenBank accession numbers for the sequences generated in this study can be found in Table S5.

References

1. Knauf, S. *et al*. Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. *Emerg Microbes Infect* **7**, 157 (2018).
2. Chuma, I. S. *et al*. Widespread *Treponema pallidum* infection in nonhuman primates, Tanzania. *Emerg Infect Dis* **24**, 1002–1009 (2018).
3. Hallmaier-Wacker, L. K., Munster, V. J. & Knauf, S. Disease reservoirs: from conceptual frameworks to applicable criteria. *Emerg Microbes Infect* **6**, e79 (2017).
4. Knauf, S., Liu, H. & Harper, K. N. Treponemal infection in nonhuman primates as possible reservoir for human yaws. *Emerg Infect Dis* **19**, 2058–60 (2013).
5. Peng, R. *et al*. Molecular typing of *Treponema pallidum*: a systematic review and meta-analysis. *PLoS Negl Trop Dis* **5**, e1273 (2011).
6. Salado-Rasmussen, K. *et al*. Molecular typing of *Treponema pallidum* in Denmark: A nationwide study of syphilis. *Acta Derm Venereol* **96**, 202–6 (2016).
7. Grange, P. *et al*. Molecular subtyping of *Treponema pallidum* in Paris, France. *Sex Transm Dis* **40**, 641–4 (2013).
8. Pope, V. *et al*. Molecular subtyping of *Treponema pallidum* from North and South Carolina. *J Clin Microbiol* **43**, 3743–3746 (2005).
9. Pillay, A. *et al*. Molecular subtyping of *Treponema pallidum* subspecies *pallidum*. *Sex Transm Dis* **25**, 408–414 (1998).

10. Grillová, L. *et al.* Molecular typing of *Treponema pallidum* in the Czech Republic during 2011 to 2013: increased prevalence of identified genotypes and of isolates with macrolide resistance. *J Clin Microbiol* **52**, 3693–3700 (2014).
11. Godornes, C., Giacani, L., Barry, A., Mitjà, O. & Lukehart, S. Development of a Multilocus sequence typing (MLST) scheme for *Treponema pallidum* subsp. *pertenue*: Application to yaws in Lihir Island, Papua New Guinea. *PLoS Neglect Trop Dis* **11**, e0006113 (2017).
12. Katz, S. S. *et al.* Molecular strain typing of the yaws pathogen, *Treponema pallidum* subspecies *pertenue*. *PLoS ONE* **13**, e0203632 (2018).
13. Knauf, S. *et al. Treponema* infection associated with genital ulceration in wild baboons. *Vet Pathol* **49**, 292–303 (2011).
14. Matějková, P. *et al.* Macrolide treatment failure in a case of secondary syphilis: a novel A2059G mutation in the 23S rRNA gene of *Treponema pallidum* subsp. *pallidum*. *J Med Microbiol* **58**, 832–836 (2009).
15. Chen, C. *et al.* Detection of the A2058G and A2059G 23S rRNA gene point mutations associated with azithromycin resistance in *Treponema pallidum* by use of a TaqMan real-time multiplex PCR assay. *J Clin Microbiol* **51**, 908–913 (2013).
16. Stamm, L. & Bergen, H. A point mutation associated with bacterial macrolide resistance is present in both 23S rRNA genes of an erythromycin-resistant *Treponema pallidum* clinical isolate. *Antimicrob Agents and Chemother* **44**, 806–807 (2000).
17. Lukehart, S. *et al.* Macrolide resistance in *Treponema pallidum* in the United States and Ireland. *N Engl J Med* **351**, 154–158 (2014).
18. Bandelt, H., Forster, P. & Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol and Evol* **16**, 37–48 (1999).
19. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221–224 (2010).
20. Nguyen, L.-T. T., Schmidt, H. A., Haeseler, A. & von Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol and Evol* **32**, 268–274 (2015).
21. Strouhal, M. *et al.* Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Ghana, Africa: Identical genome sequences in samples isolated more than 7 years apart. *PLoS Negl Trop Dis* **11**, e0005894 (2017).
22. Grillová, L. *et al.* Molecular characterization of *Treponema pallidum* subsp. *pallidum* in Switzerland and France with a new multilocus sequence typing scheme. *PloS ONE* **13**, e0200773 (2018).
23. Grillová, L. *et al.* Sequencing of *Treponema pallidum* subsp. *pallidum* from isolate UZ1974 using anti-treponemal antibodies enrichment: First complete whole genome sequence obtained directly from human clinical material. *PLoS ONE* **13**, e0202619 (2018).
24. Knauf, S. *et al.* Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. *EBioMed* **11**, 85–90 (2016).
25. Zinner, D., Chuma, I. S., Knauf, S. & Roos, C. Inverted intergeneric introgression between critically endangered kipunjis and yellow baboons in two disjunct populations. *Biol Lett*, **14** (2018).
26. Collins, D., Sindimwo, A. & Wallis, P. Reproductive disease of olive baboons (*Papio anubis*) of Gombe National Park: outbreak, time-course, and attempts to limit recurrence. In: *Climate change: challenges and opportunities towards sustainable wildlife conservation and development*. 236–248 (*Tanzania Wildlife Research Institute*, 2011).
27. Mitjà, O. *et al.* Re-emergence of yaws after single mass azithromycin treatment followed by targeted treatment: a longitudinal study. *Lancet* **391**, 1599–1607 (2018).
28. Šmajs, D. & Pospíšilová, P. Macrolide resistance in yaws. *Lancet* **391**, 1555–1556 (2018).
29. Turner, T. & Hollander, D. Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. *Monogr Ser World Health Organ* **35**, 3–266 (1957).
30. Pětrošová, H. *et al.* Whole genome sequence of *Treponema pallidum* ssp. *pallidum*, strain Mexico A, suggests recombination between yaws and syphilis strains. *PLoS Negl Trop Dis* **6**, e1832 (2012).
31. Grange, P. *et al. Treponema pallidum* 11qj subtype may correspond to a *Treponema pallidum* subsp. *endemicum* strain. *Sex Transm Dis* **43**, 517–518 (2016).
32. Mikalová, L. *et al.* Human *Treponema pallidum* 11q/j isolate belongs to subsp. *endemicum* but contains two loci with a sequence in TP0548 and TP0488 similar to subsp. *pertenue* and subsp. *pallidum*, respectively. *PLoS Negl Trop Dis* **11**, e0005434 (2017).
33. Marra, C. *et al.* Enhanced molecular typing of *Treponema pallidum*: geographical distribution of strain types and association with neurosyphilis. *J Infect Dis* **202**, 1380–1388 (2010).
34. Levrero, F., Gatti, S., Gauthier-Hion, A. & Menard, N. Yaws disease in a wild gorilla population and its impact on the reproductive status of males. *Am J Phys Anthropol* **132**, 568–575 (2007).
35. Brinkman, M. *et al.* A novel *Treponema pallidum* antigen, TP0136, is an outer membrane protein that binds human fibronectin. *Infect Immun* **76**, 1848–1857 (2008).
36. Radolf, J. D. & Kumar, S. The *Treponema pallidum* outer membrane. *Curr Top Microbiol Immunol* **415**, 1–38 (2018).

Acknowledgements

The study was funded by the German Research Foundation (DFG): KN1097/3-1 and KN1097/4-1 [SK], and RO3055/2-1 [CR]. In addition, it was partly supported by the Grant Agency of the Czech Republic (GA17-25455S; GC18-23521J) to DS. We thank the Government of the United Republic of Tanzania (URT) for permission and logistical support that was needed. Institutions include the Ministry for Education and Vocational Training (MoEVT), the Commission for Science and Technology (COSTECH), Ministry for Natural Resources and Tourism (MNRT), Ministry for of Agriculture, Natural Resources, Livestock and Fisheries (SMZ), Department of Forestry and Non-renewable Natural Resources (SMZ), Tanzania Wildlife Authority (TAWA), Tanzania Wildlife Research Institute (TAWIRI), Tanzania National Parks (TANAPA), Ngorongoro Conservation Area Authority (NCAA), Sokoine University of Agriculture (SUA), Department of Livestock Development (SMZ). We are also grateful for the support received from the Jane Goodall Institute (JGI) and the National Institute for Medical Research (NIMR). Most importantly, our gratitude goes to all people and staff-members working with government and partner institutions in Germany and Tanzania who tirelessly strived to facilitate smooth undertaking of the project at all levels from planning, execution of fieldwork as well as laboratory analyses. Sheila A. Lukehart and Charmie Godornes (University of Washington) are thanked for technical support and scientific advice.

Author Contributions

The study was designed by I.S.C., C.R., D.S. and S.K. Design of the MLST system was done by C.R., L.G., J.O., D.S. and S.K. Laboratory work took place at Sokoine University of Agriculture and the German Primate Center and was performed by I.S.C., C.R., L.H.W., S.L. and S.K. Data were analyzed by I.S.C., C.R., D.S. and S.K. All authors (I.S.C., C.R., A.A., T.B., D.A.C., L.G., L.H.W., R.R.K., J.D.K., S.L., U.M., J.O., K.J.P., A.P., F.A.S., D.S. and S.K.) contributed to the manuscript preparation.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-50779-9.

Competing Interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

7.4.2 Isolation of *Treponema* DNA from Necrophagous Flies in a Natural Ecosystem

EBioMedicine 11 (2016) 85–90

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Isolation of *Treponema* DNA from Necrophagous Flies in a Natural Ecosystem

Sascha Knauf [a,*], Jane Raphael [b], Oriol Mitjà [c], Inyasi A.V. Lejora [b], Idrissa S. Chuma [b], Emmanuel K. Batamuzi [d], Julius D. Keyyu [e], Robert Fyumagwa [e], Simone Lüert [a,f], Charmie Godornes [g], Hsi Liu [h], Christiane Schwarz [f], David Šmajs [i], Philippe Grange [j], Dietmar Zinner [k], Christian Roos [f,l,1], Sheila A. Lukehart [g,1]

[a] Work Group Neglected Tropical Diseases, Pathology Unit, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany
[b] Ecology Monitoring Department, Tanzania National Parks, P.O. Box 3134, Arusha, Tanzania
[c] ISGlobal, Barcelona Centre for International Health Research, Hospital Clínic, Universitat de Barcelona, Barcelona, Spain
[d] Department of Surgery and Theriogenology, Faculty of Veterinary Medicine, Sokoine University of Agriculture, P.O. Box 3020, Morogoro, Tanzania
[e] Tanzania Wildlife Research Institute, P.O. Box. 661, Arusha, Tanzania
[f] Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany
[g] Department of Medicine, University of Washington, Harborview Medical Center, 325 Ninth Avenue, Seattle, WA 98104, USA
[h] National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Diseases Control and Prevention, Atlanta, GA 30333, USA
[i] Department of Biology, Faculty of Medicine, Masaryk University, Kamenice 5, Brno 625 00, Czech Republic
[j] Laboratoire de Dermatologie, Centre National de Reference de la Syphilis, 75014 Paris, France
[k] Cognitive Ethology Laboratory, German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany
[l] Gene Bank of Primates, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany

ARTICLE INFO

Article history:
Received 16 June 2016
Received in revised form 26 July 2016
Accepted 27 July 2016
Available online 28 July 2016

Keywords:
Treponema pallidum
Diptera
Yaws
Nonhuman primates
Transmission

ABSTRACT

Background: Recently, the World Health Organization launched a campaign to eradicate the tropical disease yaws, caused by the bacterium *Treponema pallidum* subsp. *pertenue*; however, for decades researchers have questioned whether flies act as a vector for the pathogen that could facilitate transmission.
Methods: A total of 207 fly specimens were trapped in areas of Africa in which *T. pallidum*-induced skin ulcerations are common in wild baboons; 88 flies from Tarangire National Park and 119 from Lake Manyara National Park in Tanzania were analyzed by PCR for the presence of *T. pallidum* DNA.
Findings: We report that in the two study areas, *T. pallidum* DNA was found in 17–24% of wild-caught flies of the order Diptera. Treponemal DNA sequences obtained from many of the flies match sequences derived from nearby baboon *T. pallidum* strains, and one of the fly species with an especially high prevalence of *T. pallidum* DNA, *Musca sorbens*, has previously been shown to transmit yaws in an experimental setting.
Interpretation: Our results raise the possibility that flies play a role in yaws transmission; further research is warranted, given how important understanding transmission is for the eradication of this disfiguring disease.

1. Introduction

Human yaws, a neglected tropical disease caused by *T. pallidum* subsp. *pertenue* (TPE), has been targeted for global eradication by 2020, after previous attempts in the 1950's (Mitja et al., 2013) did not eradicate the disease (Asiedu et al., 2014). Factors involved in the natural history of yaws, including transmission by flies and a possible nonhuman reservoir (Knauf et al., 2013), may have contributed to yaws re-emergence. In humans, *Treponema pallidum* subsp. *pallidum* (TPA) causes syphilis, while TPE causes yaws (Lukehart, 2008; Giacani and Lukehart, 2014).

We reported previously high levels of treponemal infection in nonhuman primates (NHPs) with strains closely related to the human yaws-causing strains (Knauf et al., 2013). This was shown on the basis of DNA polymorphisms that are known to distinguish among the subspecies of *T. pallidum* (Harper et al., 2012; Knauf et al., 2012) as well as on the basis of the genome sequence of the Fribourg-Blanc strain (str. F-B) (Zobaníková et al., 2013), which was isolated from a baboon in Guinea in 1966 (Fribourg-Blanc and Mollaret, 1969). In addition, the subspecies of *T. pallidum* in humans (Sena et al., 2010) and those in NHPs (Knauf et al., 2015) are not distinguishable based on serology. The transmission of pathogenic treponemes by insects has been debated for over a century (Barnard, 1952a; Gudger, 1911; Gudger, 1910a; Gudger, 1910b; Barnard, 1952b; Lamborn, 1936), although there are no recent published studies.

The observation that flies function as carriers of human pathogens such as *Chlamydia trachomatis* (causing blinding trachoma) under

* Corresponding author at: Dr. Sascha Knauf, Work Group Neglected Tropical Diseases, Pathology Unit, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany.
E-mail address: sknauf@dpz.eu (S. Knauf).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ebiom.2016.07.033

natural conditions has already been reported (Emerson et al., 2000; Emerson et al., 1999). In contrast, the role of necrophagous flies in the transmission of yaws and other treponematoses, despite the early notes and models at the beginning of the last century (Lamborn, 1936; Castellani, 1907; Kumm, 1935a; Kumm et al., 1935; Kumm and Turner, 1936; Satchell and Harrison, 1953), is not clear. *Treponema* species have not yet been reported to be present on necrophagous flies caught in a natural ecosystem and, most importantly, molecular evidence of the presence of the different subspecies has not been presented in any reported studies to date. In addition, a better understanding of different modes of transmission of *T. pallidum* is important for the development of sustainable control strategies. Vector transmission of the yaws bacterium for example, would underline the importance of additional hygiene measures and wound coverage when treating infected individuals. Moreover, if the simian yaws-like treponemes were demonstrated to be infectious to humans (Knauf et al., 2013), it could add an additional route of inter-species transmission. The West African simian isolate (str. F-B) is reported to have caused a sustained infection after experimental inoculation into humans (Smith et al., 1971).

Based on previous reports, we hypothesized that flies function as a mechanical vector for the transmission of *T. pallidum* in areas of high prevalence of *Treponema*-induced skin ulceration in primates. To test this hypothesis, we performed a prospective cross-sectional study in areas that have populations of olive baboons (*Papio anubis*) with *T. pallidum*-associated genital ulcerations at Tarangire National Park (TNP; unpublished data) and Lake Manyara National Park (LMNP) (Knauf et al., 2012) in Tanzania.

2. Materials and Methods

2.1. Study Sites and Sample Size

TNP (latitude (lat.): −3·83333, longitude (long.): 36·00000) and LMNP (lat.: −3·58333, long.: 35·83333) are located in the north of Tanzania. The two parks are about 60 km apart and both belong to the same greater ecosystem (Fig. S1). TNP is mainly dominated by savanna grassland, while LMNP consists of a variety of different vegetation types: groundwater forest with freshwater sources, open grassland, and seasonal swamps in the north, savanna bush land in the center and acacia forest in the far south of the park. Though TNP is much greater in size (2600 qkm TNP vs. 330 qkm LMNP), both national parks face a substantial pressure of human activities along their borders (Kiffner et al., 2015). Wildlife corridors are narrowed down and are almost closed by human settlements. Baboons are often found crop-raiding in neighboring villages and thus are in very frequent conflict with humans.

Reports of *T. pallidum* infected NHPs at TNP were investigated in 2015 and were confirmed by clinical manifestations, serology, and PCR in 2016 (unpublished data; Chuma et al.). At LMNP, infection in olive baboons has been reported since 1994 and was extensively characterized by a study in 2007 (Harper et al., 2012; Knauf et al., 2012).

A total of 207 fly specimens were trapped in 2014 in areas inhabited by infected baboons: 88 flies from TNP and 119 from LMNP. Sampling was conducted along road transects within the national parks and only in areas where NHPs and in particular baboons were known to occur Areas with no history of NHPs were excluded from sampling.

Two different conservation methods (RNA Later Solution (Ambion, Cat# AM7020) and air dried) were used to store fly specimens until analysis in the laboratory. Further details can be found in the Supplementary material, including a detailed description of flytraps and sampling procedures.

2.2. DNA Extraction

In the laboratory, fly bodies in RNA Later Solution were separated from the liquid. Supernatant transport fluid was collected in a new reaction tube and immediately stored at 4 °C. All working steps were performed under a laminar flow workbench (BSL2). DNA extraction of fly bodies was performed using the First-DNA-All-Tissue-Kit (Gen-ial), following the manufacturer's instructions with some minor modifications. Briefly, fly bodies were smashed in their respective reaction tube using a stamp of a sterile 1-ml syringe. Tissue lysate was then incubated in a mixture of the Gen-ial Kit's lysis buffer #1 (1000 μl) and #2 (100 μl) as well as 20 μl of proteinase K and 10 μl 1 M dithiothreitol (DTT, Cleland's reagent, Sigma Aldrich, Cat# D9779). The lysate was continuously mixed at 600 rpm and incubated at 65 °C (Thermomixer comfort, Eppendorf) for 1 h. Next, temperature was reduced to 37 °C and samples were constantly shaken at 300 rpm overnight. The remaining protocol followed the manufacturer's instructions. DNA pellets were re-suspended in 20–50 μl molecular grade water. Total DNA content and purity were measured using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific).

DNA from transport fluid was extracted using the (QIAamp DNA Mini Kit, Qiagen, Cat# 51306) as follows: fluids were centrifuged at 20,000 ×*g* for 30 min at room temperature, followed by the removal of the supernatant and re-suspension of the cell pellet in 500 μl lysis buffer (10 mM Tris [pH 8·0], 0·1 M EDTA, and 0·5% SDS). After vigorous mixing, the sample was incubated with 30 μl proteinase K for 30 min at 56 °C. Next, 200 μl of the kit's AL buffer was added and incubation was prolonged for another 10 min at 70 °C. The subsequent steps followed the manufacturer's guidance. DNA was eluted from the columns with 200 μl molecular grade water.

DNA extraction from fly specimen supernatant was followed by DNA precipitation to further purify and concentrate the yielded DNA. Therefore, 2·5 volumes (500 μl) ethanol 100%, 2 μl glycogen (Thermo Fisher, Cat# R0561) and 20 μl sodium acetate (3 M, pH 4·8) were added to the DNA solution, mixed for 1 min and stored at −20 °C overnight. On the next day, centrifugation followed for 45 min at 4 °C and 17,900 ×*g*. Supernatant was discarded and the remaining DNA pellet was washed with 800 μl 70% ethanol. This last step was succeeded by immediate discard of the ethanol and drying of the DNA pellet by centrifugation in a SpeedVac (Concentrator plus, Eppendorf) for 5–10 min (until DNA was free of ethanol). DNA was re-suspended in 50 μl molecular grade water.

2.3. Amplification of Treponema DNA and Fly Species Identification

Fly specimens were analyzed by PCR for the presence of *T. pallidum* DNA, targeting three different loci: DNA polymerase I (*polA*) (Liu et al., 2001), *tp0548* (Marra et al., 2010), and *tp0574* (*tp47*) (Marra et al., 2016). Target regions were selected purposely to achieve high sensitivity and specificity (*polA* and *tp0574*) as well as to identify possible *T. pallidum* strain variation, using *tp0548*, a gene that is commonly used for multi-locus strain typing of *T. pallidum* (Marra et al., 2010).

2.3.1. DNA Polymerase I (polA)

The PCR targeting the *T. pallidum* polymerase I gene (*polA*) was performed as described elsewhere (Liu et al., 2001). Briefly, PCR was performed in a reaction volume of 30 μl, which contained 1 U Biotherm Taq 5000 DNA polymerase (Ares Bioscience, Cat# GC-002-1000), 1× reaction buffer including $MgCl_2$, 0·64 mM dNTPs, 0·33 μM for each primer, 0·6 mg/ml BSA, 0·7% TritonX-100 and 2 μl template DNA. PCR conditions comprised a pre-denaturation step at 94 °C for 2 min, followed by 50 cycles of denaturation at 94 °C for 1 min, annealing at 60 °C for 1 min, and elongation at 72 °C for another min. Amplification was completed with a final extension step at 72 °C for 5 min.

2.3.2. tp0548

Amplification of *tp0548* was performed in a reaction volume of 25 μl using primers published elsewhere (Marra et al., 2010). Briefly, the reaction mix consisted of 12·5 μl Universe Hot Start High-Fidelity 2× PCR Master Mix (Biotool, Cat# 22,101) containing Universal Hot Start High-Fidelity DNA polymerase. The mix was completed with 1 μl of

each 10 μM primer and 4 μl DNA extract, independent of the overall genomic DNA content. PCR conditions comprised a pre-denaturation step at 95 °C for 3 min, followed by 50 cycles of denaturation at 95 °C for 15 s, annealing at 62 °C for 15 s, and elongation at 72 °C for 30 s. Amplification was completed with a final extension step at 72 °C for 5 min.

2.3.3. tp0574 (tp47)

Amplification of a 132 bp-long fragment of the *tp47* gene was performed in a reaction volume of 25 μl. Primers were used as published elsewhere (Marra et al., 2016). The reaction mix was identical to the one described for *tp0548*. PCR conditions comprised a pre-denaturation step at 94 °C for 3 min, followed by 45 cycles of denaturation at 94 °C for 15 s, annealing at 55 °C for 15 s, and elongation at 72 °C for 15 s. Amplification was completed with a final extension step at 72 °C for 5 min.

2.3.4. rDNA Internal Transcribed Spacer 2 (ITS-2) for Diptera Species Identification

Fly species identification was performed utilizing a PCR targeting a 500 bp-long fragment of the *ITS-2* gene as described elsewhere (Song et al., 2008). PCR conditions were equal to the ones described for amplification of *polA*, with the exception of the annealing temperature (52 °C).

2.4. Gel Electrophoresis, Purification and DNA Sequencing

All PCR products were run on 1% agarose gels to check for PCR performance and correct amplicon size. PCR products of *polA* and *tp0548* of the correct size were excised from the gel and purified with the Qiagen Gel Extraction Kit (Qiagen, Cat# 28,706) according to the manufacturer's protocol. Unspecific PCR products were occasionally observed, but ignored for further analysis. No sequence variation was expected for *tp0574* (*tp47*). Therefore, only PCR products amplified from two independent fly samples were Sanger sequenced to receive final proof of the correct PCR product (data not shown). Sanger sequencing was performed utilizing the BigDye Terminator Cycle Sequencing kit (Applied Biosystems, Cat# 4,337,455) and the respective forward and reverse amplification primers. Sequencing was conducted on an ABI 3130xl sequencer.

2.5. Data Analysis

Sequences were analyzed and edited using 4Peaks 1.8 (Griekspoor and Groothuis, nucleobytes.com) and SeaView 4.5.4 software (Gouy et al., 2010). Sequences were compared to respective orthologous sequences available in GenBank using the standard nucleotide (nt) BLAST search option (http://blast.ncbi.nlm.nih.gov/Blast.cgi) at the NCBI homepage.

Phylogenetic tree reconstruction was conducted using 26 *tp0548* sequences obtained from fly samples (accession numbers can be found under 2.6 or in Table S2) and one of 28 identical sequences from LMNP baboon skin tissue isolates as well as TPA and TPE reference strain sequence data obtained from GenBank. Accession numbers of reference strain sequences are as follows: TPE str.: Gauthier [CP002376], and CDC-2 [CP002375]; TP str. F-B [CP003902]; TPA str.: Nichols [AE000520], Dallas 1 [CP003115], Chicago [CP001752], and Seattle 81-4 [CP003679]); *T. p.* subsp. *endemicum* (TPEN) str. Bosnia A [CP007548]. The accession number of a sequence of *tp0548* of type "j", which was obtained from a human, was not available. Data were copied from the respective publication (Grange et al., 2013) and the extension of these data to cover the full sequence (nt 40 to 344) that was obtained from the flies was added from unpublished data (DS). A maximum parsimony (MP) tree with gaps as fifth character and 1000 bootstrap replicates was constructed in SeaView. Tree visualization and editing was performed in FigTree 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). Alignment data of sequences used in this study can be found in Fig. S3.

2.6. GenBank Accession Numbers

Sequences obtained from fly specimens are deposited in GenBank under the following accession numbers: KX301264-7 and KX301269-71. The baboon skin tissue isolate from LMNP is deposited under accession number KX301268. Details can be found in Table S2.

3. Results

Overall, 17·0% of the TNP and 23·5% of the LMNP fly specimens were positive for *T. pallidum* DNA (Table 1). Not all samples were positive or were tested for all three loci (Tables 1 and 2, and Table S2). The *polA* PCR reaction was positive less frequently than the detection rates obtained from the PCRs targeting *tp0548* or *tp0574*.

In 17 fly specimens, DNA extracts from both the body and the corresponding transport fluid were tested (Table S2). Two of the specimens were positive for *tp0548* in the body extract and transport fluid and seven specimens became positive only when the transport fluid was PCR analyzed. Sequences derived from both extracts of the same fly specimen were identical (fly# 105, Table S2). For the *tp0547* gene, extracts from eight fly bodies were positive, of which seven were also positive in the transport fluid.

The seven *polA* and two *tp0574* sequences were identical to published orthologs from other *T. pallidum* subspecies (reference sequences in GenBank: U57757 [*polA*], M88769 [*tp0574*]). There was no variation at these two loci in all tested fly specimens.

In contrast, in the *tp0548* gene (nt 40 to 347, GenBank: CP000805; Fig. S3) variation was found among various *T. pallidum* subspecies and strains, which were included as reference, as well as among sequence data obtained from the fly samples. Among the 308 investigated sites, 66 were variable, of which 24 were parsimony-informative (including gaps; Fig. S3). Phylogenetic tree reconstructions based on *tp0548* sequence data (Fig. 1) show that the *T. pallidum* sequences obtained from fly specimens were diverse. This does not change even when a neighbor-joining tree is constructed (Fig. S4). Most sequences obtained from flies cluster with human reference TPE strains (str. Gauthier and str. CDC2), as well as str. F-B, and the sequences derived from *T. pallidum*-infected baboons at LMNP ($n = 28$, sample material originates from a study published elsewhere (Knauf et al., 2012)). Interestingly, 19 of the fly specimens (TNP $n = 9$, LMNP $n = 10$) exhibit a *tp0548* sequence that is identical to that obtained from samples of infected baboons as well as the type "j" sequence that was obtained from a human patient with genital ulceration and infected with a strain closely related to *T. pallidum* subsp. *endemicum* (Grange et al., 2013) (unpublished data; Smajs et al.). Four additional sequences (three haplotypes) obtained from fly specimens, all from LMNP, also cluster within the TPE-containing clade, each differing from the LMNP baboon strain sequence in just a single nucleotide.

tp0548 sequence data were also evaluated on the basis of the enhanced typing system (Marra et al., 2010), which uses only the sequence region between nt 130 and 215 (with reference to GenBank: CP000805) (Marra et al., 2010). As a result and with regard to the currently described subtypes (Read et al., 2016), the subtype distribution within the fly samples encompasses 23 sequences of dominating type "j", two sequences of subtype "a" and a yet undescribed subtype "o" ($n = 1$; Fig. S5, Table S2).

Fly species varied between the national parks (Fig. 2). Two fly species, *Chrysomya putoria* and *Musca sorbens*, seem to be disproportionally involved in the acquisition of *T. pallidum*. A full dataset of all fly specimens is presented in the Supplementary material.

4. Discussion

Our results indicate that *T. pallidum* DNA can be frequently detected on wild caught flies, at least in areas of high prevalence of *T. pallidum* infection in primates. The molecular confirmation of *T. pallidum* DNA

Table 1
Summary of genetic analysis in 207 fly specimens. Qualitative PCR results obtained from DNA extracted from individual fly bodies as well as transport fluids. "Overall" refers to the number of specimens tested positive in at least one of the gene targets. For details, see Supplementary material. Body = Fly body, TF = Transport fluid (RNA-Later Solution).

		polA	*tp0548*		*tp0574* (*tp47*)		Overall
		Body	Body	TF	Body	TF	Body + TF
TNP	n tested	88	88	8	40	5	88
	n positive (%)	0 (0%)	10 (11·4%)	4 (50·0%)	6 (15·0%)	4 (80·0%)	15 (17·0%)
LMNP	n tested	118	118	11	46	6	119
	n positive (%)	6 (5·1%)	17 (14·4%)	6 (54·5%)	10 (21·7%)	4 (66·7%)	28 (23·5%)

isolated from flies in this study, together with the recent detection of treponemal DNA on fly specimens from a yaws-endemic area of Papua New Guinea (unpublished data; Mitja et al.), supports the hypothesis of vector carriage of treponematoses, including yaws.

Variation in frequency of loci tested positive in fly specimens (Table 2) is most likely associated with a low copy number of treponemes in fly specimens as well as differences of sensitivity of the different PCR assays. However, it must be noted that also the amount of template DNA might have influenced the results since we have used 2 instead of 4 μl template DNA in the *polA* PCR. All PCRs applied in this study have been published elsewhere and are frequently used for the detection of *T. pallidum* in humans. The identity of the *tp0548* sequence isolated from olive baboons and the recovery of the same sequence from multiple fly specimens from Tanzania strongly suggests that flies often come into contact with the spirochete on NHPs. Our data are limited in that we did not attempt to isolate viable *T. pallidum* from the flies, which would be necessary to investigate the contagiousness of insects that carry the bacterium or the organism's ubiquity in the environment. *T. pallidum* is still not cultivable by standard microbiological techniques. Furthermore, based on the three loci that were amplified, it is not possible to identify the sources of contact with the bacteria, although it is most likely that, within the national park, baboon lesions are the major target for necrophagous flies near baboon groups. Our findings are consistent with earlier studies that demonstrated the presence (Kumm et al., 1935; Kumm, 1935b), and fly-associated transmission of treponemes (Lamborn, 1936; Castellani, 1907; Kumm and Turner, 1936; Satchell and Harrison, 1953; Thomson and Lamborn, 1934).

The finding that four additional *tp0548* sequences (three haplotypes) obtained from fly specimens, all from LMNP and which cluster within the TPE-containing clade, differ from the LMNP baboon strain sequence in just a single nucleotide, is likely caused by intra-strain variation and therefore does not argue for significant separation from the LMNP simian strain isolate. In contrast, three other fly-derived sequences (two from TNP and one from LMNP) cluster with human TPA reference strains (str. Nichols, str. Dallas 1, str. Chicago, and str. Seattle 81-4) as well as a TPEN str. Bosnia A. These findings may be explained either by the fact that different simian strains are present within the two baboon populations or that some flies had contact with human TPA strains circulating in villages surrounding the national parks. Laboratory contamination with TPA DNA is highly unlikely (Supplementary material).

Geographically, fly specimens with sequences that differ from the dominating *tp0548* sequence (Fig. 1), were collected in the northeastern zone of TNP (Fig. S1: A. TNP sampling sites 2 and 4) and in the north (Fig. S1: B. LMNP sampling sites 2, 6, and 7) and south (Fig. S1: B. LMNP sampling site 12) of LMNP. The finding that two of the three *tp0548* subtypes ("j" and "a") were recovered multiple times as well as at least two independent PCRs generated the same sequence of the yet undescribed subtype "o" provides evidence that these subtypes are not associated with sequencing error. Technically, *tp0548* sequencing errors were minimized by using proof-reading polymerases and by repeating PCR amplification and sequencing.

Table 2
Number of loci tested for *T. pallidum* per fly sample. TNP and LMNP samples are combined. Details can be found in Table 1 and Table S2.

	3 loci tested	2 loci tested (*polA* and *tp0548*)
n samples tested (total n = 207)	86	121
3 loci positive/sample	3	n/a
2 loci positive/sample	6	1
1 locus positive/sample	16	17

A substantial number of samples with negative test results for the fly's body DNA extract became positive for *T. pallidum* DNA when the corresponding transport fluid was analyzed. This may be explained by two different mechanisms. First, the biology of necrophagous flies forces the insects to land directly on infectious lesions and treponemes may stick to the fly's exoskeleton. Second, *Treponema*-containing lesion exudate from the fly's esophagous may have been vomited into the transport medium when numbed fly specimens were transferred from the trapping chamber into the reaction tubes. Presumably a combination of both happened. However, since many samples were positive only when the transport fluid was analyzed, this strongly suggests that the treponemes are located mainly on the outside of the exoskeleton. If the esophagous were the major source of contamination of the transport fluid, one would expect that more DNA samples extracted from the fly bodies would be positive.

The variation of fly species composition between the national parks (Fig. 2) reflects the different ecologies of the two sampling areas. While TNP is dominated by dry grassland and is heavily influenced by the Masai people herding their livestock in the surrounding area (Kiffner et al., 2015), LMNP represents a more isolated area, bordered on one side by Lake Manyara and on the other side by the Great East African Rift. At TNP, more livestock-associated flies were found (e.g., *Stomoxys* spp.), whereas LMNP seems to harbor a more primate/wildlife/wetland-dominated species composition (e.g., *Musca* spp.) (Fig. 2). In past studies, *M. sorbens* and *M. domestica* were demonstrated to transmit treponemes under experimental conditions (Lamborn, 1936; Castellani, 1907; Satchell and Harrison, 1953; Thomson and Lamborn, 1934; Robertson, 1908), although the molecular identification of the spirochetes in these old studies was not possible. Interestingly, *M. sorbens* was also one of the two fly species that were disproportionally often PCR positive for *T. pallidum* DNA in the two investigated study areas in Tanzania.

Though the risk for primates to acquire treponemal infection through fly transmission was not examined here, necrophagous flies may play a role as a potential vector for transmission. This is particularly relevant in areas where NHPs have a high prevalence of moist lesions that contain *T. pallidum* (Knauf et al., 2012). It has been shown that motile spirochetes are demonstrable in the esophagus of flies for up to 7 h following consumption of infectious lesion material (Kumm et al., 1935). Although the spirochetes in those studies were not definitively shown to be *T. pallidum*, this would provide a reasonable window of time for transmission in a setting where host abundance is high. While there are no recent reports of human yaws cases in East Africa, the area is known to have been endemic for yaws in the past and the current level of surveillance for yaws is unknown. Despite this uncertainty, if fly carriage of viable *T. pallidum* bacteria is confirmed, in combination with a possible nonhuman reservoir, flies could possibly facilitate the (inter-species) transmission of *T. pallidum*, which may result in new index cases even if yaws is eradicated in humans. This

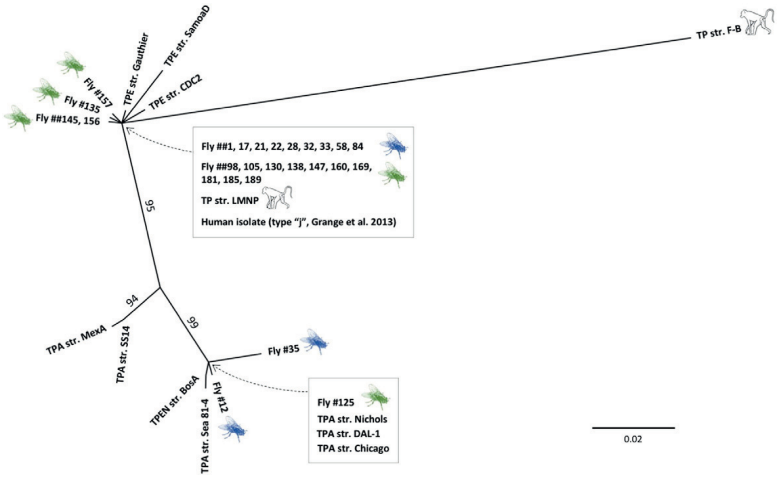


Fig. 1. Unrooted maximum parsimony tree based on *tp0548* sequence data of human TPA, TPE, and TPEN strains as well as orthologs of *T. pallidum* from LMNP baboons and fly isolates from TNP and LMNP. The tree includes *tp0548* sequences from 26 flies and 28 baboons plus TPA, TPE, and TPEN reference strains as indicated in the Materials and Methods. The type "j" sequence (Grange et al., 2013), which is identical to the baboon derived sequence, has been re-amplified and sequenced to cover a greater part of the previously published *tp0548* sequence. It should be noted, however, that despite the *tp0548* identity, other gene sequences differ from the TP str. LMNP sequence (unpublished data). The tree is based on 308 sites of which 24 are parsimony-informative. Gaps were coded as fifth character and 1000 bootstrap replicates were performed. Bootstrap values greater than 90% are displayed at respective nodes. Genbank accession numbers are provided in the Materials and Methods and Table S2, as well as the corresponding sequence data alignment, which can be seen in Fig. S3. TPA = *T. p. pallidum*, TPE = *T. p. pertenue*, TPEN = *T. p. endemicum*, str. = strain, F-B = Fribourg-Blanc, TNP = Tarangire National Park; LMNP = Lake Manyara National Park; green symbol: LMNP, blue: TNP. The bar refers to substitutions per site. Fly image source: http://www.oldskoolman.de/bilder/plog-content/images/freigestellte-bilder/natur-tiere/fliege-mit-ruessel.jpg (modified).

underlines the importance of continued intense surveillance after mass drug administration to eradicate human yaws, particularly in areas where NHPs are infected with treponemes, e.g. Cote d'Ivoire, Cameroon, Republic of Congo (Knauf et al., 2013).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ebiom.2016.07.033.

Funding Sources

This work was supported, in part, by R01AI42143 from the National Institutes of Health (to SAL) as well as partly by grants of the German Research Foundation (DFG): KN1097/3-1 and KN1097/4-1 (to SK), RO3055/2-1 (to CR), as well as ZI548/5-1 (to DZ).

The findings and conclusion in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Disclosure of Potential Conflict of Interest

The authors have nothing to declare.

Author Contributions

The study was designed by SK, JR, ISC, HL, CR, and SAL. DNA extraction and PCRs (*polA*, *tp47*, *tp0548*, and *ITS-2*) were performed at the laboratory of the German Primate Center (SK, SL, CS, CR). Additional PCR

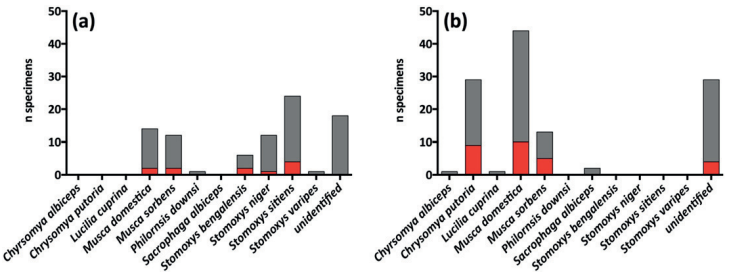


Fig. 2. Fly species composition and occurrence of *T. pallidum* DNA on flies caught at (a) TNP (n = 88) and (b) LMNP (n = 119). Fly species identification is based on the output from BLAST search using *ITS-2* sequence data. Fly specimens that did not generate identifiable sequence data are classified as "unidentified". Red = *T. pallidum* positive, grey = *T. pallidum* negative.

testing in a sample subset was run at University of Washington (CG, SAL). Data analyses and manuscript preparation was done by SK, JR, OM, IAVL, ISC, EKB, JDK, RF, HL, DS, PG, DZ, CR, and SAL.

Acknowledgments

We thank Tanzania National Parks and Tanzania Wildlife Research Institute for continuous support during the study. LMNP and TNP Headquarters are acknowledged for their hospitality and field assistance. Franziska Aron is thanked for her assistance in molecular analysis of sample material. SK and HL thank Kristin Harper for fruitful scientific discussions.

References

Asiedu, K., Fitzpatrick, C., Jannin, J., 2014. Eradication of yaws: historical efforts and achieving WHO's 2020 target. PLoS Negl. Trop. Dis. 8 (9), e3016.
Barnard, C.C., 1952a. Yaws and flies; past and present opinions on the role of flies in the transmission of framboesia tropica. *J. Trop. Med. Hyg.* 55 (5), 100–114 (contd).
Barnard, C.C., 1952b. Yaws and flies; past and present opinions on the role of flies in the transmission of framboesia tropica. II. *J. Trop. Med. Hyg.* 55 (6), 135–141.
Castellani, A., 1907. Experimental investigations on framboesia tropica (yaws). J. Hyg. 7 (4), 558–569 (Cambridge).
Emerson, P.M., Bailey, R.L., Mahdi, O.S., Walraven, G.E., Lindsay, S.W., 2000. Transmission ecology of the fly *Musca sorbens*, a putative vector of trachoma. Trans. R. Soc. Trop. Med. Hyg. 94 (1), 28–32.
Emerson, P.M., Lindsay, S.W., Walraven, G.E., et al., 1999. Effect of fly control on trachoma and diarrhoea. Lancet 353 (9162), 1401–1403.
Fribourg-Blanc, A., Mollaret, H.H., 1969. Natural treponematosis of the African primate. Primates Med. 3 (0), 113–121.
Giacani, L., Lukehart, S.A., 2014. The endemic treponematoses. Clin. Microbiol. Rev. 27 (1), 89–115.
Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27 (2), 221–224.
Grange, P.A., Allix-Beguec, C., Chanal, J., et al., 2013. Molecular subtyping of *Treponema pallidum* in Paris, France. *Sex. Transm. Dis.* 40 (8), 641–644.
Gudger, E.W., 1910a. A second early note on the transmission of yaws by flies. Science 32 (827), 632–633.
Gudger, E.W., 1910b. An early note on flies as transmitters of disease. Science 31 (784), 31–32.
Gudger, E.W., 1911. Further early notes on the transmission by flies of the disease called yaws. Science 33 (846), 427–428.
Harper, K.N., Fyumagwa, R.D., Hoare, R., et al., 2012. *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS One 7 (12), e50882.
Kiffner, C., Wenner, C., LaViolet, A., Yeh, K., Kioko, J., 2015. From savannah to farmland: effects of land-use on mammal communities in the Tarangire-Manyara ecosystem, Tanzania. *Afr. J. Ecol.* 53 (2), 156–166.
Knauf, S., Batamuzi, E.K., Mlengeya, T., et al., 2012. *Treponema* infection associated with genital ulceration in wild baboons. Vet. Pathol. 49 (2), 292–303.
Knauf, S., Dahlmann, F., Batamuzi, E.K., Frischmann, S., Liu, H., 2015. Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Negl. Trop. Dis. 9 (3), e0003637.
Knauf, S., Liu, H., Harper, K.N., 2013. Treponemal infection in nonhuman primates as possible reservoir for human yaws. Emerg. Infect. Dis. 19 (12), 2058–2060.
Kumm, H.W., 1935a. The digestive mechanism of one of the west Indian 'eye gnats': *Hippelates pallipes* Loew. Ann. Trop. Med. Parasitol. 29 (3), 283–303.
Kumm, H.W., 1935b. The natural infection of *Hippelates pallipes* Loew with the spirochaetes of yaws. Trans. R. Soc. Trop. Med. Hyg. 29 (3), 265–272.
Kumm, H.W., Turner, T.B., 1936. The transmission of yaws from man to rabbits by an insect vector, *Hippelates pallipes* Loew. Am.J.Trop. Med. Hyg. s1–16 (3), 245–271.
Kumm, H.W., Turner, T.B., Peat, A.A., 1935. The duration of motility of the spirochaetes of yaws in a small West Indian fly—*Hippelates pallipes* Loew. Am.J.Trop. Med. Hyg. s1–15 (2), 209–223.
Lamborn, W.A., 1936. The experimental transmission to man of *Treponema pertenue* by the fly *Musca sorbens*, WD. *J. Trop. Med. Hyg.* 15 (39), 235–239.
Liu, H., Rodes, B., Chen, C.Y., Steiner, B., 2001. New tests for syphilis: rational design of a PCR method for detection of *Treponema pallidum* in clinical specimens using unique regions of the DNA polymerase I gene. J. Clin. Microbiol. 39 (5), 1941–1946.
Lukehart, S., 2008. Biology of Treponemes. In: Holmes, K.K., Sparling, P.F., Stamm, W.E., et al. (Eds.), Sexually Transmitted Diseases, 4 ed. Mc Graw Hill Medical, New York, pp. 647–659.
Marra, C., Sahi, S., Tantalo, L., et al., 2010. Enhanced molecular typing of *Treponema pallidum*: geographical distribution of strain types and association with neurosyphilis. J. Infect. Dis. 202 (9), 1380–1388.
Marra, C.M., Tantalo, L.C., Sahi, S.K., Dunaway, S.B., Lukehart, S.A., 2016. Reduced Treponema pallidum-specific opsonic antibody activity in HIV-infected patients with syphilis. J. Infect. Dis. 213 (8), 1348–1354.
Mitja, O., Asiedu, K., Mabey, D., 2013. Yaws. Lancet 381 (9868), 763–773.
Read, P., Tagg, K.A., Jeoffreys, N., Guy, R.J., Gilbert, G.L., Donovan, B., 2016. *Treponema pallidum* strain-types and association with macrolide resistance in Sydney, Australia: new *tp0548* types identified. J. Clin. Microbiol.
Robertson, A., 1908. Flies as carriers of contagion in yaws (framboesia tropica). J. Trop. Med. 213.
Satchell, G.H., Harrison, R.A., 1953. Experimental observations on the possibility of transmission of yaws by wound-feeding Diptera, in Western Samoa. Trans. R. Soc. Trop. Med. Hyg. 47 (2), 148–153.
Sena, A.C., White, B.L., Sparling, P.F., 2010. Novel *Treponema pallidum* serologic tests: a paradigm shift in syphilis screening for the 21st century. Clin. Infect. Dis. 51 (6), 700–708.
Smith, J.L., David, N.J., Indgin, S., et al., 1971. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br. J. Vener. Dis. 47 (4), 226–251.
Song, Z., Wang, X., Liang, G., 2008. Species identification of some common necrophagous flies in Guangdong province, southern China based on the rDNA internal transcribed spacer 2 (ITS2). Forensic Sci. Int. 175 (1), 17–22.
Thomson, J.G., Lamborn, W.A., 1934. Mechanical transmission of trypanosomiasis, leishmaniasis, and yaws through the agency of non-biting *Haematophagous* flies: (preliminary note on experiments). Br. Med. J. 2 (3845), 506–509.
Zobaníková, M., Strouhal, M., Mikalova, L., et al., 2013. Whole genome sequence of the Treponema Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Negl. Trop. Dis. 7 (4), e2172.

7.4.*3 Haemophilus ducreyi* DNA is Detectable on the Skin of Asymptomatic Children, Flies and Fomites in Villages of Papua New Guinea

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Haemophilus ducreyi DNA is detectable on the skin of asymptomatic children, flies and fomites in villages of Papua New Guinea

Wendy Houinei[1☯], Charmie Godornes[2☯], August Kapa[3], Sascha Knauf[4], Eric Q. Mooring[5], Camila González-Beiras[6], Ronald Watup[3], Raymond Paru[3], Paul Advent[3], Sivauk Bieb[1], Sergi Sanz[7], Quique Bassat[7,8,9], Stanley M. Spinola[10], Sheila A. Lukehart[2,11☯], Oriol Mitjà[3,7,12☯*]

1 Disease Control Branch, National Department of Health, Port Moresby, Papua New Guinea, **2** Department of Medicine, University of Washington, Seattle, Washington, United States of America, **3** Lihir Medical Centre– International SOS, Newcrest Mining, Lihir Island, Papua New Guinea, **4** German Primate Center, Leibniz-Institute for Primate Research, Pathology Unit, Working Group Neglected Tropical Diseases, Göttingen, Germany, **5** Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States of America, **6** Global Public Health PhD Programme, Nova University of Lisbon, Lisbon, Portugal, **7** Barcelona Institute for Global Health (ISGlobal), Hospital Clinic-University of Barcelona, Barcelona, Spain, **8** Centro de Investigação em Saúde de Manhiça (CISM), Maputo, Mozambique, **9** Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain, **10** Departments of Microbiology and Immunology, Medicine, and Pathology and Laboratory Medicine Indiana University School of Medicine, Indianapolis, Indiana, United States of America, **11** Department of Global Health, University of Washington, Seattle, Washington, United States of America, **12** Division of Public Health, School of Medicine and Health Sciences, University of Papua New Guinea, Port Moresby, Papua New Guinea

☯ These authors contributed equally to this work.
* oriol.mitja@isglobal.org

OPEN ACCESS

Citation: Houinei W, Godornes C, Kapa A, Knauf S, Mooring EQ, González-Beiras C, et al. (2017) *Haemophilus ducreyi* DNA is detectable on the skin of asymptomatic children, flies and fomites in villages of Papua New Guinea. PLoS Negl Trop Dis 11(5): e0004958. https://doi.org/10.1371/journal.pntd.0004958

Editor: Patrick J. Lammie, Task Force for Child Survival and Developmentorce for Global Health, UNITED STATES

Received: December 25, 2015

Accepted: August 8, 2016

Published: May 10, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported in part by Newcrest Mining Company (to OM), by R01AI42143 from the National Institutes of Health (to SAL), by T32AI007535-16A1 from the National Institutes of Health (to EQM), and by the Michael von Clemm Traveling Fellowship (to EQM). The

Abstract

Background

Haemophilus ducreyi and *Treponema pallidum* subsp. *pertenue* are major causes of leg ulcers in children in Africa and the Pacific Region. We investigated the presence of DNA (PCR positivity) from these bacteria on asymptomatic people, flies, and household linens in an endemic setting.

Methodology/Principal findings

We performed a cross-sectional study in rural villages of Lihir Island, Papua New Guinea during a yaws elimination campaign. Participants were asymptomatic subjects recruited from households with cases of leg ulcers, and from households without cases of leg ulcers. We rubbed swabs on the intact skin of the leg of asymptomatic individuals, and collected flies and swabs of environmental surfaces. All specimens were tested by PCR for *H. ducreyi* and *T. p. pertenue* DNA. Of 78 asymptomatic participants that had an adequate specimen for DNA detection, *H. ducreyi*-PCR positivity was identified in 16 (21%) and *T. p. pertenue*-PCR positivity in 1 (1%). In subgroup analyses, *H. ducreyi*-PCR positivity did not differ in participants exposed or not exposed to a case of *H. ducreyi* ulcer in the household (24% vs 18%; p = 0.76). Of 17 cultures obtained from asymptomatic participants, 2 (12%) yielded a

funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: August Kapa, Ronald Watup, Raymond Paru, Paul Advent and Oriol Mitjà are employed by the Lihir Medical Centre, which is jointly funded by the government of Papua New Guinea and Newcrest Mining. The authors have declared that no competing interests exist.

definitive diagnosis of *H. ducreyi*, proving skin colonization. Of 10 flies tested, 9 (90%) had *H. ducreyi* DNA and 5 (50%) had *T. p. pertenue* DNA. Of 6 bed sheets sampled, 2 (33%) had *H. ducreyi* DNA and 1 (17%) had *T. p. pertenue* DNA.

Conclusions/Significance

This is the first time that *H. ducreyi* DNA and colonization has been demonstrated on the skin of asymptomatic children and that *H. ducreyi* DNA and *T. p. pertenue* DNA has been identified in flies and on fomites. The ubiquity of *H. ducreyi* in the environment is a contributing factor to the spread of the organism.

Author summary

Children in rural communities of tropical countries often suffer skin ulcers that are caused by the bacteria *Haemophilus ducreyi*–causative agent of chancroid- and *Treponema pallidum* subsp. *pertenue* -causative agent of yaws-. The currently recommended strategy for yaws eradication is one round of mass drug administration (MDA) with azithromycin. We attempted to find reasons for the limited impact of yaws MDA on the prevalence of *H. ducreyi* leg ulcers by examining potential sources of infection in healthy carriers, flies, and bed linen. *H. ducreyi* DNA was found in skin swabs from 20% of asymptomatic children, in 9/10 flies, and 3/6 bed sheets from the houses of children with ulcers. While *H. ducreyi* DNA has been detected in the genital tract of asymptomatic women without genital ulcers, this is the first report of such detection on the skin of asymptomatic individuals. Importantly, skin cultures obtained from two asymptomatic children yielded viable *H. ducreyi*, confirming colonization and a potential reservoir of infection. If confirmed to contain viable bacteria, flies and fomites may also contribute to the continued presence of this infection after mass treatment with azithromycin. Our findings provide evidence that persistence of *H. ducreyi* ulcers after antibiotic MDA is due to the ubiquity of the organism in the environment. Improved hygiene and additional strategies such as repeated rounds of MDA could be able to control such a reservoir.

Introduction

During recent yaws eradication studies, *Haemophilus ducreyi* was shown to be a major cause of chronic cutaneous ulcers in rural tropical-regions in the South Pacific islands and equatorial Africa [1–5]. Studies from yaws-endemic villages in Papua New Guinea (PNG), Vanuatu, and Ghana reported that 27% to 60% of all skin ulcers were caused by *H. ducreyi*, while 15% to 34% had detectable *Treponema pallidum* subsp. *pertenue* [1, 3–5]. Mixed infections with both pathogens were seen in 3% to 13% of skin ulcers. Unlike yaws, *H. ducreyi* infection appears to be restricted to the skin and does not result in inflammatory lesions of the bones [6]. The infectivity of *H. ducreyi* is high [7], and reinfections following clinical and microbiological cure of *H. ducreyi* genital ulcers are common [8].

A single oral dose of 30 mg/kg azithromycin is highly effective for treatment of yaws [9], while 15 mg/kg generally provides effective treatment and prophylaxis against *H. ducreyi* genital ulcer strains [10]. Cutaneous strains of *H. ducreyi* are also susceptible to macrolides in vitro [11]. In the context of yaws eradication, the use of mass drug administration (MDA) with oral

azithromycin might be expected to be effective for both yaws and *H. ducreyi*. However, in rural villages of Lihir Island, MDA drastically decreased the proportion of ulcers containing *T. p. pertenue* DNA, but had limited impact on *H. ducreyi* [12].

Both infections are thought to be exclusively transmitted through non-intact skin contact with infectious lesions. Nevertheless, the persistence of *H. ducreyi* skin ulcers after mass azithromycin treatment [12,13] raises the possibility that the bacteria may exist in a natural reservoir. If *H. ducreyi* adheres and survives on the healthy skin of asymptomatic carriers, this could enhance its persistence in the community and its transmission, because skin-colonizing bacteria might escape systemic azithromycin treatment and could infect the skin after a minor abrasion. Asymptomatic colonization of exposed skin may serve as a reservoir for transmission to family members, classmates, or playmates. Asymptomatic carriage of *H. ducreyi* was demonstrated by PCR in the genital mucosa of about 2% of sex workers in The Gambia [14], but no studies address whether either *H. ducreyi* or *T. p. pertenue* potentially colonize non-genital skin.

Another source of re-infection after MDA could be fomites, such as household linens, or insects, such as flies. Linens harboring bacteria could facilitate colonization of skin or allow the bacteria to gain access to new hosts with wounds or abrasions. Indirect transmission of yaws by non-biting flies has been suggested in the older literature on the basis that *Musca* spp. and *Hippelates* spp. flies fed on scrapings from yaws lesions produced infection in experimental animals [15–17]. Infected flies were shown to have motile spirochetes (not formally shown to be *T. p. pertenue*) in their esophageal diverticula that could be regurgitated and deposited on the skin or into wounds [15,18]. Nevertheless, there is no clear evidence of *H. ducreyi* or yaws transmission by insects or fomites.

The demonstration of bacterial colonization or carriage usually requires a positive culture as an indication of presence and multiplication of microorganisms. However, culture of *H. ducreyi* is expensive and technically difficult and *T. p. pertenue* cannot be grown in vitro. Highly sensitive PCR methods cannot determine whether the source of DNA is viable organisms, but identification of bacterial DNA on the skin or a body surface is useful for exploratory analyses and helps to formulate hypotheses that could lead to new experimental studies.

The primary objective of this study was to examine the skin of asymptomatic children for the presence of *H. ducreyi* and *T. p. pertenue* DNA. Secondary objectives were to demonstrate the presence of viable *H. ducreyi* in cultures from a subset of these children, to examine whether contacts of an ulcer case are more likely to carry bacterial DNA than persons from households without ulcer disease, and to identify potential environmental sources, such as linens and flies, for the presence of *H. ducreyi* and *T. p. pertenue* DNA.

Methods

Study setting and participants

During October 2014 and May 2015 we performed a cross sectional study examining the skin of asymptomatic children and environmental sources for the presence of *H. ducreyi* and *T. p. pertenue* DNA in villages of Lihir Island, PNG. All villages in Lihir Island had received mass azithromycin treatment for yaws elimination in May 2013 followed by active surveillance and treatment of ulcer cases and their contacts at 6-month intervals.

Children with skin ulcers were identified during active case surveillance during the yaws elimination campaign as previously reported [12], and these were designated as "index cases". Index cases with leg ulcers or their parents identified household relatives. We enrolled a convenience sample of asymptomatic children and young adults without ulcers from households of index cases. We also enrolled a convenience sample of asymptomatic subjects from households

without cases of leg ulcers in randomly selected villages. Villages were selected with probability proportional to size sampling.

Ethics statement

All participants, or their parents or guardians, provided written informed consent to be enrolled in this study. The protocol was approved by the National Medical Research Advisory Committee of the PNG National Department of Health (MRAC no. 17.01).

Procedures

To assess participants for the presence of bacterial DNA, we rubbed a sterile unmoistened dacron swab on intact skin of the anterior aspect of the lower legs over a 5- by 5-cm area. Swab specimens were placed into tubes containing 1 mL of lysis buffer (10mM Tris-HCl, 0.1M EDTA, and 0.5% SDS) to stabilize DNA prior to shipment for PCR testing. Cultures were attempted on a subset of swab specimens. A structured questionnaire was administered to all asymptomatic participants to collect household-level sociodemographic information and to assess health-related, hygiene-related, and other household-level hypothesized risk factors for intra-household transmission of skin pathogens.

To assess environmental sources, we collected flies from verandas and surrounding areas immediately outside the houses of patients with cutaneous ulcers, and we swabbed the surface of their bed linens to cover 5 separate areas of 10 by 10 cm. We considered fomite (bed linens) transmission to family members, because of the standard practice of bed sharing in PNG. Flies and swabs from the linens were placed in tubes containing Tris-EDTA-SDS buffer and frozen for transportation to the laboratory.

Laboratory assessment

The skin swab samples were sent to the University of Washington (Seattle, WA, USA) for PCR testing to detect *T. p. pertenue* and *H. ducreyi* DNA. As described [1], *T. pallidum* DNA was assayed by standard PCR for TP0548 and by TaqMan real time PCR for T47 (TP0574). The pertenue subspecies was confirmed by TprL PCR amplicon size. *H. ducreyi* DNA was assayed by standard PCR and by TaqMan PCR. All samples were tested by PCR for human beta-globin DNA, as a control for sample adequacy and DNA integrity.

We cultured a subset of swabs that were collected from ulcers of index cases and intact skin of asymptomatic participants for *H. ducreyi* on C-HgCh (Mueller Hinton agar with 5% chocolatized hemoglobin + 3 mg/ml vancomycin) culture plates in the field; these were transported in a candle jar to the laboratory within 4 hours and transferred to an incubator with BD GasPak system and incubated at 33˚C for 48 h. If cultures grew organisms presumptively identified as *H. ducreyi* (small yellow-gray colonies whose gram stain showed small gram negative rods [6]), we transferred the plate grown colonies to a tube containing Assay Assure medium for transport, DNA extraction, and definitive species identification by PCR [19].

The flies and the Tris-EDTA-SDS buffer used for transport of flies were tested separately. Organisms passively adhering to the outside of the fly would likely be reflected in the DNA extracted from the buffer, while regurgitating pathogens would be inside the gut. Fly DNA was extracted using Gen-ial All-Tissue DNA Extraction kit. Briefly, the flies were completely disrupted in Lysis Buffer #1 using a pestle. Proteinase K, Lysis Buffer #2, and 0.01M DTT were added to the individual tubes and incubated for 48 hours at 37˚C. Following centrifugation at 16,000 g, the supernatant was treated with chloroform and the aqueous phase was washed with Lysis Buffer #3. DNA was precipitated using glycogen and isopropanol. We tested each fly and transport buffer specimen using the standard PCR and TaqMan assays mentioned above. In

the TaqMan assays, DNA from each fly was tested in 21 replicates (7 from the transport buffer and 14 from the extracted fly DNA). Multiple negative control PCR reactions were run in each assay, and were uniformly negative.

Statistical analyses

The primary outcome measure was PCR-positivity rate for *T. p. pertenue* or *H. ducreyi* in asymptomatic participants. We compared PCR-positivity rates of contacts of an *H. ducreyi* ulcer case to those of participants from a household without an *H. ducreyi* skin ulcer case with Fisher exact tests. We reported odds ratios with 95% CIs from univariate logistic regression to compare the living conditions of PCR positive and negative participants. We accounted for clustering by household among exposed subjects using a penalized maximum likelihood estimation method in the regression model [20]. All analyses were done with Stata version 14.0.

Results

Household visits were completed for 21 patients who had ulcerative lesions and agreed to participate. Of 21 patients with leg ulcers, 12 (57%) had detectable *H. ducreyi* DNA and 1 (5%) had detectable *T. p. pertenue* DNA in their ulcers; neither pathogen was detected in 8 (38%) of the index cases. Of 7 cultures obtained from leg ulcers, 2 (29%) yielded no growth and 5 (71%) yielded small gram negative rods or coccobacilli. Of the 5 gram negative organisms, 2 were confirmed by PCR to be *H. ducreyi*.

We identified 71 asymptomatic subjects exposed to a skin ulcer case, and these were all enrolled (Fig 1). We also enrolled 20 asymptomatic subjects from households without an ulcer case. Of 91 asymptomatic participants tested, 12 (17%) in the group exposed to an ulcer case, and 1 (5%) in the group not exposed to an ulcer case in the household had negative beta-globin and bacterial-DNA amplification results, and were excluded from subsequent comparative analyses, so a total of 78 asymptomatic participants were evaluated.

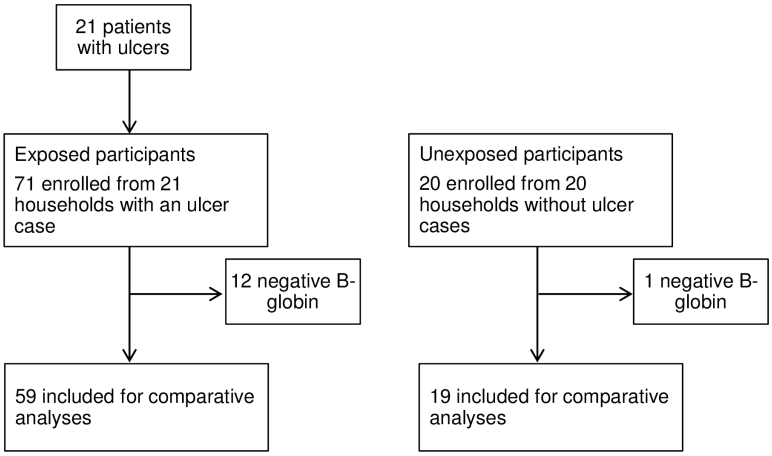


Fig 1. Trial profile.

https://doi.org/10.1371/journal.pntd.0004958.g001

Table 1. *H. ducreyi* skin PCR positivity rates.

	Group 1	Group 2	Group3	Odds ratio	Odds ratio
	Exposed to HD positive case (n = 34)	Exposed to HD negative case (n = 25)	Not exposed to ulcer case in the household (n = 19)	(95%CI) Group 1 vs. Group 2	(95%CI) Group 1 vs. Group 3
H. ducreyi DNA amplified	8 (24%)	4 (16%)	4 (21%)	1.62 (0.42–6.18)	1.15 (0.30–4.49)

HD, *H. ducreyi*. Data are n (%).

https://doi.org/10.1371/journal.pntd.0004958.t001

Overall, 16 (21%, 95%CI 12–31) of 78 asymptomatic participants had *H. ducreyi* bacterial DNA on the skin, and 1 (1%) had *T. p. pertenue* DNA. Of 17 cultures obtained from asymptomatic contacts, 8 (47%) yielded no growth and 9 (53%) yielded small gram negative rods or coccobacilli. Six of the gram negative organisms were tested by PCR and 2 were confirmed to be *H. ducreyi*.

By exposure status, 8 (24%) of 34 asymptomatic participants in the group exposed to an *H. ducreyi* positive ulcer, and 8 (18%) of 44 participants not exposed to an *H. ducreyi* positive ulcer were *H. ducreyi* DNA positive (OR 1.38; 95%CI 0.46–4.17; p = 0.76) (Table 1, analyses by subgroups—exposed to HD negative case and not exposed to an ulcer case). The single asymptomatic participant with detectable *T. p. pertenue* DNA was identified in a household of an index case with *H. ducreyi* mono-infection.

When we compared individual household-level socio-demographic factors, the living conditions of asymptomatic PCR-positive participants and PCR-negative participants were very similar (Table 2). Mean age (SD) of PCR-positive participants was 7.7 (3.7) years, and of PCR-negative participants was 10.3 (8.0) years. PCR-positive participants were significantly more likely to change bed linens less than once per week than participants without detectable bacterial DNA.

Flies (n = 10) were caught outside 10 houses of patients with ulcers; 9 (90%) of 10 flies had *H. ducreyi* DNA and 5 (50%) had amplifiable *T. p. pertenue* DNA (Table 3). For 2 flies,

Table 2. Socioeconomic, environmental and behavioral factors that may predispose to PCR positivity with *H. ducreyi*.

Variable	PCR-negative (n = 61)	PCR-positive (n = 17)	OR (95%CI)	P value
Exposed to an ulcer case	46 (75%)	13 (76%)	1.06 (0.30–3.75)	0.819
Sex (male)	31 (51%)	8 (47%)	1.15 (0.41–3.28)	0.788
Age (mean, SD)[1]	10.3 (8.0)	7.7 (3.7)	0.95 (0.86–1.04)	0.280
No fixed indoor shower or bath[2]	60 (98%)	17 (100%)	0.93 (0.03–27.10)	0.967
No indoor toilet[2]	48 (79%)	17 (100%)	9.61 (0.54–169.69)	0.122
Sand or mud in the floor[2]	54 (89%)	17 (100%)	5.13 (0.27–95.88)	0.274
Traditional (temporary) housing[2]	24 (39%)	9 (53%)	1.71 (0.60–4.88)	0.319
Livestock in the house[2]	54 (89%)	17 (100%)	4.86 (0.26–89.47)	0.288
Walking bare-foot[2]	53 (87%)	17 (100%)	5.65 (0.31–103.16)	0.243
Frequency of changing bed-linen less 1 per week[3]	19 (31%)	10 (59%)	3.03 (1.03–8.87)	0.043
Frequency of bathing less 1 per day[4]	32 (52%)	9 (53%)	1.02 (0.35–2.99)	0.811
Children share towel [2]	35 (57%)	10 (59%)	1.06 (0.36–3.10)	0.919

Univariate model adjusted for exposure to an ulcer case

[1] Odds ratio per unit increase,

[2] Odds ratio for yes vs. no,

[3] Odds ratio for less 1 per week vs. more than 1 per week

[4] Odds ratio for less 1 per day vs. more than 1 per day

https://doi.org/10.1371/journal.pntd.0004958.t002

Table 3. Presence of *H. ducreyi* and *T. p. pertenue* DNA on flies.

Setting by specimen	*H. ducreyi* (n = 10)	*T. p. pertenue* (n = 10)
All specimens	9 (90%)	5 (50%)
Type of PCR		
Standard	5 (50%)	4 (40%)
Taqman	9 (90%)	5 (50%)
Specimen		
Whole fly	7 (70%)	5 (50%)
Lysis buffer	6 (60%)	0 (0%)

Data are n (%)

https://doi.org/10.1371/journal.pntd.0004958.t003

H. ducreyi was detected reproducibly by standard PCR and in 21/21 replicates by real-time PCR, suggesting relatively abundant *H. ducreyi* on these flies. Lower amounts of *H. ducreyi* DNA (1–3 positive tests of 21 replicates) were found in 7 other flies. We detected *T. p. pertenue* positive results in standard and real-time PCR for 4 flies and only in real-time PCR for the 5th fly. All of the *T. p. pertenue*-positive flies also carried detectable *H. ducreyi* DNA. Three of the 5 flies with detectable *T. p. pertenue* DNA were collected from a community where active yaws prevalence remained high after MDA. The buffer used to transport the flies yielded amplifiable *H. ducreyi* DNA in 6 (60%) of 10 specimens, but none yielded amplifiable *T. p. pertenue* DNA, perhaps suggesting different location of the two bacteria in flies (i.e. *T. p. pertenue* as a regurgitating pathogen in the gut), or dilution effect related to differential bacterial load of *H. ducreyi* and *T. p. pertenue* carried by affected flies.

Environmental contamination was observed in 3 (50%) of 6 bed sheets sampled from 6 households with index cases; 2 sheets had detectable *H. ducreyi* DNA and 1 had detectable *T. p. pertenue* DNA. One of the 2 sheets with *H. ducreyi* and the *T. p. pertenue* sheet were from households with *H. ducreyi* positive index cases, while the other *H. ducreyi* sheet was from a household with a PCR-negative ulcer case.

Discussion

We used PCR techniques to demonstrate that 20% of asymptomatic children living in *H. ducreyi*-endemic communities had detectable *H. ducreyi* DNA on the skin. These children had no evidence of cutaneous ulceration, and no other symptoms or signs of infection. This is the first report of asymptomatic carriage of *H. ducreyi* DNA on skin. Notably, people from households without an ulcer case had positivity rates that were similar to people with an *H. ducreyi* ulcer case in the household. Hence, a member of a household without a case should still have been exposed to an infectious case in the community, such as in school.

To detect possible *H. ducreyi* colonization, we chose PCR as the primary diagnostic method instead of culture because of our limited capability of performing a large number of cultures for this fastidious organism. Compared to PCR, single plate culture systems are only ~ 50% sensitive for detecting *H. ducreyi* in patients with chancroid [21,22], which is why PCR is the preferred diagnostic test for chancroid. However, PCR can detect nonviable bacteria or contaminating DNA; hence, a swab positive for DNA does not give definitive evidence of colonization. In our attempts to culture the organism from a subset of samples, we presumptively identified *H. ducreyi* in approximately half of the cultures, and two of them were positive by PCR for definitive species identification. Further studies using optimal culturing techniques are required to elucidate the true extent of the biological reservoir of *H. ducreyi* on the skin in this population.

Our study was significantly limited in its ability to assess asymptomatic carriage of *T. p. pertenue* because the mass azithromycin treatment conducted earlier in these communities was so effective in reducing the prevalence of yaws ulcers [12]. Since only few contacts of one yaws case were included in this study, our findings should be interpreted with caution. The viability of *T. p. pertenue* on skin surfaces or sheets is unknown. This organism is thought to be highly fragile outside of a susceptible host and would not be expected to survive on fomites.

We previously reported that approximately 2% of the total population in Lihir Island and 7% of the children aged 5–15 years had ulcers containing *H. ducreyi* DNA. Although *H. ducreyi* cutaneous ulcer strains might be exclusively transmitted through contact of wounds with infectious lesions, it seems unlikely that this mode of transmission could account for the high prevalence of infection or persistence after MDA using azithromycin. Our findings suggest that *H. ducreyi* survives on healthy non-genital skin where even minor trauma could initiate infection. In human volunteers, placement of up to 10^6 colony forming units (CFU) of the genital ulcer strain 35000HP on intact skin fails to cause infection; however, as few as 1 CFU delivered by a 2 mm puncture wound is sufficient to initiate infection [7,23]. Other factors may impact the initiation of disease in carriers of *H. ducreyi*. For example, recent human inoculation experiments raise the possibility that the composition of the skin microbiome could influence host susceptibility to *H. ducreyi* skin infection and ulcer formation [24].

Our findings suggest a possible role for flies in the transmission of *H. ducreyi* and yaws under natural conditions. Flies have also been discussed as possible vectors of the related treponematoes, pinta and bejel [25,26] and *T. pallidum* DNA has been detected on flies related to treponemal infections in non-human primates [27], but direct evidence of transmission is lacking. Although we did not test whether flies carry viable bacteria, a high proportion of flies collected had detectable *H. ducreyi* DNA and half carried both *H. ducreyi* and *T. p. pertenue* DNA. Further study is necessary to unravel whether bacterial DNA in the flies reflects the ubiquity of the organism in the environment or carriage of live organisms.

The minimal effect of MDA with azithromycin on the prevalence of skin ulcers due to *H. ducreyi*, compared to the profound effect on the prevalence of yaws, is puzzling. *H. ducreyi* colonization of asymptomatic villagers, flies, and fomites could explain the continued presence of this infection after MDA. If azithromycin does not reach the outer surface of skin, it may not interrupt colonization, and these sources may perpetuate the infection in the community. Given the prolonged prophylactic effect of azithromycin against experimental *H. ducreyi* infection, it is plausible that repeated mass treatment can confer a prophylactic effect to the population for long enough to clear the asymptomatic reservoir [10].

New strategies to control *H. ducreyi* along with yaws need to be explored. Syndromic care for ulcers using azithromycin and multiple rounds of MDA could be included in future iterations of the Morges Strategy for yaws eradication. Indeed mathematical modeling has shown the value of multiple rounds of MDA to reduce *T. p. pertenue* infection [28], and this could, in parallel, reduce *H. ducreyi* infection and potentially skin carriage. In addition, skin hygiene and effective wound management using non-adherent dressings must be emphasized; given the potential carriage of *H. ducreyi* and *T. p. pertenue* by flies, covering ulcers may also help to prevent transmission.

Acknowledgments

We would like to thank the study participants and their parents or legal guardians and the field teams of the Lihir Medical Centre.

Author Contributions

Conceptualization: SMS SAL OM.

Formal analysis: SS OM.

Funding acquisition: SAL OM.

Investigation: WH CG AK EQM RW RP PA.

Methodology: WH CG SMS SAL OM.

Resources: CG SAL.

Supervision: OM.

Writing – original draft: WH SMS SAL OM.

Writing – review & editing: EQM CGB SB QB SK.

References

1. Mitjà O, Lukehart SA, Pokowas G, et al. *Haemophilus ducreyi* as a cause of skin ulcers in children from a yaws-endemic area of Papua New Guinea: a prospective cohort study. Lancet Glob Health. 2014; 2:235–241.
2. Marks M, Chi KH, Vahi V, et al. *Haemophilus ducreyi* associated with skin ulcers among children, Solomon Islands. Emerg Infect Dis. 2014; 20:1705–1707. https://doi.org/10.3201/eid2010.140573 PMID: 25271477
3. Chi KH, Danavall D, Taleo F, et al. Molecular differentiation of *Treponema pallidum* subspecies in skin ulceration clinically suspected as yaws in Vanuatu using real-time multiplex PCR and serological methods. Am J Trop Med Hyg. 2015; 92:134–138. https://doi.org/10.4269/ajtmh.14-0459 PMID: 25404075
4. Fegan D, Glennon MJ, Kool J, Taleo F. Tropical leg ulcers in children: more than yaws. Trop Doct 2015: pii: 0049475515599326.
5. Gonzalez-Beiras C, Marks M, Chen C, et al. Epidemiology of *Haemophilus ducreyi* infection: from genital ulcer in adults to skin ulcer in children. Emerg Infect Dis. 2016; 22.
6. Spinola SM, Ballard RC. Chancroid. In: Morse SA, Holmes KK, and Ballard RC, eds. Atlas of sexually transmitted diseases and AIDS, 4th edn. Philadelphia: Saunders, 2010.
7. Janowicz DM, Ofner S, Katz BP, Spinola SM. Experimental infection of human volunteers with *Haemophilus ducreyi*: fifteen years of clinical data and experience. J Infect Dis, 2009; 199: 1671–1679. https://doi.org/10.1086/598966 PMID: 19432549
8. Ballard RC, Ye H, Matta A, Dangor Y, Radebe F. Treatment of chancroid with azithromycin. Int J STD AIDS. 1996; 7:9–12.
9. Mitjà O, Hays R, Ipai A, et al. Single-dose azithromycin versus benzathine benzylpenicillin for treatment of yaws in children in Papua New Guinea: an open-label, non-inferiority, randomised trial. Lancet. 2012; 379:342–347. https://doi.org/10.1016/S0140-6736(11)61624-3 PMID: 22240407
10. Thornton AC, O'Mara EM Jr, Sorensen SJ et al. Prevention of experimental *Haemophilus ducreyi* infection: a randomized, controlled clinical trial. J Infect Dis. 1998; 177:1608–1613. PMID: 9607840
11. Gangaiah D, Webb KM, Humphreys TL, et al. *Haemophilus ducreyi* cutaneous ulcer strains are nearly identical to class I genital ulcer strains. PLoS Negl Trop Dis. 2015; 9:e0003918. https://doi.org/10.1371/journal.pntd.0003918 PMID: 26147869
12. Mitjà O, Houinei W, Moses P, et al. Mass treatment with single-dose azithromycin for yaws. N Engl J Med. 2015; 372:703–710. https://doi.org/10.1056/NEJMoa1408586 PMID: 25693010
13. Ghinai R, El-Duah P, Chi KH, et al. A cross-sectional study of 'yaws' in districts of Ghana which have previously undertaken azithromycin mass drug administration for trachoma control. PLoS Negl Trop Dis. 2010; 9:e0003496.
14. Hawkes S, West B, Wilson S, Whittle H, Mabey D. Asymptomatic carriage of *Haemophilus ducreyi* confirmed by the polymerase chain reaction. Genitourin Med. 1995; 71:224–227. PMID: 7590712
15. Kumm HW, Turner TB. The transmission of yaws from man to rabbits by an insect vector, Hippelates Pallipes. Am J Trop Med Hyg. 1936; 16:245–271.
16. Castellani A. Experimental investigations on Framboesia Tropica (Yaws). J hygiene. 1907; 7:558–569.

17. Lamborn WA. The experimental transmission to man of *Treponema pertenue* by the fly Musca sorbens. J Trop Med Hyg. 1936; 15:235–239.

18. Satchell GH, Harrison RA. Experimental observations on the possibility of transmission of yaws by wound-feeding Diptera in Western Samoa. Trans R Soc Trop Med Hyg. 1953; 47:148–153. PMID: 13077710

19. Mackay IM, Harnett G, Jeoffreys N, Bastian I, Sriprakash KS, Siebert D, Sloots T. Detection and Discrimination of Herpes Simplex Viruses, *Haemophilus ducreyi*, *Treponema pallidum*, and *Calymmatobacterium (Klebsiella) granulomatis* from Genital Ulcers. Clin Infect Dis. 2006; 42:1431–1438. https://doi.org/10.1086/503424 PMID: 16619156

20. Varin C, Reid NM, Firth D. An overview of composite likelihood methods. Statistica Clinica. 2011; 21:5–42.

21. Morse SA, Trees DL, Htun Y, et al. Comparison of clinical diagnosis and standard laboratory and molecular methods for the diagnosis of genital ulcer disease in Lesotho: association with human immunodeficiency virus infection. J Infect Dis. 1997; 175:583–589. PMID: 9041329

22. Nsanze H, Pummer FA, Maggwa ABN, Dylewski J, Piot P, Ronald AR. Comparison of media for primary isolation of Haemophilus ducreyi. Sex Transm Dis. 1984; 11:6–9. PMID: 6608798

23. Spinola SM, Wild LM, Apicella MA, Gaspari AA, Campagnari AA. Experimental human infection with *Haemophilus ducreyi*. J Infect Dis. 1994; 169:1146–1150. PMID: 8169411

24. van Rensburg JJ, Lin H, Gao X, et al. The human skin microbiome associates with the outcome of and is influenced by bacterial infection. MBio. 2015; 6:e1315–1315.

25. Blanco L. La noción del contagio y la idea del vector en el mal de pinto. Medicina. 1940; 20:162.

26. Perine PL, Hopkins DR, Niemel PLA, St John RK, Causse G, Antal GM. Handbook of endemic treponematoses: yaws, endemics syphilis, and pinta. Geneva: World Health Organization, 1984

27. Knauf S, Raphael J, Mitjà O, et al. Isolation of Treponema DNA from Necrophagous Flies in a Natural Ecosystem. EBioMedicine. 2016; 11:85–90. https://doi.org/10.1016/j.ebiom.2016.07.033 PMID: 27488881

28. Marks M, Mitja O, Fitzpatrick C et al. Mathematical Modeling of Programmatic Requirements for Yaws Eradication. Emerg Infect Dis. 2017; 23:22–28. https://doi.org/10.3201/eid2301.160487 PMID: 27983500

7.4.4 Mating Avoidance in Female Olive Baboon (*Papio anubis*) Infected by *Treponema pallidum*

ECOLOGY

Mating avoidance in female olive baboons (*Papio anubis*) infected by *Treponema pallidum*

F. M. D. Paciência[1,2]*, J. Rushmore[3], I. S. Chuma[4], I. F. Lipende[4], D. Caillaud[5], S. Knauf[2,6†], D. Zinner[1,7†]

Sexually transmitted infections (STIs) are ubiquitous within wild animal populations, yet it remains largely unknown whether animals evolved behavioral avoidance mechanisms in response to STI acquisition. We investigated the mating behavior of a wild population of olive baboons (*Papio anubis*) infected by the bacterium *Treponema pallidum*. This pathogen causes highly conspicuous genital ulcerations in males and females, which signal infectious individuals. We analyzed data on 876 mating attempts and associated acceptance or rejection responses in a group of about 170 baboons. Our findings indicate that females are more likely to avoid copulation if either the mating partner or females themselves have ulcerated genitals. We suggest that this outcome is linked to the overall higher choosiness and infection-risk susceptibility typically exhibited by females. Our results show that selection pressures imposed by pathogens induce individual behavioral modifications, leading to altered mate choice and could reduce promiscuity in a wild nonhuman primate population.

INTRODUCTION

Infectious diseases are pervasive in the animal kingdom and pose a serious threat to many wildlife populations (*1*). Fitness costs associated with infectious diseases (e.g., reduced fecundity and increased mortality rates) constitute important selection pressures on individuals and have driven the evolution of sophisticated physiological defenses through complex immune systems (*2*). In addition, many species also exhibit behavioral strategies (e.g., grooming avoidance of parasitized conspecifics and selective foraging to avoid contaminated grazing areas) that can serve as the first line of defense against directly or environmentally transmitted pathogens (*3*–*6*).

Sexually transmitted infections (STIs) represent a particularly interesting case, as they are tightly linked to mating behavior and consequently to reproductive fitness. Costs associated with STIs include chronic infections with low recovery rates, reduced offspring survival, sterility, and costly immune defenses (*7*, *8*). However, unexpectedly, little is known whether individuals avoid mating with infectious partners.

Syphilis, a disease caused by the bacterium *Treponema pallidum* subsp. *pallidum*, is one of the most common STIs in humans. Individuals become infected by direct contact, usually sexual, with an infectious lesion (i.e., primary chancre) in the genital area (*9*). In Tanzania, at Lake Manyara National Park (LMNP), olive baboons (*Papio anubis*) are infected with a closely related bacterium [*T. pallidum* subsp. *pertenue* (*TPE*) (*10*, *11*)]. Other nonhuman primates (NHPs) such as yellow baboons (*Papio cynocephalus*), blue monkeys (*Cercopithecus mitis*), and vervet monkeys (*Chlorocebus pygerythrus*) are also found infected (*11*). In humans, *TPE* is known to cause yaws, a nonvenereal disease that spreads via skin-to-skin contact and causes skin ulcers in different body regions [e.g., on the face, arms, and legs (*9*)]. In contrast, clinical signs of *TPE* in NHPs in Tanzania mainly manifest themselves in the anogenital region, resembling a syphilis-like infection. Infected individuals can be identified by the appearance of genital ulcerations that lead to partial or complete mutilation of the external genitalia (Fig. 1) (*10*, *12*, *13*). Nevertheless, in rare cases, facial lesions can also be observed among TPE-infected baboons (*13*). *T. pallidum* is very sensitive to temperature changes and desiccation outside of its host (*9*). Therefore, genital skin-to-skin transmission is a highly effective transmission pathway as the genital area provides a constant moist environment compared to other regions of the body. This, together with the high frequency of observed genital ulcerations in sexually mature NHPs in Tanzania, suggests that *TPE* in those is most likely sexually transmitted (*10*, *12*).

Every time a susceptible individual mates with an infectious partner, it exposes itself to pathogens and subsequent loss of fitness, but the overall benefit of avoidance behavior might be counterbalanced by the cost of missed mating opportunities. Theoretical models have analyzed the potential impacts that STIs might have on mate choice and on the evolution of mating systems (*14*–*17*), but empirical data on animal host–STI interactions are lacking.

In this study, we investigated the sexual behavior of a wild olive baboon population to test whether the genital health status (GHS) had an impact on mate choice and mating behavior. We hypothesized that olive baboons at LMNP can discriminate between genitally ulcerated and non-ulcerated individuals and consequently adapt their mating behavior according to the GHS of the sexual partner. We assumed that individuals were infectious when genital ulcers were present, as is the case in humans infected with *T. pallidum* (*9*). We postulated that (i) non-ulcerated individuals avoid mating with ulcerated conspecifics and that (ii) mating patterns (i.e., the number of mating attempts and copulations) of ulcerated individuals are reduced when compared to non-ulcerated individuals.

RESULTS

Frequency of genital ulcers

The prevalence of genitally ulcerated individuals, determined visually, remained almost stable throughout the 18-month study period. At

[1]Cognitive Ethology Laboratory, Leibniz Institute for Primate Research, German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany. [2]Work Group Neglected Tropical Diseases, Infection Biology Unit, Leibniz Institute for Primate Research, German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany. [3]EpiCenter for Disease Dynamics, One Health Institute, School of Veterinary Medicine, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA. [4]Sokoine University of Agriculture, P.O. Box 3000, Chuo Kikuu, Morogoro, Tanzania. [5]Department of Anthropology, University of California, One Shields Avenue, Davis, CA 95616, USA. [6]Division of Microbiology and Animal Hygiene, Georg-August-University, Göttingen, Germany. [7]Leibniz ScienceCampus Primate Cognition, Göttingen, Germany.
*Corresponding author. Email: fpaciencia@dpz.eu
†These authors shared last authorship.

Fig. 1. Genital ulcerations caused by *Treponema pallidum*. Clinical signs of infection in adult females (top) and males (bottom) of our study group. Photo credit: F. M. D. Paciência, German Primate Center.

the end of the study, 44% of the 53 adult and subadult females and 47% of the 35 adult and subadult males displayed genital ulcers and were therefore considered as infectious (Fig. 1). Six baboons (three males and three females) that showed no ulcers at the beginning of the study developed genital ulcers, and no recovery was observed. No clinical signs were detected in infants, but five juveniles (one female and four males) were observed with genital ulcerations (of an estimated 70 immature individuals). Genital ulcers were observed in 40% of the females participating in sexual behaviors (of n = 32 cycling and well-habituated females) and in 53% of the males (n = 35).

Mating patterns

The majority of the mating behaviors (91% of n = 876 mating attempts) occurred when a female was in peak estrus, i.e., when a female's sexual swelling was at its maximum size. Males initiated most of the mating attempts (86%; Fig. 2). We observed 876 mating attempts between 32 females and 35 males (Fig. 3) of which 540 resulted in copulations. Males with genital ulcers performed fewer pelvic thrusts during copulations than non-ulcerated males (median of 5 and 8 pelvic thrusts, respectively; Mann-Whitney U test, W = 63.948; P < 0.001).

Among the 32 focal females (table S1), we observed an average of 2.94 cycles (range, 1 to 6) within the 18-month study period. During this period, each female had an average of 1.8 mating partners (range, 1 to 6; Fig. 3), with a median of 1 partner per 3 cycles. Females spent 95% of the maximum sexual swelling time in consortship with a male partner.

Contrary to our expectations, the number of mating attempts observed for each possible female-male dyad revealed no significant effect of the GHS of the male or the female (fig. S1), indicating a lack of support for the hypothesis that GHS status affects mating attempts. This result was obtained irrespective of whether the mating attempts were initiated by the males [model TA-1 (total attempt−1); Table 1] or by the females [model TA-2 (total attempt−1); Table 1]. However, as we predicted, acceptance of mating attempts was influenced by the GHS of both baboons in a dyad. Attempts initiated by males [model SA-1 (successful attempt−1); Table 2] were significantly less likely to result in copulations if either the male or the female was genitally ulcerated (male GHS, P = 0.007; female GHS, P = 0.021). The odds of a successful male-initiated copulation were 3.2 times higher if the male was non-ulcerated versus ulcerated and 3.1 times higher if the female was non-ulcerated versus ulcerated. In contrast, acceptance of mating attempts initiated by females [model SA-2 (successful attempt−2); Table 2] was not significantly influenced by either male or female GHS (male GHS, P = 0.55; female GHS, P = 0.37), despite the high number of mating avoidance events by males (Fig. 2).

DISCUSSION

In animal populations, mating is tightly linked to fitness maximization, and therefore, the selection of healthy partners is crucial to avoid fitness costs associated with STIs. We examined whether wild olive baboons avoided mating with individuals affected by conspicuous genital ulcers caused by *T. pallidum*. Our results demonstrate a significant effect of clinically apparent *T. pallidum* infection on the mating behavior of female baboons. Females were more likely to avoid copulation if approached by males with ulcerated genitals, indicating behavioral avoidance of diseased conspecifics. NHP females are often more susceptible to STI acquisition than males, as

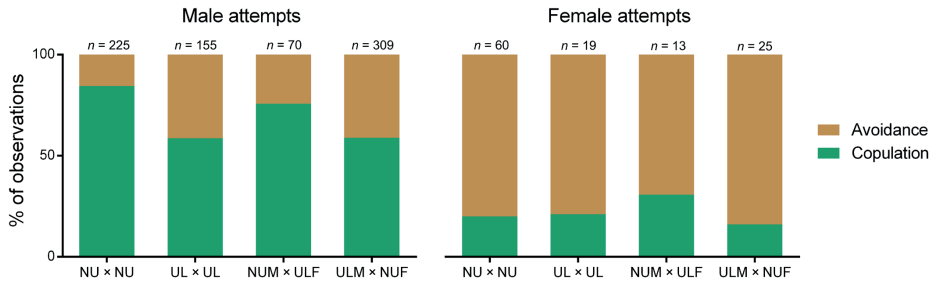


Fig. 2. Percentage of male and female mating attempts according to their GHS. Each attempt could result in either copulation (green bar) or avoidance (brown bar). The percentage of observations is shown on the *y* axis. The attempts between individuals according to their GHS are shown on the *x* axis. NU × NU (non-ulcerated × non-ulcerated), UL × UL (ulcerated × ulcerated), NUM × ULF (non-ulcerated male × ulcerated female), ULM × NUF (ulcerated male × non-ulcerated female). Sample sizes are shown at the top of each bar.

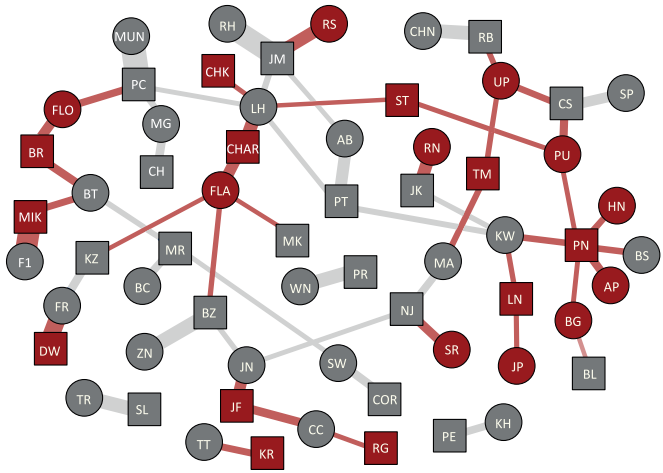


Fig. 3. Olive baboon mating network. Nodes represent individual baboons and are colored according to each baboon's GHS. Red nodes represent ulcerated individuals and gray non-ulcerated. Individuals who switched GHS are colored according to the status which had a higher number of copulations. Squares and circles represent males and females, respectively. Edges are weighted by the number of copulations among dyads and respective focal time. Edge colors correspond to the pairwise GHS of the nodes they connect, with red edges representing copulations where at least one of the individuals is ulcerated, while gray edges represent copulations between non-ulcerated individuals.

is observed in humans (*2*). Female-biased disease risk avoidance is widespread across species, and infections can significantly affect mate choice (*18*). In general, males that are infected with directly transmitted parasites (e.g., nematodes, protozoa, bacteria, and viruses) are less preferred mating partners by their conspecific females (*18*). As compared to men, women also exhibit a higher level of disgust toward potential disease threats (e.g., contaminated environment and spoiled food) and show a higher tendency to avoid infected sexual partners [e.g., individuals exhibiting genital lesions (*19*, *20*)]. The results of our study are in accordance with this generalized higher disgust reported in females since mating with STI-infected partners can entail greater costs to females due to obligatory investment in gestation, lactation, and infant rearing.

We also found that avoidance by females was more frequent if females themselves had ulcerated genitals. The progressive scarification of the genitalia due to infection can lead to a permanently open state of the vagina and anus, which increases urogenital infections (i.e., promoting miscarriages) and risk of dystocia. Thus, in addition to affecting the mate choice of non-ulcerated individuals, *T. pallidum* appears to affect mating behavior of ulcerated females, as genital

Table 1. Total attempt models. Weibull mixed-effects models evaluating whether the total number of mating attempts is influenced by the dyad's GHS of the male and/or the female. Mating attempts initiated by males and females are shown in TA-1 and TA-2, respectively. Given are posterior means, SDs, 2.5 and 97.5% confidence intervals (CIs), and effective sample sizes (*n*). Rhat = 1 (model convergence) for all parameters. Models were run using four chains, each with 2000 iterations, a burn-in of 1000, and thinning set to 1.

TA-1 model

Term	Mean	SD	CI lower	CI upper	*n* (effective)
Intercept *q*	−3.40	0.24	−3.89	−2.94	2799
Intercept β	−2.11	0.20	−2.53	−1.75	2550
Main effects:					
GHS females *q*	0.04	0.28	−0.52	0.59	4000
GHS females β	0.17	0.24	−0.31	0.64	4000
GHS males *q*	0.37	0.28	−0.18	0.92	2942
GHS males β	0.07	0.24	−0.40	0.56	3025
Random effects:					
Male ID *q*	0.24	0.17	0.01	0.63	1803
Male ID β	0.14	0.11	0.01	0.40	2027
Female ID *q*	0.17	0.13	0.01	0.47	2424
Female ID β	0.11	0.09	0.00	0.33	4000

TA-2 model

Term	Mean	SD	CI lower	CI upper	*n* (effective)
Intercept *q*	−3.56	0.28	−4.15	−3.05	4000
Intercept β	−1.32	0.25	−1.80	−0.83	4000
Main effects:					
GHS females *q*	0.12	0.32	−0.51	0.73	4000
GHS females β	0.50	0.34	−0.18	1.17	4000
GHS males *q*	−0.21	0.35	−0.92	0.44	4000
GHS males β	−0.27	0.35	−1.00	0.37	4000
Random effects:					
Male ID *q*	0.43	0.27	0.02	1.02	1222
Male ID β	0.19	0.16	0.01	0.59	1982
Female ID *q*	0.24	0.18	0.01	0.65	2192
Female ID β	0.32	0.22	0.01	0.83	1385

ulcers are likely painful and potentially cause afflicted females to refrain from mating.

While we found no evidence to support the hypothesis that males avoid mating opportunities based on their own or their partner's GHS, genitally ulcerated individuals performed significantly fewer pelvic thrusts during copulation than non-ulcerated males. Fewer pelvic thrusts might result in no or fewer ejaculation events. In addition, some symptomatic individuals were observed showing unusual body contractions with simultaneous vocalizations (e.g., kecking) during urination. As observed in females, genital ulcers seem to be equally painful for males, thus, to a certain extent, ulcerated males might also incur fitness costs.

STIs have been hypothesized to constitute a key selection pressure in shaping the evolution of mating strategies (*14*, *15*, *21*). A model based on a human hunter-gatherer population indicated that a high prevalence of STIs in large group sizes could foster the emergence of socially imposed monogamy (*22*). Cycling females of yellow and olive baboons have been reported to copulate with most of the males comprising their group (*23*, *24*). In contrast, in our study, female olive baboons only consorted and mated with few partners (average of 1.8), despite the large pool of males available ($n = 35$). While the reason for this low promiscuity is unclear, and studies reporting on olive baboon promiscuity are scarce, it would be important to investigate whether factors such as group size, pathogen incidence, pathogen virulence, and fitness consequences due to *T. pallidum* exert a selective pressure on our baboon population, which, in return, could lead to altered mating strategies.

A potential limitation of our study is that we were unable to collect data on male dominance rank; however, we think it is unlikely that including rank data in our models would significantly affect our findings. First, the vast majority of sexually mature males (83%) in our study group engaged in copulations and established consortships with peak-estrous females, indicating a lack of mating skew according to rank. Second, despite the fact that dominance rank has

Table 2. Successful attempt models. GLMMs evaluating whether the success of a mating attempt (i.e., likelihood of copulation) is influenced by the dyad's GHS of the male and/or the female. Mating attempts initiated by males and females are shown in SA-1 and SA-2, respectively. Estimates, SEs, df, and 2.5 and 97.5% confidence intervals are shown for fixed effects. Significant variables ($P < 0.05$) are shown in bold. Intercept with reference category for non-ulcerated individuals.

Term	Estimate	SE	CI lower	CI upper	X^2	df	P
SA-1 model							
Intercept	2.371	0.477	1.628	3.326	*	*	*
GHS females	−1.127	0.480	−1.962	−0.325	7.187	1	**0.007**
GHS males	−1.167	0.481	−2.151	−0.228	5.315	1	**0.021**
SA-2 model							
Intercept	−1.542	0.469	−2.782	−0.735	*	*	*
GHS females	0.577	0.647	−0.738	1.983	0.794	1	0.373
GHS males	−0.396	0.657	−1.727	1.055	0.348	1	0.555

*Not shown due to very limited interpretation

been considered a predictor of mating success in yellow and olive baboons, it is also a subject of high variation among groups and individuals, and it appears to depend on multiple factors (*25*, *26*). This highly contrasts with chacma baboons, where rank and access to estrous females are strongly connected (*27*). In addition, in yellow and olive baboons, group size and composition are important determinants of mating access among male baboons: High-ranking males tend to lose their monopoly over cycling females in large groups (*28*), and male dominance ranks are less pronounced in large groups with many cycling females (*29*). Given that olive baboon groups at LMNP are extraordinarily large (averaging approximately 150 individuals), we consider that rank does not play a major role regarding access to reproductive females in our study population. Instead, our results indicate that partner preference is crucially important at LMNP, as we observed cycling females maintaining a consistent male partner across consecutive estrus cycles.

Empirical studies on mating behavior in relation to STIs are scarce and mainly confined to arthropods. In these, no evidence of mating avoidance or discrimination between healthy and sick conspecifics has been found (*30*–*32*). Fitness disadvantages, disgust, and pain anticipation might be influencing the mating patterns of our group. Our study reporting on the ability of a long-lived vertebrate to discriminate among mating partners according to health status sheds light on how sexually transmitted pathogens can markedly shape mating dynamics in a group of NHPs. Last, given that *T. pallidum* is present in humans in overlapping areas with infected NHPs (*11*), the investigation of the underlying mechanisms affecting pathogen transmission is of the uttermost importance to mitigate health risks for both human and wildlife communities.

MATERIALS AND METHODS

Study site and subjects

Fieldwork was conducted at LMNP, Northern Tanzania (3°28′ S 35°46′E), during two field seasons (April to December) in 2015 and 2016. LMNP is a small protected area (approximately 580 km^2) with almost 220 km^2 of lake coverage. In 2016, we conducted a survey estimating the population size of the LMNP baboons in the park excluding the recently added area of the Marang forest (250 km^2). The population was estimated to consist of approximately 5200 olive baboons, and individuals with ulcerated genitals were observed in most groups within LMNP.

We habituated our study group ($n \approx 170$ baboons) over a 4-month period before data collection. To enhance location of the group, we radio-collared three adult females (Advanced Telemetry Systems Inc., Isanti, MN, USA). Proceedings on immobilization and anesthesia are described in (*13*). Our analyses focused on 26 adult and 11 subadult females (after excluding data of females observed <1.5 hours) and 28 adult and 7 subadult males. Age categories were defined as in (*33*). Adult males were identified by their large body size and fully developed secondary sexual traits; subadult males were larger than females but lacked secondary sexual characteristics (e.g., large shoulder mane and elongated canines). Adult females were identified as individuals that have reached full body size, whereas subadult females were smaller and lacked elongated nipples (but were already cycling).

Genital health status

Within our study group, baboons were categorized as either "genitally ulcerated" or "non-ulcerated," based on their GHS, using macroscopic visual cues [as in (*10*)]. In both males and females, genital ulcerations were observed ranging from small-medium ulcers to severe necrotizing dermatitis and mutilation of the outer genital structures (Fig. 1).

Behavioral data

We conducted focal follows (*34*) from dawn to dusk on subadult and adult females. We aimed for full-day focal follows, but if a focal individual was out of sight for more than 10 min, then another baboon was selected. We recorded 597 hours of observation data, with an average of 16.40 ± 10.02 hours (mean ± SD; range, 1.50 to 39.00 hours) per focal female. We prioritized following females in their peak estrus [denoted by maximal tumescence of the anogenital area and bright pink color (*35*)] to maximize the number of mating events observed. We collected data on the sexual behavior of focal females and their partners (table S1), recording when dyads participated in consortships in which a male maintained close proximity to a female and attempted to prevent other males from mating with her (*24*). In addition, we collected data on the frequency and success of mating attempts led by either males or females. Male-led mating

attempts were identified as a male trying to mount a female with the performance of pelvic thrusts (*36*). Attempts led by females were documented when a female presented her perineum to a male while lifting her tail (*37*). Attempts by either sex were considered successful if they resulted in copulation. Because of the loss of the corpus penis or phimosis, some males were unable to engage in intromission and/or ejaculation, precluding our ability to use these behaviors to define successful mating attempts. An unsuccessful attempt was defined as a female rejecting a male's attempt to mount (e.g., by sitting or fleeing) or a male refraining from mounting a female after she presented to him. Behavioral data were recorded on a hand-held device (Samsung Galaxy Hand Note) in the field using Pendragon 5.1.2 software (Pendragon Software Corporation, USA) and transferred daily onto computers for error checking and data storage.

Statistical analyses

Total attempt model

We first investigated whether the number of mating attempts was affected by the GHS of a dyad. To examine this, we developed a dataset that indicated the corresponding number of attempts made by the male and female of every possible dyad in the study group, along with their respective GHSs. Including every possible dyad allowed us to consider all potential mating attempts, as mating could theoretically occur between any cycling female and any male in the group. This dataset included a high proportion (96.3%) of zero attempt data, with data for nonzero attempts showing extremely high variance ($\sigma^2 = 149.9$). Distributions commonly used in instances of zero inflation [e.g., zero-inflated Poisson or negative binomial distributions (*38*)] were a poor fit for our data; however, the discrete Weibull distribution (*39*) proved to be a good fit (fig. S2). In particular, the discrete Weibull distribution is highly flexible in modeling under- and overdispersed data relative to the Poisson distribution (*40*). The discrete Weibull probability mass function is defined as

$$P(X = x) = q^{x^{\beta}} - q^{(x+1)^{\beta}}$$

with positive shape parameter β and parameter q satisfying $0 < q < 1$. As no current R packages offered the functionality to fit mixed-effects models with a discrete Weibull distribution, we used the Bayesian programming language "Stan" to develop discrete Weibull mixed-effects models (*41*). We built two models: one predicting the number of attempts led by males (model TA-1; Table 1) and a second for attempts led by females (model TA-2; Table 1). Each model assumed that q and β were a linear combination of the male GHS and the female GHS (fig. S2). Both linear predictors also included male and female identities as crossed random effect variables affecting model intercepts. We used uninformative or weakly informative priors. Fixed effect coefficients had normally distributed priors and SDs equal to 100. Random effects were assumed to be normally distributed with SDs sampled from Cauchy prior distributions with location parameters equal to zero and scale parameters equal to 25. All models were fitted using R v 3.4.4 (*42*) with the "Rstan" package (*43*).

Successful attempt model

Here, we examined whether the success of a mating attempt is affected by the GHS of the dyad. Our dataset included the success (1/0) of each mating attempt observed, along with the GHS of the male and female involved in the mating attempt. We built two generalized linear mixed models (GLMMs) (*44*): one model for attempts led by males (model SA-1; Table 2) and a second model for attempts led by females (model SA-2; Table 2). Each model assumed that the probability of a successful mating event depended on the GHS of the male and GHS of the female. Both models had a binomial error structure with a logit link function and included the individual and pair identities as random effects. In addition, we included the focal observation time as an offset term to account for variation in observation effort among cycling females. The interactions between male and female GHS were nonsignificant and were excluded from the final models. Models were run in R v 3.4.4 (*42*) with the lme4 package v 1.1-15 (*45*), and P values are shown based on likelihood ratio tests of individual fixed effects [function drop1 with argument *test* set to "Chisq", (*46*)].

Network visualization

We constructed a mating network (Fig. 3) using a force-directed (Fruchterman-Reingold) layout in R v 3.4.4 (*42*) with the igraph package v 1.2.2. Connections between nodes (edges) were weighted according to each dyad's copulation rate (i.e., the number of copulations after controlling for the female's observation time). To aid in visualizing connections among nodes, we made minor adjustments to the final network graphic (e.g., preventing overlapping nodes).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/12/eaaw9724/DC1

Fig. S1. Posterior probabilities with q and β estimation parameters.
Fig. S2. Observed and predicted values using a discrete Weibull distribution.
Table S1. Focal females and their respective mating partners.

REFERENCES AND NOTES

1. K. F. Smith, K. Acevedo-Whitehouse, A. B. Pedersen, The role of infectious diseases in biological conservation. *Anim. Conserv.* **12**, 1–12 (2009).
2. C. L. Nunn, S. Altizer, *Infectious Diseases in Primates: Behavior, Ecology and Evolution* (Oxford Univ. Press, Oxford, 2006).
3. V. O. Ezenwa, Host social behavior and parasitic infection: A multifactorial approach. *Behav. Ecol.* **15**, 446–454 (2004).
4. C. Sarabian, A. J. J. MacIntosh, Hygienic tendencies correlate with low geohelminth infection in free-ranging macaques. *Biol. Lett.* **11**, 20150757 (2015).
5. C. Poirotte, F. Massol, A. Herbert, E. Willaume, P. M. Bomo, P. M. Kappeler, M. J. E. Charpentier, Mandrills use olfaction to socially avoid parasitized conspecifics. *Sci. Adv.* **3**, e1601721 (2017).
6. D. C. Behringer, M. J. Butler, J. D. Shields, Avoidance of disease by social lobsters. *Nature* **441**, 421 (2006).
7. A. B. Lockhart, P. H. Thrall, J. Antonovics, Sexually transmitted diseases in animals: Ecological and evolutionary implications. *Biol. Rev.* **71**, 415–471 (1996).
8. C. L. Nunn, J. L. Gittleman, J. Antonovics, Promiscuity and the primate immune system. *Science* **290**, 1168–1170 (2000).
9. K. K. Holmes, P. F. Sparling, W. E. Stamm, P. Piot, J. N. Wasserheit, L. Corey, M. S. Cohen, D. H. Watts, R. A. Nelson, *Sexually Transmitted Diseases* (McGraw Hill Professional, 2007).
10. S. Knauf, E. K. Batamuzi, T. Mlengeya, M. Kilewo, I. A. V. Lejora, M. Nordhoff, B. Ehlers, K. N. Harper, R. Fyumagwa, R. Hoare, K. Failing, A. Wehrend, F. J. Kaup, F. H. Leendertz, K. Mätz-Rensing, *Treponema* infection associated with genital ulceration in wild baboons. *Vet. Pathol.* **49**, 292–303 (2012).
11. S. Knauf, J. F. Gogarten, V. J. Schuenemann, H. M. De Nys, A. Düx, M. Strouhal, L. Mikalová, K. I. Bos, R. Armstrong, E. K. Batamuzi, I. S. Chuma, B. Davoust, G. Diatta, R. D. Fyumagwa, R. R. Kazwala, J. D. Keyyu, I. A. V. Lejora, A. Levasseur, H. Liu, M. A. Mayhew, O. Mediannikov, D. Raoult, R. M. Wittig, C. Roos, F. H. Leendertz, D. Šmajs, K. Nieselt, J. Krause, S. Calvignac-Spencer, Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. *Emerg. Microbes Infect.* **7**, 1–4 (2018).
12. T. Mlengeya, Distribution pattern of a sexually transmitted disease (STD) of olive baboons in Lake Manyara National Park, Tanzania, thesis, College of African Wildlife Management (2004).
13. I. S. Chuma, E. K. Batamuzi, D. A. Collins, R. D. Fyumagwa, L. K. Hallmaier-Wacker, R. R. Kazwala, J. D. Keyyu, I. A. Lejora, I. F. Lipende, S. Lüert, F. M. D. Paciência, A. Piel, F. A. Stewart, D. Zinner, C. Roos, S. Knauf, Widespread *Treponema pallidum* infection in nonhuman primates, Tanzania. *Emerg. Infect. Dis.* **24**, 1002–1009 (2016).

14. P. H. Thrall, J. Antonovics, J. D. Bever, Sexual transmission of disease and host mating systems: Within-season reproductive success. *Am. Nat.* **149**, 485–506 (1997).
15. H. Kokko, E. Ranta, G. Ruxton, P. Lundberg, Sexually transmitted disease and the evolution of mating systems. *Evolution* **56**, 1091–1100 (2002).
16. M. Boots, R. J. Knell, The evolution of risky behaviour in the presence of a sexually transmitted disease. *Proc. Biol. Sci.* **269**, 585–589 (2002).
17. D. V. McLeod, T. Day, Sexually transmitted infection and the evolution of serial monogamy. *Proc. Biol. Sci.* **281**, 20141726 (2014).
18. S. Beltran-Bech, F.-J. Richard, Impact of infection on mate choice. *Anim. Behav.* **90**, 159–170 (2014).
19. J. M. Tybur, S. W. Gangestad, Mate preferences and infectious disease: Theoretical considerations and evidence in humans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 3375–3388 (2011).
20. V. Curtis, M. de Barra, The structure and function of pathogen disgust. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, 20170208 (2018).
21. B. C. Sheldon, Sexually transmitted disease in birds: Occurrence and evolutionary significance. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **339**, 491–497 (1993).
22. C. T. Bauch, R. McElreath, Disease dynamics and costly punishment can foster socially imposed monogamy. *Nat. Commun.* **7**, 11219 (2016).
23. G. Hausfater, *Dominance and Reproduction in Baboons (Papio cynocephalus)* (Karger, 1975).
24. F. B. Bercovitch, Female cooperation, consortship maintenance, and male mating success in savanna baboons. *Anim. Behav.* **50**, 137–149 (1995).
25. S. C. Alberts, H. E. Watts, J. Altmann, Queuing and queue-jumping: Long-term patterns of reproductive skew in male savannah baboons, *Papio cynocephalus*. *Anim. Behav.* **65**, 821–840 (2003).
26. J. T. Walz, Competition, Coercion, and Choice: The Sex Lives of Female Olive Baboons (*Papio anubis*), Ohio State University (2016).
27. J. B. Bulger, Dominance rank and access to estrous females in male savanna baboons. *Behaviour* **127**, 67–103 (1993).
28. G. Cowlishaw, R. I. M. Dunbar, Dominance rank and mating success in male primates. *Anim. Behav.* **41**, 1045–1056 (1991).
29. I. DeVore, Male dominance and mating behavior in baboons, in *Sex and Behavior*, F.A. Beach Eds. (New York, J. Wiley, 1965).
30. P. Abbot, L. M. Dill, Sexually transmitted parasites and sexual selection in the milkweed leaf beetle, *Labidomera clivicollis*. *Oikos* **92**, 91–100 (2001).
31. M. Webberley, G. D. D. Hurst, J. Buszko, M. E. N. Majerus, Lack of parasite mediated sexual selection in a ladybird/sexually transmitted disease system. *Anim. Behav.* **63**, 131–141 (2002).
32. L. T. Luong, H. K. Kaya, Sexually transmitted parasites and host mating behavior in the decorated cricket. *Behav. Ecol.* **16**, 794–799 (2005).
33. T. Rowell, Long-term changes in a population of Ugandan baboons. *Folia Primatol.* **11**, 241–254 (1969).
34. J. Altmann, Observational study of behavior: Sampling methods. *Behaviour* **49**, 227–267 (1974).
35. D. Zinner, C. L. Nunn, C. P. van Schaik, P. M. Kappeler, Sexual selection and exaggerated sexual swellings of female primates, in *Sexual Selection in Primates. New and Comparative Perspectives*, P. M. Kappeler, C. P. van Schaik CP, Eds. (Cambridge Univ. Press, 2004).
36. F. Nitsch, S. Stueckle, D. Stahl, D. Zinner, Copulation patterns in captive hamadryas baboons: A quantitative analysis. *Primates* **52**, 373 (2011).
37. L. Rigaill, J. P. Higham, P. C. Lee, A. Blin, C. Garcia, Multimodal sexual signaling and mating behavior in olive baboons (*Papio anubis*). *Am. J. Primatol.* **75**, 774–787 (2013).
38. A. F. Zuur, E. N. Ieno, N. Walker, A. A. Saveliev, G. M. Smith, *Mixed Effects Models and Extensions in Ecology with R* (Springer, 2009).
39. T. Nakagawa, S. Osaki, The Discrete Weibull Distribution. *IEEE. Trans. Reliab.* **R-24**, 300–301 (1975).
40. H. S. Klakattawi, V. Vinciotti, K. Yu, A simple and adaptive dispersion regression model for count data. *Entropy* **20**, 142 (2018).
41. B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, A. Riddell, Stan: A probabilistic programming language. *J. Stat. Softw.* **76**, 1 (2017).
42. R Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2018).
43. Stan Development Team, RStan: The R interface to Stan. R package version 2.17.4 (2018).
44. R. H. Baayen, *Analyzing Linguistic data: A Practical Introduction to Statistics Using R* (Cambridge Univ. Press, New York, 2008).
45. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1 (2015).
46. D. J. Barr, R. Levy, C. Scheepers, H. J. Tily, Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).

Acknowledgments: We thank Tanzania Wildlife Research Institute (TAWIRI), especially J. D. Keyyu and R. Fyumagwa; Tanzania National Parks (TANAPA), especially I. A. V. Lejora; LMNP headquarters staff, particularly R. Kaitila. We thank the LMNP rangers who helped with fieldwork: P. Mkama, D. Baluya, P. Mbaryo, and J. Bitulo, without them, this study would not have been possible. We also thank the Tanzania Commission for Science and Technology (COSTECH) for their support. We thank the editor and two anonymous reviewers for their helpful comments. The Animal Welfare and Ethics Committee of the German Primate Center approved the entire study. Certified veterinarians performed all baboon immobilizations and radio collar placements. All procedures in which animals were handled were performed according to the rules and regulations of "Good Veterinary Practice" and in compliance with the Tanzanian and German laws. **Funding:** Financial support was provided by the German Science Foundation (DFG) and conducted as part of the research group Sociality and Health in Primates (KN1097/4-1 to S.K. and ZI548/5-1 to D.Z.). **Author contributions:** F.M.D.P., S.K., and D.Z. designed the study. F.M.D.P., I.S.C., I.F.L., and S.K. performed field work including the collaring of the baboon females. F.M.D.P. collected the data in the field. Data analysis was performed by F.M.D.P., S.K., J.R., and D.C. The paper was written by F.M.D.P., J.R., D.C., S.K., and D.Z. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from F.M.D.P. (fpaciencia@dpz.eu) or D.Z. (dzinner@dpz.eu).

Submitted 11 February 2019
Accepted 18 October 2019
Published 4 December 2019
10.1126/sciadv.aaw9724

Citation: F. M. D. Paciência, J. Rushmore, I. S. Chuma, I. F. Lipende, D. Caillaud, S. Knauf, D. Zinner, Mating avoidance in female olive baboons (*Papio anubis*) infected by *Treponema pallidum*. *Sci. Adv.* **5**, eaaw9724 (2019).

ScienceAdvances

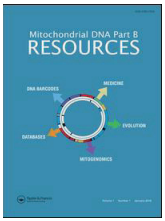
Mating avoidance in female olive baboons (*Papio anubis*) infected by *Treponema pallidum*

F. M. D. Paciência, J. Rushmore, I. S. Chuma, I. F. Lipende, D. Caillaud, S. Knauf and D. Zinner

Sci Adv **5** (12), eaaw9724.
DOI: 10.1126/sciadv.aaw9724

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

7.4.5 Complete Mitochondrial Genome of an Olive Baboon (*Papio anubis*) From Gombe National Park, Tanzania



Taylor & Francis
Taylor & Francis Group

Mitochondrial DNA Part B
Resources

ISSN: (Print) 2380-2359 (Online) Journal homepage: http://www.tandfonline.com/loi/tmdn20

Complete mitochondrial genome of an olive baboon (Papio anubis) from Gombe National Park, Tanzania

Christian Roos, Idrissa S. Chuma, D. Anthony Collins, Sascha Knauf & Dietmar Zinner

To cite this article: Christian Roos, Idrissa S. Chuma, D. Anthony Collins, Sascha Knauf & Dietmar Zinner (2018) Complete mitochondrial genome of an olive baboon (Papio anubis) from Gombe National Park, Tanzania, Mitochondrial DNA Part B, 3:1, 177-178

To link to this article: https://doi.org/10.1080/23802359.2018.1437813

Published online: 09 Feb 2018.

MITOCHONDRIAL DNA PART B: RESOURCES, 2018
VOL. 3, NO. 1, 177–178
https://doi.org/10.1080/23802359.2018.1437813

Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

OPEN ACCESS Check for updates

Complete mitochondrial genome of an olive baboon (*Papio anubis*) from Gombe National Park, Tanzania

Christian Roos[a,b], Idrissa S. Chuma[b,c,d], D. Anthony Collins[e], Sascha Knauf[d] and Dietmar Zinner[f]

[a]Gene Bank of Primates, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany; [b]Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany; [c]Department of Veterinary Medicine and Public Health, College of Veterinary and Biomedical Sciences, Sokoine University of Agriculture, Morogoro, Tanzania; [d]Work Group Neglected Tropical Diseases, Infection Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany; [e]Gombe Stream Research Centre, Jane Goodall Institute, Kigoma, Tanzania; [f]Cognitive Ethology Laboratory, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany

ABSTRACT

The olive baboon (*Papio anubis*) is the most widely distributed baboon species. We report here on the complete mitochondrial genome of an olive baboon from the south-eastern edge of the species' range from Gombe National Park (NP), Tanzania. The genome (GenBank accession number MG787545) has a length of 16,490 bp and exhibits the typical structure of mammalian mitochondrial genomes. Phylogenetically, the olive baboon from Gombe NP is most closely related to eastern *P. anubis*, northern *P. cynocephalus* and *P. hamadryas*. The data are an important addition to further clarify the phylogeography of baboons and phylogeny of papionins in general.

ARTICLE HISTORY
Received 16 January 2018
Accepted 18 January 2018

KEYWORDS
Sanger sequencing; Cercopithecidae; non-human primates

Baboons, genus *Papio*, are widely distributed Old World monkeys that occur over almost all sub-Saharan Africa and parts of the Arabian Peninsula. Among the six species, the olive baboon (*Papio anubis*) has the largest range and occurs from Mali and Guinea in the West to Eritrea in the East, and South to Uganda, Democratic Republic of Congo and Tanzania (Anandam et al. 2013). Phylogenetic studies on baboons using fragments of the mitochondrial genome revealed seven major clades that display a geographic pattern, but disagree with the six species taxonomy. Using these fragments, the branching pattern among the clades remained largely unresolved, while complete mitochondrial genome data revealed a well-supported phylogeny (Zinner et al. 2013).

We report here on the sequencing of a mitochondrial genome of an olive baboon from its southernmost range from Gombe National Park (NP), Tanzania. The skin sample (individual ID 19GNM2220916; S04°40′41″, E29°37′15″) was collected for the purpose of screening for *Treponema pallidum* infection in non-human primates and not specifically for this study. Collection was in accordance with Tanzanian and German laws and guidelines. DNA was extracted with methods outlined in Knauf et al. (2016) and the complete mitochondrial genome was PCR amplified, sequenced and assembled following methods described in Zinner et al. (2013).

The newly generated mitochondrial genome exhibits and A + T content of 56.32% and contains 13 protein-coding genes, 22 transfer RNAs, two ribosomal RNAs and the control region in the order typically found in mammals (Anderson et al. 1981).

We performed phylogenetic reconstructions by adding mitochondrial genome data of other baboons representing all species and major clades (Zinner et al. 2009, 2013) and *Theropithecus gelada* as an outgroup. Sequences were aligned with Muscle 3.8.31 (Edgar 2010) in SeaView 4.5.4 (Gouy et al. 2010), and indels and poorly aligned positions were removed with Gblocks 0.91b (Castresana 2000). A maximum-likelihood tree was generated in IQ-TREE 1.5.2 (Nguyen et al. 2015) using the optimal substitution model (TrN + I+G) as selected by ModelFinder (Chernomor et al. 2016; Kalyaanamoorthy et al. 2017) and 10,000 ultrafast bootstrap replicates (Minh et al. 2013). In the obtained tree (Figure 1), the olive baboon from Gombe NP forms a strongly supported (100% bootstrap) sister lineage to a clade consisting of eastern *P. anubis*, northern *P. cynocephalus* and *P. hamadryas*, thus, further supporting the previously detected polyphyly of olive baboons (Zinner et al. 2009, 2013).

The mitochondrial genome of an olive baboon from Gombe NP, Tanzania is an important addition to understand the phylogeography of baboons and to further investigate the phylogeny papionins in general.

CONTACT Christian Roos croos@dpz.eu Gene Bank of Primates, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany

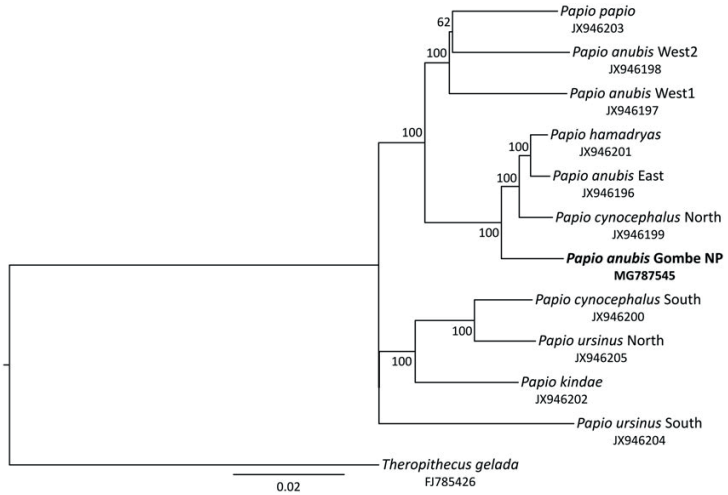


Figure 1. Maximum-likelihood tree showing phylogenetic relationships among baboon lineages. The olive baboon from Gombe NP (highlighted in bold) clusters with eastern *P. anubis*, northern *P. cynocephalus* and *P. hamadryas*. Numbers on nodes refer to bootstrap values and the bar indicates substitutions per site. GenBank accession numbers are listed below species.

Acknowledgements

We thank the Government of the United Republic of Tanzania and its relevant institutions (Tanzania Wildlife Research Institute, Tanzania National Parks, Commission for Science and Technology) for permission and support in undertaking the study. Our gratitude goes to all Gombe National Park staff for support of the sampling activities. In the laboratory, we further thank Simone Lueert and Christiane Schwarz for their help to generate the sequence data.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

Sample collection was supported by grants from the German Research Foundation (DFG) [Grant Numbers KN1097/3-1 and RO3055/2-1].

ORCID

Christian Roos http://orcid.org/0000-0003-0190-4266

References

Anandam MV, Bennett EL, Davenport TRB, Davies NJ, Detwiler KM, Engelhardt A, Eudey AA, Gadsby EL, Groves CP, Healy A, et al. 2013. Family Cercopithecidae (Old World monkeys) – species accounts of Cercopithecidae. In: Mittermeier RA, Rylands AB, Wilson DE, editors. Handbook of the mammals of the world, volume 3: primates. Barcelona: Lynx Edicions; p. 628–753.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. 1981. Sequence and organization of the human mitochondrial genome. Nature. 290:457–465.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 17:540–552.

Chernomor O, von Haeseler A, Minh BQ. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Syst Biol. 65:997–1008.

Edgar RC. 2010. Quality measures for protein alignment benchmarks. Nucleic Acids Res. 38:2145–2153.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 27:221–224.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14:587–589.

Knauf S, Raphael J, Mitjà O, Lejora IAV, Chuma IS, Batamuzi EK, Keyyu JD, Fyumagwa R, Lüert S, Godornes C, et al. 2016. Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. EBioMedicine. 11:85–90.

Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol. 30:1188–1195.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32:268–274.

Zinner D, Groeneveld LF, Keller C, Roos C. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.): indication for introgressive hybridization? BMC Evol Biol. 9:83.

Zinner D, Wertheimer J, Liedigk R, Groeneveld LF, Roos C. 2013. Baboon phylogeny as inferred from complete mitochondrial genomes. Am J Phys Anthropol. 150:133–140.

7.4.6 Distribution of Mitochondrial Clades and Morphotypes of Baboons *Papio* spp. (Primates: Cercopithecidae) in Eastern Africa

Distribution of Mitochondrial Clades and Morphotypes of Baboons *Papio* spp. (Primates: Cercopithecidae) in Eastern Africa

Author(s): Dietmar Zinner, Christina Keller, Julius W. Nyahongo, Thomas M. Butynski, Yvonne A. de Jong, Luca Pozzi, Sascha Knauf, Rasmus Liedigk & Christian Roos
Source: Journal of East African Natural History, 104(1-2):143-168.
Published By: Nature Kenya/East African Natural History Society
DOI: http://dx.doi.org/10.2982/028.104.0111
URL: http://www.bioone.org/doi/full/10.2982/028.104.0111

Journal of East African Natural History **104**(1&2)**:** 143–168 (2015)

DISTRIBUTION OF MITOCHONDRIAL CLADES AND MORPHOTYPES OF BABOONS *PAPIO* SPP. (PRIMATES: CERCOPITHECIDAE) IN EASTERN AFRICA

Dietmar Zinner, Christina Keller
Cognitive Ethology Laboratory, German Primate Center
37077 Göttingen, Germany
dzinner@gwdg.de

Julius W. Nyahongo
School of Biological Sciences, University of Dodoma
P.O. Box 259, Dodoma, Tanzania
nyhwjulius@yahoo.co.uk

Thomas M. Butynski, Yvonne A. de Jong
Lolldaiga Hills Research Programme &
Eastern Africa Primate Diversity and Conservation Program
P.O. Box 149, Nanyuki, 10400, Kenya
TButynski@aol.com

Luca Pozzi
Behavioural Ecology & Sociobiology Unit, German Primate Center
37077 Göttingen, Germany
&
Department of Anthropology, University of Texas at San Antonio
One UTSA Circle, San Antonio, TX 78249
luca.pozzi@utsa.edu

Sascha Knauf
Pathology Unit, Work Group Neglected Tropical Diseases, German Primate Center
37077 Göttingen, Germany
SKnauf@dpz.eu

Rasmus Liedigk
Cognitive Ethology Laboratory, German Primate Center
37077 Göttingen, Germany
RLiedigk@dpz.eu

Christian Roos
Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center
37077 Göttingen, Germany
CRoos@dpz.eu

ABSTRACT

Recent genetic studies, using maternally inherited mitochondrial DNA, indicate a complex evolutionary history for baboons *Papio* spp. in general, and for eastern African baboons in particular. To further address this topic and to improve our understanding of phylogeographic patterns of baboons in eastern Africa, mitochondrial cytochrome *b* sequence data were analysed from 148 baboon samples from 103 locations in eastern Africa. The resultant phylogenetic reconstructions suggest an initial split of baboons into four main clades: southern chacma baboons, baboons from Mahale Mountains in Tanzania, main southern, and main northern. We confirm that the boundary between southern and northern clades lies along the Ugalla-Malagarasi River and Ruaha-Rufiji River of central Tanzania. We detected new mitochondrial haplogroups, most notably the Mahale Mountains clade, and refined haplogroup distributions. The evolutionary divergence of baboons in eastern Africa was most likely triggered and maintained by numerous episodes of population division and reconnection, probably related mainly to climate change. To better understand these processes, nuclear DNA information is required, especially to assess gene flow among populations.

Keywords: phylogeography, cytochrome *b,* Ruaha-Rufiji River, Ugalla-Malagarasi River

INTRODUCTION

Baboons (genus *Papio* Erxleben, 1777) are widely distributed over most of sub-Saharan Africa, occupying the greatest diversity of habitats of any genus of non-human primates. Baboons are found in all terrestrial habitats from moist forest to the edge of deserts, and from sea level to >3300 m (Swedell, 2011; Butynski *et al.*, 2013; Jolly, 2013). In the savannas and woodlands of eastern Africa, from northeast Sudan southward to Malawi and Zambia, baboons are, second to humans, the most abundant catarrhine species. Several taxa of *Papio* have been described for eastern Africa (Elliot, 1913; Napier & Napier, 1967; Hill, 1970). Their validity and taxonomic ranks are, however, still disputed (Jolly, 1993, 2013; Sarmiento, 1998; Groves, 2001; Grubb *et al.*, 2003; Butynski *et al.*, 2013). The main eastern African forms are olive baboon *Papio anubis* (Lesson, 1827) and yellow baboon *Papio cynocephalus* (Linnaeus, 1766) (figure 1; appendix 1).

Across the distribution of baboons in Africa, morphological clines (Frost *et al.*, 2003; De Jong & Butynski, 2009) and evidence of interspecific hybridization in contact zones (*P. kindae* x *P. ursinus griseipes, e.g.* Jolly *et al.*, 2011; *P. anubis* x *P. cynocephalus, e.g.* Charpentier *et al.*, 2012; *P. hamadryas* x *P. anubis, e.g.* Bergey, 2015) complicate this problem. Taxa are identified primarily by the colour, length, and texture of the pelage (including their mane), body size, body shape, and skull morphology (Hill, 1970; Kingdon, 1971, 2015; Jolly, 1993; Rowe, 1996; Alberts & Altmann, 2001; Groves, 2001; De Jong & Butynski, 2009, 2012; Butynski *et al.,* 2013) (see appendix 1 and visit http://wildsolutions.nl/photomaps/Papio/ to view a large selection of photographs of *Papio* spp. with localities depicted on an interactive digital map). In eastern African baboon populations, comparisons of cranial and dental morphologies indicate that the morphotypes are distinct, but that intermediate forms do exist (Hayes *et al.*, 1990; Frost *et al.*, 2003; Jolly, 2003). Molecular studies, mainly applying mitochondrial (mt) markers, have not

solved this "taxonomic tangle" (Groves, 2001). Phylogenetic reconstructions, based on parts or even complete mt-genomes, reveal several mt-haplogroups or clades. These, however, are only marginally concordant with the morphological variation or taxa (Newman *et al*., 2004; Zinner *et al*., 2009a, 2011, 2013).

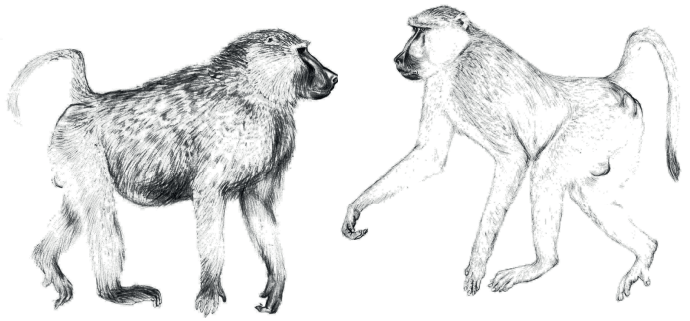


Figure 1. Adult male eastern Africa baboons; olive baboon Papio anubis *(left) and yellow baboon* Papio cynocephalus. *Drawings by Jonathan Kingdon (Kingdon,1971).*

The deepest split in the phylogenetic tree of *Papio* divides the genus into two main mtDNA clades, a northern clade and a southern clade (Burrell, 2008; Zinner *et al*., 2009a, 2013). The northern clade includes Guinea baboon *P. papio* (Desmarest, 1820) from West Africa, hamadryas baboon *P. hamadryas* (Linnaeus, 1758) from the Horn of Africa and southwest Arabia, *P. anubis* from central and northeast Africa, and *P. cynocephalus* from Somalia and southeast Kenya, as well as east and central Tanzania. The southern clade includes *P. cynocephalus* from south Tanzania to north Mozambique, as well as Kinda baboon *P. kindae* (Lönnberg, 1919) from Zambia, and chacma baboon *P. ursinus* (Kerr, 1792) from southern Africa. The two main mtDNA clades come into contact within the distribution of *P. cynocephalus* in central Tanzania, along a line following the Ruaha-Rufiji River from the coast inland westward to the eastern shore of Lake Tanganyika at the mouth of the Malagarasi River (5.25°S, 29.81°E; Zinner *et al*., 2009a). No morphological differences have been reported among *P. cynocephalus* within this contact area. Hill (1970) reports only *P. cynocephalus cynocephalus* for Tanzania (figure 2). Others report *P. anubis* in northwest Tanzania (Groves, 2001; Anandam *et al*., 2013; Butynski *et al*., 2013) and *P. kindae* in central west Tanzania (Butynski & De Jong, 2009; De Jong & Butynski, 2012). Based on present knowledge, the Ibean baboon *P. cynocephalus ibeanus* Thomas, 1893 of Somalia and Kenya, and the Nyasa baboon *P. cynocephalus strepitus* Elliot, 1907 from Malawi (Hill, 1970), are distributed at least 500 km north and south, respectively, from where the two main mtDNA clades meet in central Tanzania.

Although the *Papio* mtDNA phylogeny has been recently intensively studied, only a few samples from eastern Africa have been analysed. In particular, central Tanzania, where the southern and northern clades meet, has not been sampled in detail. To fill this geographic

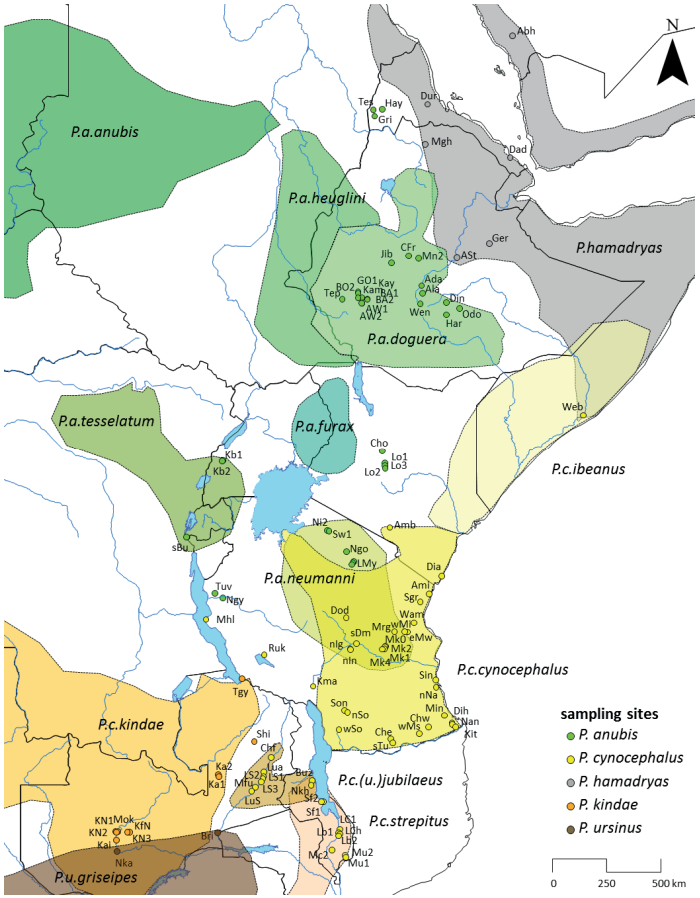


Figure 2. Distribution of Papio *taxa in eastern Africa according to Hill (1970), and locations of sampling sites for this study. A photograph of an adult male baboon from each of 21 widely-spaced sites on this map is presented in appendix 1. More photographs of baboons in the wild from many sites across Africa can be viewed at: http://wildsolutions.nl/photomaps/Papio/*

sampling gap, this study focused on *P. cynocephalus* in Tanzania, Malawi and Zambia, but also includes samples from *P. anubis* and *P. hamadryas* in Eritrea, Ethiopia, Kenya, Uganda, Democratic Republic of Congo (DRC), and northwest Tanzania. The aim of this study is to clarify the geographic distribution of *Papio* mtDNA clades in eastern Africa, and to relate and combine these clades with previous maps of *Papio* distributions (*e.g.* Hill, 1970) and photographs of *Papio* from the region (*e.g.* appendix 1).

MATERIAL AND METHODS

Sample Collection

We obtained 147 non-invasively collected baboon faecal samples from sites in eastern Africa (figure 2, appendix 2). To these we added one *P. hamadryas* tissue sample from southwest Saudi Arabia and one museum tissue sample from Somalia [see Zinner *et al.* (2008) and Kopp *et al.* (2014) for detailed information on these samples]. The samples originate from 104 locations (103 in eastern Africa and one from southwest Saudi Arabia). Of the 149 samples, 45 were used in previous studies, including the samples from Saudi Arabia and Somalia (Zinner *et al.*, 2008, 2009a; Kopp *et al.*, 2014). For those 45 samples we retrieved respective mtDNA sequence information from GenBank. We also included mtDNA sequence information for two southern chacma *P. ursinus ursinus*, two grey-footed chacma baboons *P. ursinus griseipes* Pocock, 1911, two *P. papio,* and, as outgroup, gelada *Theropithecus gelada* (Rüppell, 1835)*,* from GenBank (accession numbers in appendix 2).

Faecal samples were collected and stored following the two-step protocol of Roeder *et al.* (2004) and Nsubuga *et al.* (2004). Samples were stored at ambient temperature for up to 6 months in the field and at -20°C upon arrival in the laboratory. For each sample, consecutive number, date, location, and GPS coordinates were recorded. Sample collection complied with the laws of the respective countries of origin and Germany, and with the guidelines of the International Primatological Society.

Laboratory Work

We extracted DNA from 104 faecal samples using the QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany) and First DNA All Tissue Kit (Gen-Ial, Troisdorf, Germany) according to the manufacturers' protocols. DNA was eluted in 20μl water (HPLC grade) and stored at -20°C until further processing. To avoid contamination, all working steps (*i.e.* DNA extraction, PCR set-up, PCR amplification, gel electrophoresis, gel extraction and sequencing) were performed in separate laboratories. All PCR reactions were performed with negative (HPLC-grade water) controls.

We amplified and sequenced the complete mitochondrial cytochrome *b* gene (cyt *b*, 1140 bp). This allowed us to include published data in the statistical analyses. In addition, this marker can be reliably amplified from low quality samples, such as faecal material and tissue samples from museum specimens. We used established protocols (Zinner *et al.*, 2009a) and amplified cyt *b* via two over-lapping fragments to ensure that sequences were obtained even if DNA was degraded. Zinner *et al.* (2009a) showed that the primers and PCR conditions applied here solely amplify mtDNA and not nuclear mitochondrial pseudogenes (numts). PCR reactions with a total volume of 30 μl included 1 U BiothermTaq 5000 DNA polymerase (Genecraft, Cologne, Germany), 1x reaction buffer, 0.16 mM dNTPs, and 0.33µM of each primer. PCR conditions for amplification comprised a pre-denaturation step at 94°C for 2 min, followed by

40 cycles at 94°C for 1 min, 60°C for 1 min, 72°C for 1 min, and a final extension step at 72°C for 5 min. Results of the PCR amplifications were checked on 1% agarose gels. PCR products were cleaned with the Qiagen PCR Purification Kit and subsequently sequenced on an ABI 3130xL sequencer using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, USA) and both amplification primers.

Phylogenetic Reconstruction

Sequences were checked, edited, and aligned manually using BIOEDIT 7.5.0.2 (Hall, 1999). The final alignment comprised 156 sequences of which 149 derived from eastern African and Saudi Arabian baboons, two sequences each from *P. papio*, *P. u. ursinus,* and *P. u. griseipes,* and as outgroup, one sequence from *T. gelada*. All sequences were deposited in GenBank (details of samples and accession numbers are given in Table S1).

Phylogenetic trees were reconstructed with maximum-likelihood (ML) and Bayesian methods using RAxML 8 (Stamatakis, 2014) as implemented in raxmlGUI 1.5b.1 (Silvestro & Michalak, 2012) and MrBayes 3.2.6 (Ronquist *et al.,* 2012), respectively. ML calculations were run with the GTR-CAT-I model and 1,000 rapid bootstrapping replications. For Bayesian tree reconstructions, we conducted four Markov Chain Monte Carlo (MCMC) runs with a default temperature of 0.2 and the HKY+I+G model as selected as best-fit model in jModeltest 2.1.7 (Darriba *et al.,* 2012) under the Bayesian information criterion (BIC) and the Decision Theory Performance-based Selection (DT). All repetitions were run for 1 million generations with tree and parameter sampling setting in every 100 generations. The first 25% of samples were discarded as burn-in. The adequacy of the burn-in and convergence of all parameters was assessed via the uncorrected Potential Scale Reduction Factor (PSRF) (Gelman & Rubin, 1992) as calculated by MrBayes and by visual inspection of the trace of the parameters across generations using TRACER 1.6 (Rambaut *et al.,* 2014). To check whether posterior clade probabilities were also converging, AWTY (Nylander *et al.*, 2008) was applied. Posterior probabilities for each split and a phylogram with mean branch lengths were calculated from the posterior density of trees. Trees were visualized and edited in FigTree 1.4.2. Additionally, a haplotype network was built in POPART 1.7 (Leigh & Bryant, 2015) using the median-joining network algorithm (Bandelt *et al.*, 1999).

RESULTS

Among the 149 eastern African and Arabian baboon sequences, we detected 69 unique cyt *b* haplotypes. Phylogenetic analyses were conducted with these unique sequences, plus the additional six non-eastern African sequences and the outgroup sequence. The final alignment comprised 76 sequences. Phylogenetic trees derived from Bayesian inference and the ML algorithm yielded identical tree topologies and mostly well-supported terminal clades (ML bootstrap values: >75%, Bayesian posterior probabilities: >0.95) (figure 3). The analysis suggests that baboons initially divided into four clades [southern chacma, Mahale Mts (central west Tanania), main southern, main northern], although the branching pattern among them remains largely unknown. The main southern clade contains Kinda, northern chacma, and southern yellow baboons, while the main northern clade comprises Guinea, hamadryas, eastern olive, and northern yellow baboons.

For eastern Africa, our reconstruction shows nine well-supported mt-clades or haplogroups. These indicate paraphyletic or polyphyletic relationships within at least the two main East African

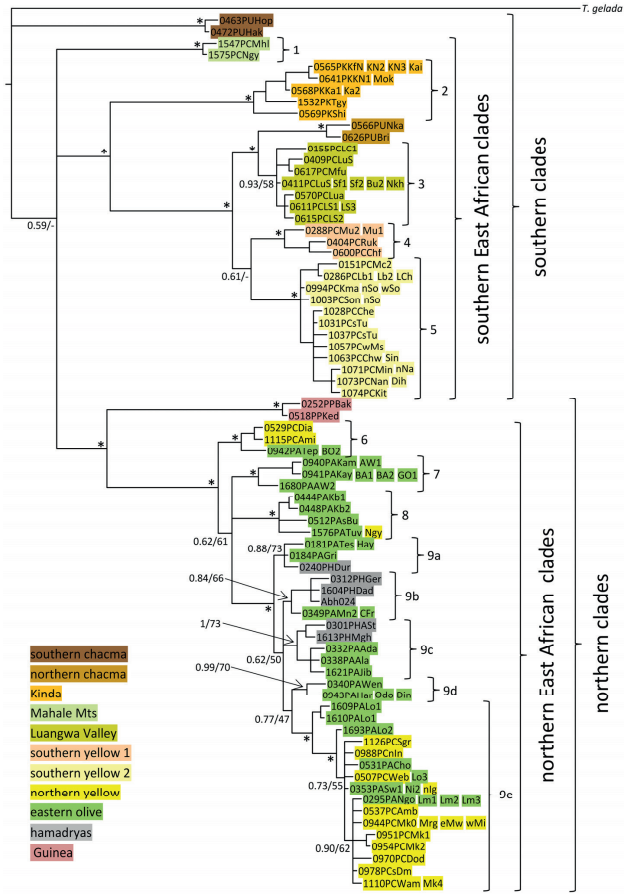


Figure 3. Bayesian phylogram of Papio *cyt b haplotypes (Bayesian tree reconstruction for identification of clades) (* Bayesian posterior probabilities > 0.95, ML bootstrap support values >75%). Tip labels refer to the species (as identified by phenotypic characters; PA =* P. anubis; *PC =* P. cynocephalus; *PH =* P. hamadryas; *PK =* P. kindae; *PP =* P. papio; *PU =* P. ursinus) *together with the sampling location. Several haplotypes were found at more than one site. This is indicated by the listing of more than one site abbreviation at some branch tips.*

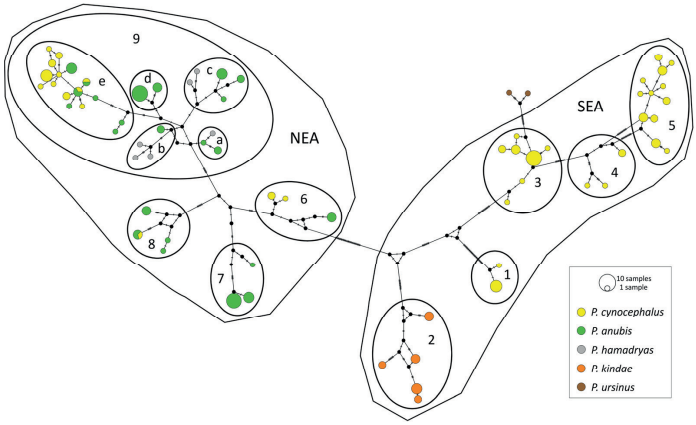


Figure 4. Median-joining mt-haplotype network of eastern African baboons (1–9 = haplogroups or clades of East African baboon as in figure 3; SEA = main southern East African clade; NEA = main northern East African clade). Papio ursinus *are northern chacma baboons from Zambia. These appear to be* P. u. griseipes.

baboon taxa (*P. cynocephalus* and *P. anubis*). The nine mt-clades (figures 3 & 4) are: (1) Mahale Mts baboons; (2) Kinda baboons; (3) Luangwa Valley (Zambia) and west Malawi baboons (with northern chacma baboons as a sister clade); (4) southern yellow baboons 1 from southwest Tanzania, north Zambia, and south Malawi; (5) southern yellow baboons 2 from southeast Tanzania and south Malawi east of the Shire River; (6) yellow baboons from northeast Tanzania and olive baboons from central Ethiopia; (7) olive baboons from central Ethiopia; (8) olive baboons from Uganda, east DRC and west Tanzania, and one yellow baboon from the Ugalla region; and (9) hamadryas baboons from Saudi Arabia, Eritrea, and Ethiopia, and eastern and northern clades of olive baboons and yellow baboons from Eritrea, Ethiopia, Somalia, Kenya, and north Tanzania. Clades 2–5 represent the main southern East African clade, while clades 6–9 represent the main northern East African clade. Within clade 9, distinctions can be made among possible clades 9a to 9e. Hamadryas baboons cluster with olive baboons from Eritrea (9a) and Ethiopia (9b and 9c). Olive baboons from east of the Ethiopian Rift Valley form another clade (9d). Northern yellow baboons from Somalia, Kenya and north Tanzania cluster with olive baboons from Kenya and north Tanzania (9e).

The nine mt-clades are also visible within the haplotype network (figure 4), and the subdivision within clade 9 becomes even more obvious than in the tree reconstruction. Among the northern East African clades, olive baboon haplotypes appear in every clade—usually in combination with yellow baboon or hamadryas baboon haplotypes. Interestingly, hamadryas and yellow baboons do not share the same clade or subclade.

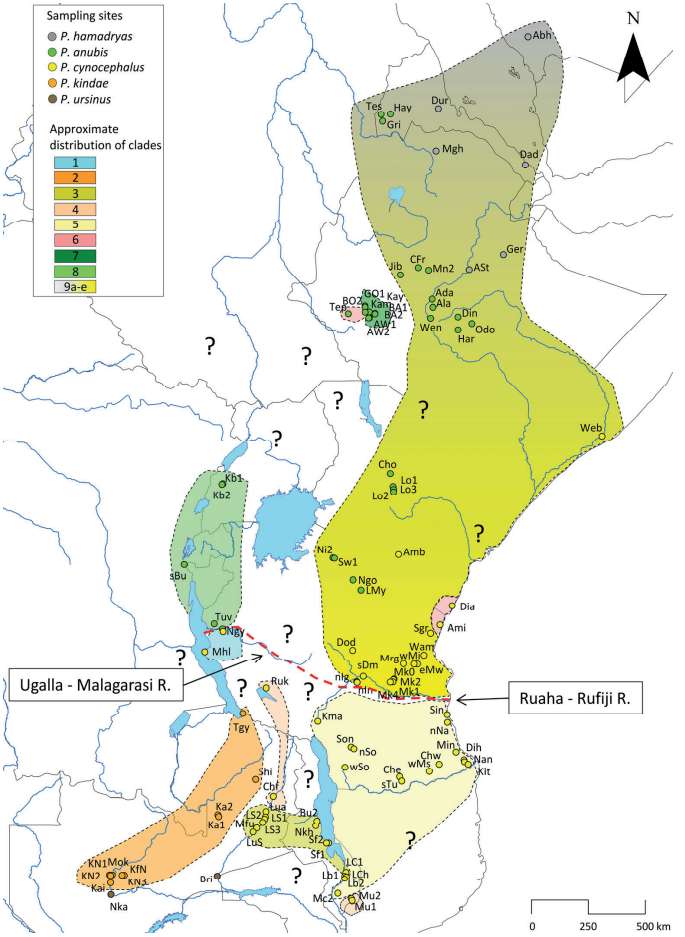


Figure 5. Approximate distribution of the nine Papio *mt-haplogroups or clades within eastern Africa and sampling sites. Broken red line depicts the approximate position of the boundary between the distributions of northern East African and southern East African mt-clades. "?" indicates an important geographic sampling gap (see Discussion).*

Identical mt-haplotypes were mainly found in samples collected at the same locality (same social group or neighbouring groups), but identical sequences were found at different sites and in different taxa. Median distance between sites with identical sequences was 63 km (quartiles 39–121; n = 49 sites with distance > 10 km). Largest distances among sites (figures 2 & 5) with identical haplotypes were between Webi Shebelli River, south Somalia (Web, yellow baboon) and Lolldaiga Hills Ranch, central Kenya (Lo3, olive baboon; *ca*. 950 km), and between Serengeti National Park, north Tanzania (Ni2, olive baboon) and Dodoma-Iringa, central Tanzania (nIg, yellow baboon; *ca*. 550 km).

DISCUSSION

The general topology of our phylogenetic tree for *Papio* in eastern Africa is similar to those found in earlier studies (Zinner *et al.,* 2009a, 2011, 2013; Liedigk *et al.,* 2014).

There are unresolved relationships between the two main East African clades, and with the southern chacma and Mahale Mts clades (figures 3 & 4). The major split separates southern and northern clades. This is the first study to reveal the Mahale Mts clade (as no samples from Mahale Mts were included in previous studies).

We confirm the previously assumed boundary between the distributions of southern and northern clades in Tanzania (figure 5; Zinner *et al.,* 2011). This boundary runs from the lower Rufiji River in the east across central Tanzania to the mouth of the Malagarasi River at Lake Tanganyika in the west. These two large rivers are biogeographic barriers, or at least boundaries, for a number of genera, species, and subspecies of terrestrial mammals (Kingdon, 1971–1982, 2015; Kingdon *et al*., 2013; Butynski & De Jong, in press). As concerns primates:

- The Ugalla-Malagarasi River appears to be a barrier between *P. anubis* (north) and *P. cynocephalus* (south) (Kano, 1971).
- The Ruaha-Rufiji River represents the south limit for colobus monkeys (subfamily Colobinae) in coastal eastern Africa (Rodgers, 1981; Butynski *et al*., 2013; Butynski & De Jong, in press).
- The Ruaha-Rufiji River seems to be the south limit in coastal Tanzania for the Zanzibar dwarf galago *Galagoides zanzibaricus* (Matschie, 1893) (Butynski *et al*., 2006, 2013; De Jong, 2012; Butynski & De Jong, in press).
- The Ruaha-Rufiji River appears to be the north limit for the Mozambique dwarf galago *Galagoides granti* (Thomas & Wroughton, 1907) (Butynski *et al*., 2006, 2013; De Jong, 2012; Butynski & De Jong, in press).

Our phylogenetic reconstruction and the haplotype network reveal nine reasonably well-supported mt-haplogroups or clades within eastern African baboons. Tree topology and statistical support values suggest five clades for southern East African baboons: Mahale Mts, Kinda, Luangwa Valley, and two southern yellow clades. Luangwa Valley baboons form the sister group of northern chacmas, whereas southern chacmas are either basal among baboons or represent the first split in the southern baboon clade (Zinner *et al.,* 2009a, 2013). For the main northern East African clade, the situation looks different. Here we found only three well-supported haplogroups, but a diverse clade consisting of hamadryas, eastern olive, and northern yellow baboons stretching from Eritrea and southwest Saudi Arabia south to the Ruaha-Rufiji River in central Tanzania (clade 9). Within the distribution of clade 9, identical

haplotypes occur at sites that are almost 1,000 km apart and in different baboon species (*e.g.* yellow and olive baboons [this study], and olive and hamadryas baboons [Hapke *et al.,* 2001]). In general, the distribution pattern of baboon haplotypes appears to be in a geographic cline—irrespective of their taxonomic relationships.

Mapping the approximate distributions of the mt-haplogroups reveals some disjunct distributions. Among southern East African baboons, clade 4 was found in southeast Tanzania and northeast Zambia, but also east of the Shire River in south Malawi. Between these two distributions, there are haplotypes belonging to clade 3 (Luangwa Valley baboons). Among northern East African baboons, haplotypes of clade 6 occur in yellow baboons at the south Kenyan and north Tanzanian coast, and in olive baboons in west Ethiopia. A possible explanation for such disjunct distributions is that they are relicts of a former wider distribution; the respective haplotypes are now extinct, having been replaced by haplotypes of other clades or haplogroups in parts of the former distribution. Alternatively, the disjunct pattern is a result of incomplete geographic sampling. Increasing the number of samples from areas in between the two distributions might resolve this question.

As in previous studies, the mt-clades only partly match with recognized baboon taxa, most obvious in the case of southern and northern yellow baboons. Their mt-genomes diverged around 2 million years ago (Zinner *et al.,* 2013), though no taxonomic differentiation was reported among yellow baboons from northeast and southeast Tanzania; both are regarded as *P. c. cynocephalus*. Several of the clades, however, occur in areas where Hill (1970) identified different morphotypes or subspecies of baboons, *e.g.* Luangwa Valley baboon *P. c. (u.) jubilaeus* Schwarz, 1928 and central olive baboon *P. a. tesselatum* Elliot, 1909 (figure 2, appendix 1). Whether *P. c. strepitus* Elliot, 1907 is at least partly equivalent with our clade 4 needs further investigation. MtDNA information is often used to delimit species within taxonomic groups (*e.g.* barcoding; Blaxter, 2004; Galimberti *et al.*, 2015; Raupach *et al.,* 2016). This study, however, indicates that (at least) the two main East African baboon taxa are paraphyletic or polyphyletic groups. As such, mtDNA information alone cannot delimit species within *Papio* in East Africa.

The geographic pattern of mt-haplogroup distributions suggests a complicated biogeographic history for baboons. Like other savanna-living mammals (Lorenzen *et al.,* 2012), baboons were impacted by multiple cycles of expansion and retreat of savanna biomes during Pleistocene glacial and inter-glacial periods, and other climatic changes (Maslin *et al.,* 2015; Trauth *et al.,* 2015). Recurrent fragmentation and reconnection of populations, extinctions, and distribution shifts of demes likely led to multiple phases of isolation, hybridization, and introgression among populations (Zinner *et al.,* 2009a, 2011).

The geographic distribution pattern of mt-haplogroups can provide insights into the biogeographic history of taxa (Avise, 2000, 2004). We found better supported clades, and geographically more confined clades, among southern East African baboons than among northern East African baboons. This suggests that southern East African baboons were affected more by ecological change, leading to longer periods of isolation (thereby hampering movement among demes and/or dispersal into new areas), than northern East Africa baboons. Northern East African baboons, on-the-other-hand, seem to have experienced more introgression due to more frequent reconnection of populations. Here the possible introgression of olive baboons into hamadryas and yellow baboons might be of particular interest (see Zinner *et al.,* 2011 for a more detailed scenario). There is some evidence that ancient introgression occurred even beyond the species level, as found in kipunji *Rungwecebus kipunji* (Ehardt, Butynski, Jones & Davenport, 2005) (Jones *et al.*, 2005).

Individuals from the Mt Rungwe population carry mitochondrial DNA-sequences that are highly similar to those in south Tanzanian yellow baboons (Zinner *et al.*, 2009b), whereas the Udzungwa Mts population seems not to be affected by baboon introgression (Roberts *et al.*, 2010).

It also might be that at least part of the difference between southern East African and northern East African populations is the result of incomplete geographic sampling. It is, therefore, important to collect more samples in under-sampled areas: (1) central and west Kenya, easten and central Uganda, south South Sudan into Ethiopia; (2) Somalia, north coast of Kenya; (3) west Uganda, east DRC; (4) southwest Tanzania; (5) north Mozambique, east of Lake Malawi, and (6) Mahale Mts, Ugalla, east DRC west of Lake Tanganyika. The latter is of particular interest since the Mahale Mts baboons share morphological traits with Kinda baboons. Here are two scenarios for this similarity: (1) Kinda baboons once occurred east of Lake Tanganyika as far north as Mahale Mts; (2) during times of extreme lake level lowstands a land bridge connected the east and west shores of Lake Tanganyika (Scholz & Rosendahl, 1988; Lezzar *et al.*, 1996; Cohen *et al.*, 1997; Nevado *et al.*, 2013) that enabled Kinda baboons from the west side to come into contact with yellow baboons from the east side—and the possibility of gene flow.

CONCLUSIONS

This study provides further insights into the evolutionary history of eastern African baboons and confirms the Ugalla-Malagarasi River and Ruaha-Rufiji River as the boundary between the main southern East African and main northern East African baboon clades. Subclades are obvious in both of these main clades. Interestingly, the distribution of the subclades seems to be more geographically structured in southern East African baboons, while a more clinal pattern is evident in northern East African baboons. A possible explanation is that southern East African baboons were historically affected more by population division due to changing ecological conditions, while northern East African baboons were affected more by frequent reconnection and gene flow among populations.

Although we obtained new information and insights on the phylogeographic pattern of baboons in eastern Africa, there are several topics that remain to be addressed. For instance, the disjunct distribution of several subclades could be a relict of a former wide distribution, but the possibility that the distribution pattern of subclades is due to incomplete geographic sampling cannot be ruled out. Likewise, it remains unknown (1) how Mahale Mts baboons are related to Kinda baboons on the west shore of Lake Tanganyika; (2) which yellow baboon haplotypes occur around Lake Malawi, particularly in Mozambique, and (3) how baboons in central and west central Ethiopia are related to baboons in Uganda and DRC. To address these topics, additional sampling is required. Further, since gene flow and hybridization among baboon populations most likely occurred repeatedly, nuclear DNA information, ideally nuclear genome data, are required to better reconstruct the evolutionary history and phylogeography of *Papio*. In combination with baboon genomics, spatial and ecological modelling of historic and current distributions of *Papio* taxa and clades will provide insights into habitat preferences (niches) and adaptations.

ACKNOWLEDGMENTS

We thank the national administrations of countries in which samples were collected. We are particular grateful to the University of Dodoma, TAWIRI, and COSTECH in Tanzania (Research Permit issued 24th September 2008, 2008-272-NA-2008-64), and ZAWA in Zambia (Research Permit issued 13th July 2007, Export Certificate 212007007240). We are grateful to the many colleagues and field assistants who helped collect the samples, to Lorna Depew and two anonymous reviewers for their comments on the manuscript, and to Hakan Pohlstrand, Chris Roche, Katerina Guschanski, Ian Salisbury, Dean Gaffigan, Celesta von Chamier, Daniel Cara, Colleen Begg, Anna Weyher, Jim Auburn, and Mike Haworth for the use of their photographs in appendix 1. Jonathan Kingdon kindly allowed the reproduction of his drawings of *Papio anubis* and *Papio cynocephalus* in figure 1. This paper is dedicated to Jonathan Kingdon on the event of his 80th birthday and in gratitude for all that he has accomplished on behalf of the conservation of East Africa's mammals.

REFERENCES

Alberts, S.C. & J. Altmann (2001). Immigration and hybridization patterns of yellow and anubis baboons in and around Amboseli, Kenya. *American Journal of Primatology* **53**: 139–154.

Anandam, M.V., E.L. Bennett, T.R.B. Davenport, N.J. Davies, K.M. Detwiler, A. Engelhardt, A.A. Eudey, E.L. Gadsby, C.P. Groves, A. Healy, K.P. Karanth, S. Molur, T. Nadler, M.C. Richardson, E.P. Riley, C. Roos, A.B. Rylands, L.K. Sheeran, N. Ting, J. Wallis, S.S. Waters, D.J. Whittaker & D. Zinner (2013). Species accounts of Cercopithecidae. In R.A. Mittermeier, A.B. Rylands & D.E. Wilson (eds.), *Handbook of the Mammals of the World. Vol. 3. Primates*. Lynx Editions, Barcelona. Pp. 628–753.

Avise, J.C. (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Avise, J.C. (2004). *Molecular Markers, Natural History, and Evolution*. Sinauer Associates, Sunderland, Massachusetts.

Bandelt, H.J., P. Forster & A. Rohl (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37–48.

Bergey, C.M. (2015). Population Genomics of a Baboon Hybrid Zone. PhD thesis, New York University, New York.

Blaxter, M.L. (2004). The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**: 669–679. DOI: 10.1098/rstb.2003.1447

Burrell, A.S. (2008). Phylogenetics and Population Genetics of Central African Baboons. PhD thesis, New York University, New York.

Butynski, T.M. & Y.A. de Jong (2009). *Primates of Mahale Mountains National Park, Tanzania*. Nomad, Arusha, Tanzania. DOI: 10.13140/RG.2.1.1654.4403

Butynski, T.M. & Y.A. De Jong (in press). Primates of Africa's coastal deltas and their conservation. In A.A. Barnett, I. Matsuda & K. Nowak (eds.), *Primates in Flooded Habitats: Ecology and Conservation*. Cambridge University Press, Cambridge, UK.

Butynski, T.M., Y.A. de Jong, A.W. Perkin, S.K. Bearder & P.E. Honess (2006). Taxonomy, distribution, and conservation status of three species of dwarf galagos (*Galagoides*) in eastern Africa. *Primate Conservation* **21**: 63–79. DOI: 10.1896/0898-6207.21.1.63

Butynski, T.M., J. Kingdon & J. Kalina (eds.) (2013). *Mammals of Africa. Vol. II. Primates*. Bloomsbury, London.

Charpentier, M.J.E., M.C. Fontaine, E. Cherel, J.P. Renoult, T. Jenkins, L. Benoit, N. Barthès, S.C. Alberts & J. Tung (2012). Genetic structure in a dynamic baboon hybrid zone corroborates behavioural observations in a hybrid population. *Molecular Ecology* **21**: 715–731. DOI: 10.1111/j.1365-294X.2011.05302.x

Cohen, A.S., K.E. Lezzar, J.J. Tiercelin & M. Soreghan (1997). New palaeogeographic and lake-level reconstructions of Lake Tanganyika: implications for tectonic, climatic and biological evolution in a rift lake. *Basin Research* **9**: 107-132. DOI: 10.1046/j.1365-2117.1997.00038.x

Darriba, D., G.L. Taboada, R. Doallo & D. Posada (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772. DOI: 10.1038/nmeth.2109

De Jong, Y.A. (2012). Taxonomy, Diversity, Biogeography and Conservation of the Primates of Kenya and Tanzania. PhD thesis, Oxford Brookes University, Oxford, UK.

De Jong, Y.A. & T.M. Butynski (2009). Primate Biogeography, Diversity, Taxonomy and Conservation of the Coastal Forests of Kenya. Unpublished report to the Critical Ecosystem Partnership Fund and Eastern Africa Primate Diversity and Conservation Program, Nanyuki, Kenya. DOI: 10.13140/RG.2.1.5062.3128

De Jong, Y.A. & T.M. Butynski (2012). The primates of East Africa: country lists and conservation priorities. *African Primates* **7**: 135–155. DOI: 10.1896/052.029.0107

Elliot, D.G. (1913). *A Review of the Primates. Vol. 2. Anthropoidea* (Aotus *to* Lasiopyga). Monograph Series, American Museum of Natural History, New York.

Frost, S.R., L.F. Marcus, F.L. Bookstein, D.P. Reddy & E. Delson (2003). Cranial allometry, phylogeography, and systematics of large-bodied papionins (Primates: Cercopithecinae) inferred from geometric morphometric analysis of landmark data. *The Anatomical Record. Part A. Discoveries in Molecular, Cellular, and Evolutionary Biology* **275A**: 1048–1072. DOI: 10.1002/ar.a.10112

Galimberti, A, A. Sandionigi, A. Bruno, A. Bellati & M. Casiraghi (2015). DNA barcoding in mammals: what's new and where next? *Hystrix* **26**: 13–14. DOI: 10.4404/hystrix-26.1-11347

Gelman, A. & D.B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* **7**: 457–472. http://www.jstor.org/stable/2246093

Groves, C.P. (2001). *Primate Taxonomy*. Smithsonian Institution Press, Washington, DC.

Grubb, P., T.M. Butynski, J.F. Oates, S.K. Bearder, T.R. Disotell, C.P. Groves & T.T. Struhsaker (2003). Assessment of the diversity of African primates. *International Journal of Primatology* **24**: 1301–1357. DOI: 10.1023/B:IJOP.0000005994.86792.b9

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium* **41**: 95–98.

Hapke, A., D. Zinner & H. Zischler (2001). Mitochondrial DNA variation in Eritrean hamadryas baboons (*Papio hamadryas hamadryas*): life history influences population genetic structure. *Behavioral Ecology and Sociobiology* **50**: 483–492. DOI: 10.1007/s002650100393

Hayes, V.J., L. Freedman & C.E. Oxnard (1990). The taxonomy of savannah baboons: an odontomorphometric analysis. *American Journal of Primatology* **22**: 171–190. DOI: 10.1002/ajp.1350220304

Hill, W.C.O. (1970). *Primates: Comparative Anatomy and Taxonomy. Vol. 8. Cynopithecinae:* Papio, Mandrillus, Theropithecus. Edinburgh University Press, Edinburgh, Scotland.

Jolly, C.J. (1993). Species, subspecies, and baboon systematics. In W.H. Kimbel & L.B. Martin (eds.), *Species, Species Concepts, and Primate Evolution*. Plenum Press, New York. Pp. 67–101.

Jolly, C.J. (2003). Cranial anatomy and baboon diversity. *The Anatomical Record. Part A. Discoveries in Molecular, Cellular, and Evolutionary Biology* **275A**: 1043–1047. DOI: 10.1002/ar.a.10121

Jolly, C.J. (2013). Genus *Papio* baboons. In T.M. Butynski, J. Kingdon & J. Kalina (eds.), *Mammals of Africa. Vol II. Primates*. Bloomsbury, London. Pp. 217–218.

Jolly, C.J., A.S. Burrell, J.E. Phillips-Conroy, C.M. Bergey & J. Rogers (2011). Kinda baboons (*Papio kindae*) and grayfoot chacma baboons (*P. ursinus griseipes*) hybridize in the Kafue river valley, Zambia. *American Journal of Primatology* **73**: 291–303. DOI: 10.1002/ajp.20896

Jones, T., C.L. Ehardt, T.M. Butynski, T.R.B. Davenport, N.E. Mpunga, S.J. Machaga & D.W. De Luca (2005). The highland mangabey *Lophocebus kipunji*: a new species of African monkey. *Science* **308**: 1161–1164. DOI: 10.1126/science.1109191

Kano, T. (1971). Distribution of the primates on the eastern shore of Lake Tanganyika. *Primates* **12**: 281–304. DOI: 10.1007/BF01730416

Kingdon, J. (1971–1982). *East African Mammals: An Atlas of Evolution in Africa* (7 vols.). The University of Chicago Press, Chicago.

Kingdon, J. (1971). *East African Mammals: An Atlas of Evolution in Africa Vol. 1. Primates*. The University of Chicago Press, Chicago.

Kingdon, J. (2015). *The Kingdon Field Guide to African Mammals* (2nd ed.). Academic Press, London.

Kingdon, J., D. Happold, T. Butynski, M. Hoffmann, M. Happold & J. Kalina (eds.). (2013). *Mammals of Africa* (6 vols.). Bloomsbury, London.

Kopp, G.H., C. Roos, T.M. Butynski, D.E. Wildman, A.N. Alagaili, L.F. Groeneveld & D. Zinner (2014). Out of Africa, but how and when? The case of hamadryas baboons (*Papio hamadryas*). *Journal of Human Evolution* **76**: 154–164. DOI: 10.1016/j.jhevol.2014.08.003

Leigh, J.W. & D. Bryant (2015). POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**: 1110–1116. DOI: 10.1111/2041-210X.12410

Lezzar, K.E., J.J. Tiercelin, M. De Batist, A.S. Cohen, T. Bandora, P. Van Rensbergen, C. Le Turdu, W. Mifundu & J. Klerkx (1996). New seismic stratigraphy and Late Tertiary history of the North Tanganyika Basin, East African Rift system, deduced from multichannel and high-resolution reflection seismic data and piston core evidence. *Basin Research* **8**: 1–28. DOI: 10.1111/j.1365-2117.1996.tb00112.x

Liedigk, R., C. Roos, M. Brameier & D. Zinner (2014). Mitogenomics of the Old World monkey tribe Papionini. *BioMed Central: Evolutionary Biology* **14**: 176. DOI: 10.1186/s12862-014-0176-1

Lorenzen, E.D., H. Heller & H.R. Siegismund (2012). Comparative phylogeography of African savannah ungulates. *Molecular Ecology* **21**: 3656–3670. DOI: 10.1111/j.1365-294X.2012.05650.x

Maslin, M.A., S. Shultz & M.H. Trauth (2015). A synthesis of the theories and concepts of early human evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **370**. DOI: 10.1098/rstb.2014.0064

Napier, J.R. & P.H. Napier (1967). *A Handbook of Living Primates*. Academic Press, London.

Nevado, B., S. Mautner, C. Sturmbauer & E. Verheyen (2013). Water-level fluctuations and metapopulation dynamics as drivers of genetic diversity in populations of three Tanganyikan cichlid fish species. *Molecular Ecology* **22**: 3933–3948. DOI: 10.1111/mec.12374

Newman, T.K., C.J. Jolly & J. Rogers (2004). Mitochondrial phylogeny and systematics of baboons (*Papio*). *American Journal of Physical Anthropology* **124**: 17–27. DOI: 10.1002/ajpa.10340

Nsubuga, A.M., M.M. Robbins, A.D. Roeder, P.A. Morin, C. Boesch & L.Vigilant (2004). Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. *Molecular Ecology* **13**: 2089–2094. DOI: 10.1111/j.1365-294X.2004.02207.x

Nylander, J.A.A., J.C. Wilgenbusch, D.L. Warren & D.L. Swofford (2008). AWTY (are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**: 581–583. DOI: 10.1093/bioinformatics/btm388

Rambaut, A., M.A. Suchard, D. Xie & A.J. Drummond (2014). Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer

Raupach, M.J., R. Amann, Q.D. Wheeler & C. Roos (2016). The application of "-omics" technologies for the classification and identification of animals. *Organisms Diversity & Evolution* **16**: 1–12. DOI: 10.1007/s13127-015-0234-6

Roberts, T.E., T.R.B. Davenport, K.B.P. Hildebrandt, T. Jones, W.T. Stanley, E.J. Sargis & L.E. Olson (2010). The biogeography of introgression in the critically endangered African monkey *Rungwecebus kipunji*. *Biology Letters* **6**: 233–237. DOI: 10.1098/rsbl.2009.0741

Roeder, A.D., F.I. Archer, H.N. Poiner & P.A. Morin (2004). A novel method for collection and preservation of faeces for genetic studies. *Molecular Ecology Notes* **4**: 761–764. DOI: 10.1111/j.1471-8286.2004.00737.x

Rodgers, W.A. (1981). The distribution and conservation status of colobus monkeys in Tanzania. *Primates* **22**: 33–45. DOI: 10.1111/j.1471-8286.2004.00737.x

Ronquist, F., M. Teslenko, P. van der Mark, D. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard & J.P. Huelsenbeck (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**: 539–542. DOI: 10.1093/sysbio/sys029

Rowe, N. (1996). *The Pictorial Guide to the Living Primates*. Pogonias Press, East Hampton, New York.

Sarmiento, E. (1998). Current problems with *Papio* taxonomies. *African Primates* **3**: 48–52.

Scholz, C.A. & B.R. Rosendahl (1988). Low lake stands in Lakes Malawi and Tanganyika, East Africa, delineated with multifold seismic data. *Science* **240**: 1645–1648. DOI: 10.1126/science.240.4859.1645

Silvestro, D. & I. Michalak (2012). raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution* **12**: 335–337. DOI: 10.1007/s13127-011-0056-0

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313. DOI: 10.1093/bioinformatics/btu033

Swedell, L. (2011). African papionins: diversity of social organization and ecological flexibility. In C.J. Campbell, A. Fuentes, K.C. MacKinnon, S.K. Bearder & R.M. Stumpf (eds.), *Primates in Perspective*. Oxford University Press, New York. Pp. 241–277.

Trauth, M.H., A. Bergner, V. Foerster, A. Junginger, M.A. Maslin & F. Schaebitz (2015). Episodes of environmental stability vs. instability in Late Cenozoic lake records of eastern Africa. *Journal of Human Evolution* **87**: 21–31. DOI: 10.1016/j.jhevol .2015.03.011

Zinner, D., R. Kraft & C. Roos (2008). The phylogenetic position of "*Papio ruhei*" – a unique baboon taxon from Somalia? *Der Zoologische Garten (Neue Folge)* **77**: 303–311. DOI: 10.1016/j.zoolgart.2008.05.002

Zinner, D., L.F. Groeneveld, C. Keller & C. Roos (2009a). Mitochondrial phylogeography of baboons (*Papio* spp.) - Indication for introgressive hybridization? *BioMed Central: Evolutionary Biology* **9**: 83. DOI: 10.1186/1471-2148-9-83

Zinner, D., M.L. Arnold & C. Roos (2009b). Is the new primate genus *Rungwecebus* a baboon? *PLoS ONE* **4** : e4859. DOI:10.1371/journal.pone.0004859

Zinner, D., U. Buba, S. Nash & C. Roos (2011). Pan-African voyagers. The phylogeography of baboons. In V. Sommer & C. Ross (eds.), *Primates of Gashaka: Socioecology and Conservation in Nigeria's Biodiversity Hotspot*. *Developments in Primatology: Progress and Prospects* **35**: 267–306.

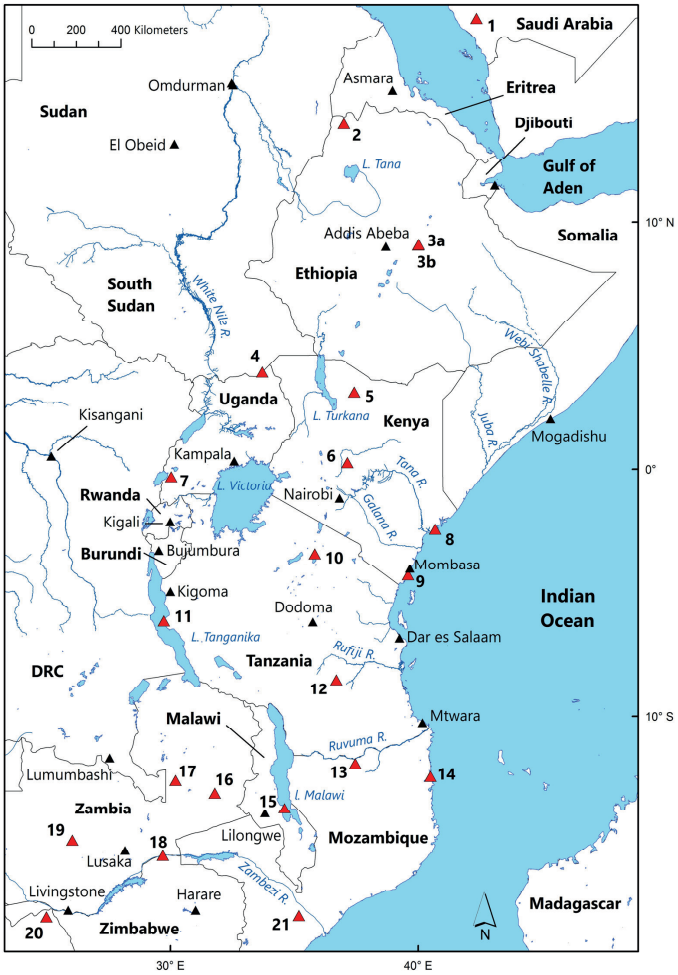
Zinner, D., J. Wertheimer, R. Liedigk, L.F. Groeneveld & C. Roos (2013). Baboon phylogeny as inferred from complete mitochondrial genomes. *American Journal of Physical Anthropology* **150**: 133–140. DOI: 10.1002/ajpa.22185

Appendix 1a. Supporting information for 22 photographs of adult male baboons from 21 sites in eastern Africa. The baboons in these 22 photographs represent five Papio species *and 11* Papio *subspecies. The taxonomy followed here is that of Hill (1970), except for the Kinda baboon, which is here recognized as species* P. kindae. *Numbers in the first column correspond to the numbers on the map (appendix 1b) and with the photographs (appendix 1c). More than 650 photographs of baboons in the wild from many sites across Africa can be viewed at: http://wildsolutions.nl/photomaps/Papio/*

Number	Locality	Taxon	mt-clade	Photographer(s)
1	Raidah Forest, Saudi Arabia	*Papio hamadryas* (Linnaeus, 1758)	9a or 9b	Tom Butynski & Yvonne de Jong
2	Kafta Sheraro National Park, Ethiopia	*Papio anubis heuglini* Matschie, 1898	9a	Hakan Pohlstrand
3a	Awash National Park, Ethiopia	*Papio anubis doguera* (Pucheran & Schimper, 1856)	9c	Hakan Pohlstrand
3b	Awash National Park, Ethiopia	*Papio hamadryas* (Linnaeus, 1758)	9c	Chris Roche
4	Kidepo Valley National Park, Uganda	*Papio anubis furax* Elliot, 1907	?	Yvonne de Jong & Tom Butynski
5	Kalacha, Kenya	*Papio anubis furax* Elliot, 1907	?	Yvonne de Jong & Tom Butynski
6	Lolldaiga Hills Ranch, Kenya	*Papio anubis neumanni* Matschie, 1897	9e	Yvonne de Jong & Tom Butynski
7	Kigezi Game Reserve, Uganda	*Papio anubis tesselatum* Elliot, 1909	8	Katerina Guschanski
8	Mpeketoni, Kenya	*Papio cynocephalus ibeanus* Thomas, 1893	?	Yvonne de Jong & Tom Butynski
9	Diani, Kenya	*Papio cynocephalus cynocephalus* (Linnaeus, 1766)	6	Yvonne de Jong & Tom Butynski
10	Lake Manyara National Park, Tanzania	*Papio anubis neumanni* Matschie, 1897	9e	Tom Butynski & Yvonne de Jong
11	Mahale Mountains National Park, Tanzania	*Papio c.f. kindae* Lönnberg, 1919	1	Yvonne de Jong & Tom Butynski
12	Mahenge, Tanzania	*Papio cynocephalus cynocephalus* (Linnaeus, 1766)	5	Yvonne de Jong & Tom Butynski
13	Niassa National Reserve, Mozambique	*Papio cynocephalus* ssp.?	?	Colleen Begg
14	Quirimbas National Park, Mozambique	*Papio cynocephalus* ssp.?	?	Daniel Cara
15	Senga Hills Forest Reserve, Malawi	*Papio cynocephalus* (*u.*) *jubilaeus* Schwarz, 1928	3	Jim Auburn
16	Mfuwe, South Luangwa National Park, Zambia	*Papio cynocephalus* (*u.*) *jubilaeus* Schwarz, 1928	3	Ian Salisbury
17	Kasanka National Park, Zambia	*Papio kindae* Lönnberg, 1919	2	Anna Weyher
18	Lower Zambezi, Zambia	*Papio ursinus griseipes* Pocock, 1911	northern chacma*	Dean Gaffigan
19	Lufupa River, Kafue National Park, Zambia	*Papio kindae* Lönnberg, 1919	2	Jim Auburn
20	Chobe National Park, Botswana	*Papio ursinus chobiensis* Roberts, 1932	northern chacma*	Mike Haworth
21	Catapu, Mozambique	*Papio cynocephalus strepitus* Elliot, 1907	?	Celesta von Chamier

* The northern chacma clade is not regarded as an eastern African baboon clade and, therefore, was not given a number.

Appendix 1b. Map of eastern Africa with localities where the photographs (1–21) presented in appendix 1c were taken. See appendix 1a for locality and taxon names.



Appendix 1c. Photographs of adult male baboons in eastern Africa. See appendix 1a for locality, taxon, and name of the photographer. A map depicting the locations at which these photographs were taken is presented in appendix 1b. More photographs of baboons in the wild from many sites across Africa can be viewed at: http://wildsolutions.nl/photomaps/Papio/

1

2

3a

3b

4

5

6

7

8

9

10

11

12

13

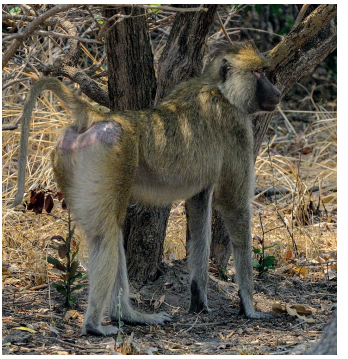
14

15

16

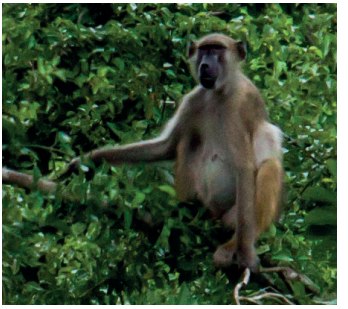
17

18



19

20



21

Appendix 2. List of Papio *spp. samples. The following information is given: taxon (based on morphotype and location), sample ID, country and site of origin (coordinates in decimal degrees), haplotype ID, and GenBank accession number. Abbreviations for taxa are given at the end of this table.*

Taxon	ID	Country	Site Code	Site	Longitude	Latitude	Haplo	Acc. Num.
Tg	149*							EU885487
Pu	463	RSA	Hop	DeHoop Nature Reserve	20.40658	-34.45621		EU885486
Pu	472	NAM	Hak	Hakos Gästefarm	16.36463	-23.23708		EU885480
Pp	252	GUI	Bak	Bakaria, Haute Niger	-10.31542	10.54267		EU885463
Pp	518	SEN	Ked	Kedougou	-12.12472	12.57556		EU885449
Pg	566	ZAM	Nka	Kafue Middle, New Kalala	26.01077	-15.77360		EU885464
Pg	626	ZAM	Bri	Lower Zambezi (Bridge Camp)	30.21516	-15.00530		GQ148683
Pc	961	TZA	Mrg	north of Morogoro	37.58342	-6.62653	1	KU871141
Pc	963	TZA	Mrg	north of Morogoro	37.58342	-6.62653	1	KU871142
Pc	1106	TZA	eMw	east of Mwindu	38.13203	-6.63574	1	KU871165
Pc	1109	TZA	wMi	west of Mwindu	38.06753	-6.63987	1	KU871166
Pc	944	TZA	Mk0	Mikumi Nat. Park	37.20976	-7.23480	1	KU871137
Pc	970	TZA	Dod	30 km west of Dodoma	35.57563	-6.05992	2	KU871143
Pc	978	TZA	sDm	Dodoma->Iringa	35.98829	-7.13016	3	KU871144
Pc	1110	TZA	Wam	Wami River	38.38786	-6.24847	4	KU871167
Pc	1112	TZA	Wam	Wami River	38.38786	-6.24847	4	KU871168
Pc	951	TZA	Mk1	Mikumi Nat. Park	37.18905	-7.26596	5	KU871138
Pc	960	TZA	Mk4	Mikumi Nat. Park	37.16463	-7.34651	6	KU871140
Pc	954	TZA	Mk2	Mikumi Nat. Park	37.06127	-7.35522	6	KU871139
Pc	537	KEN	Amb	Amboseli	37.39000	-2.29000	7	EU885431
Pa	295	TZA	Ngo	Ngorongoro	35.59039	-3.28206	8	KU871123
Pa	1696	TZA	Lmy	Lake Manyara National Park	35.78283	-3.49483	8	KU871218
Pa	1697	TZA	Lmy	Lake Manyara National Park	35.78283	-3.49483	8	KU871219
Pa	1698	TZA	Lmy	Lake Manyara National Park	35.78283	-3.49483	8	KU871220
Pa	351	TZA	Ni2	Serengeti National Park 2	34.79356	-2.42233	9	KU871126
Pa	353	TZA	Sw1	Serengeti National Park 1	34.85236	-2.43100	9	EU885427
Pc	983	TZA	nIg	Dodoma->Iringa	35.73175	-7.33854	9	KU871145
Pc	507	SOM	Web	Webi Shebelli	45.43333	2.42083	10	EU885428
Pa	1699	KEN	Lo3	Lolldaiga Hills Ranch	37.12340	0.26432	10	KU871221
Pc	988	TZA	nIn	Dodoma->Iringa	35.74325	-7.37304	11	KU871146
Pc	990	TZA	nIn	Dodoma->Iringa	35.74325	-7.37304	11	KU871147
Pa	531	KEN	Cho	Cholollo Ranch	37.06000	0.91000	12	KU871131
Pc	1126	TZA	Sgr	Segera	38.65102	-5.38707	13	KU871170
Pa	1609	KEN	Lo1	Lolldaiga Hills Ranch north	37.09806	0.33667	14	KU871183
Pa	1610	KEN	Lo1	Lolldaiga Hills Ranch north	37.09806	0.33667	15	KU871184
Pa	1693	KEN	Lo2	Lolldaiga Hills Ranch cental	37.13736	0.21279	16	KU871217
Pa	340	ETH	Wen	Wendo Genet	38.64965	7.07127	17	KM267363
Pa	341	ETH	Wen	Wendo Genet	38.64965	7.07127	17	KU871124
Pa	343	ETH	Wen	Wendo Genet	38.64965	7.07127	17	KU871125
Pa	1624	ETH	Din	Dinsho, Bale Mts. Nat. Park	39.73464	7.12239	18	KU871189
Pa	1625	ETH	Din	Dinsho, Bale Mts. Nat. Park	39.73464	7.12239	18	KU871190
Pa	1626	ETH	Din	Dinsho, Bale Mts. Nat. Park	39.73464	7.12239	18	KU871191
Pa	1627	ETH	Din	Dinsho, Bale Mts. Nat. Park	39.73464	7.12239	18	KU871192
Pa	1628	ETH	Din	Dinsho, Bale Mts. Nat. Park	39.73464	7.12239	18	KU871193
Pa	1629	ETH	Din	Dinsho, Bale Mts. Nat. Park	39.73464	7.12239	18	KU871194
Pa	1630	ETH	Din	Dinsho, Bale Mts. Nat. Park	39.73464	7.12239	18	KU871195
Pa	1623	ETH	Odo	Odobulu Forest	40.28572	6.86631	18	KU871188
Pa	943	ETH	Har	Harenna 2	39.73718	6.61577	18	KU871136
Ph	1604	ERI	Dad	Dada (Bolo) Assab Area	42.40194	13.15556	19	KM267378
Ph	Abh24	ARA	Abh	Abha	42.50523	18.21639	20	KM267380
Ph	312	ETH	Ger	Gerba Luku/ Erer Gota	41.53400	9.58740	21	KM267351
Pa	509	ETH	CFr	Chilimo Forest	38.16333	9.07167	22	KU871128
Pa	349	ETH	Mn2	Managasha Nat. Park	38.57125	8.96838	22	EU885424
Pa	200	ERI	Hay	Haykota, R. Gash	37.06600	15.15695	23	KM267341
Pa	181	ERI	Tes	Tesseney, R. Gash	36.70142	15.14510	23	KM267339

Taxon	ID	Country	Site Code	Site	Longitude	Latitude	Haplo	Acc. Num.
Ph	240	ERI	Dur	Durfo, SC1	38.96458	15.37370	24	KM267348
Pa	184	ERI	Gri	Griset, R.	36.76018	14.88322	25	EU885422
Ph	301	ETH	ASt	Awash Station	40.17775	8.99268	26	EU885444
Ph	1613	ETH	Mgh	3,3 km NW of Menghi, Tigray	38.86666	13.70865	27	KU871185
Pa	1621	ETH	Jib	Jibat Forest	37.46302	8.77871	28	KU871186
Pa	1622	ETH	Jib	Jibat Forest	37.46302	8.77871	28	KU871187
Pa	338	ETH	Ala	Alambada	38.74768	7.50463	29	KM267362
Pa	332	ETH	Ada	Adami Tulu	38.71493	7.82558	30	KM267358
Pa	334	ETH	Ada	Adami Tulu	38.71493	7.82558	30	KM267359
Pa	335	ETH	Ada	Adami Tulu	38.71493	7.82558	30	KM267360
Pa	336	ETH	Ada	Adami Tulu	38.71493	7.82558	30	KM267361
Pa	1576	TZA	Tuv	Tubila	30.09739	-5.01438	31	KU871181
Pa	1578	TZA	Tuv	Tubila	30.09739	-5.01438	31	KU871182
Pc	1574	TZA	Ngy	Nguye	30.42222	-5.22556	31	KU871179
Pa	444	UGA	Kb1	Kibale Forest 1	30.43333	0.51667	32	KU871127
Pa	448	UGA	Kb2	Kibale Forest 2	30.40000	0.48333	33	EU885420
Pa	512	DCR	sBu	south Bukavu	28.91092	-2.68258	34	EU885421
Pa	513	DCR	sBu	south Bukavu	28.91092	-2.68258	34	KU871129
Pa	516	DCR	sBu	south Bukavu	28.91092	-2.68258	34	KU871130
Pc	529	KEN	Dia	opp. TwoFishes Hotel, Diani Beach	39.55000	-4.32000	35	EU885429
Pc	533	KEN	Dia	opp. TwoFishes Hotel, Diani Beach	39.55000	-4.32000	35	EU885430
Pc	1115	TZA	Ami	Amboni River	39.02847	-5.05471	36	KU871169
Pa	1688	ETH	BO2	Boginda Forest	36.06189	7.50829	37	KU871214
Pa	1692	ETH	BO2	Boginda Forest	36.06189	7.50829	37	KU871216
Pa	942	ETH	Tep	Tepi	35.40529	7.25132	37	KU871135
Pa	1680	ETH	AW2	Bonga, Sheaka Forest, Awurada Valley 2	36.20730	7.08558	38	KU871207
Pa	940	ETH	Kam	Kama1	36.07014	7.31731	39	KU871133
Pa	1675	ETH	AW1	Bonga, Sheaka Forest, Awurada Valley 1	36.22529	7.09305	39	KU871204
Pa	1677	ETH	AW1	Bonga, Sheaka Forest, Awurada Valley 1	36.22529	7.09305	39	KU871205
Pa	1679	ETH	AW1	Bonga, Sheaka Forest, Awurada Valley 1	36.22529	7.09305	39	KU871206
Pa	1689	ETH	GO1	Saja Forest, Gojeb wetland	36.06092	7.55529	40	KU871215
Pa	941	ETH	Kay	Kayakela3	36.22385	7.32184	40	KU871134
Pa	1681	ETH	BA1	Bonga, Bamboo Forest1	36.45549	7.26829	40	KU871208
Pa	1682	ETH	BA2	Bonga, Bamboo Forest2	36.45636	7.24365	40	KU871209
Pa	1683	ETH	BA2	Bonga, Bamboo Forest2	36.45636	7.24365	40	KU871210
Pa	1684	ETH	BA2	Bonga, Bamboo Forest2	36.45636	7.24365	40	KU871211
Pa	1685	ETH	BA2	Bonga, Bamboo Forest2	36.45636	7.24365	40	KU871212
Pa	1686	ETH	BA2	Bonga, Bamboo Forest2	36.45636	7.24365	40	KU871213
Pc	1547	TZA	Mhl	Kasiha, Mahale Mts Nat. Park	29.72477	-6.11688	41	KU871174
Pc	1548	TZA	Mhl	Kasiha, Mahale Mts Nat. Park	29.72477	-6.11688	41	KU871175
Pc	1549	TZA	Mhl	Kasiha, Mahale Mts Nat. Park	29.72477	-6.11688	41	KU871176
Pc	1550	TZA	Mhl	Kasiha, Mahale Mts Nat. Park	29.72477	-6.11688	41	KU871177
Pc	1551	TZA	Mhl	Kasiha, Mahale Mts Nat. Park	29.72477	-6.11688	41	KU871178
Pc	1575	TZA	Ngy	Nguye	30.42222	-5.22556	42	KU871180
Pk	641	ZAM	KN1	North Kafue (Tar road)	25.95237	-14.95338	43	GQ148708
Pk	645	ZAM	Mok	North Kafue (Mokambi Lodge)	25.99480	-14.98034	43	GQ148707
Pk	642	ZAM	KN2	North Kafue (Tar road)	26.06321	-14.95601	44	GQ148709
Pk	565	ZAM	KfN	North Kafue	26.53577	-14.96779	44	EU885438
Pk	649	ZAM	KN3	North Kafue (Tar road)	26.44522	-14.97154	44	GQ148710
Pk	658	ZAM	Kai	between Kaingu & New Kalala	25.99285	-15.30998	44	GQ148711
Pk	581	ZAM	Ka2	Kasanka Nat Park 2	30.24556	-12.57860	45	GQ148705
Pk	582	ZAM	Ka2	Kasanka Nat Park 2	30.24556	-12.57860	45	GQ148706
Pk	568	ZAM	Ka1	Kasanka Nat Park 1	30.25202	-12.59059	45	EU885439
Pk	1532	ZAM	Tgy	Kalambo Falls, south Lake Tanganyika	31.23650	-8.59283	46	KU871171

Taxon	ID	Country	Site Code	Site	Longitude	Latitude	Haplo	Acc. Num.
Pk	1533	ZAM	Tgy	Kalambo Falls, south L. Tanganyika	31.23650	-8.59283	46	KU871172
Pk	1536	ZAM	Tgy	Kalambo Falls, south L. Tanganyika	31.23650	-8.59283	46	KU871173
Pk	569	ZAM	Shi	Shiwa N`gandu	31.73892	-11.19677	47	EU885440
Pk	593	ZAM	Shi	Shiwa N`gandu	31.73892	-11.19677	47	GQ148704
Pc	1031	TZA	sTu	Amani (south of Tunduru)	37.51363	-11.26054	48	KU871154
Pc	1037	TZA	sTu	Amani (south of Tunduru)	37.51363	-11.26054	49	KU871155
Pc	1057	TZA	wMs	west of Masasi	38.60093	-10.86915	50	KU871156
Pc	1096	TZA	Sin	east of Sinza	39.30755	-8.63751	51	KU871164
Pc	1063	TZA	Chw	Chiwata	38.98838	-10.59470	51	KU871157
Pc	1028	TZA	Che	Chem-Chem (south of Tunduru)	37.43273	-11.08541	52	KU871153
Pc	1092	TZA	nNa	15 km north of Nangurukuru	39.31910	-8.95165	53	KU871163
Pc	1071	TZA	Min	Mingoyo	39.65751	-10.10855	53	KU871158
Pc	1074	TZA	Kit	north of Kitaya	40.13528	-10.59522	54	KU871160
Pc	1080	TZA	Dih	Dihimbo	39.97846	-10.41007	55	KU871162
Pc	1077	TZA	Dih	Dihimbo	39.97846	-10.41007	55	KU871161
Pc	1073	TZA	Nan	Nanguruwe	39.99297	-10.49506	55	KU871159
Pc	292	MLW	LCh	Liwonde Nat. Park	35.26033	-15.01481	56	KU871122
Pc	286	MLW	Lb1	Liwonde Nat. Park	35.25392	-15.03905	56	KU871120
Pc	282	MLW	Lb2	Liwonde Nat. Park	35.25343	-15.04134	56	KU871119
Pc	151	MLW	Mc2	Michiru Mountains Cons. Area 2	34.98668	-15.72307	57	EU885433
Pc	994	TZA	Kma	Kamani	34.19086	-8.90926	58	KU871148
Pc	1007	TZA	nSo	north of Songea	35.62780	-9.98347	58	KU871150
Pc	1010	TZA	wSo	west of Songea	35.26643	-10.71346	58	KU871152
Pc	1003	TZA	Son	100 km north of Songea	35.50796	-9.90943	59	KU871149
Pc	1009	TZA	nSo	north of Songea	35.62780	-9.98347	59	KU871151
Pc	288	MLW	Mu2	Mulanje Mt. 2	35.51849	-15.93561	60	EU885434
Pc	287	MLW	Mu1	Mulanje Mt. 1	35.52112	-15.94492	60	KU871121
Pc	404	TZA	Ruk	Tanzania, NO Ufer Lake Rukwa	32.15517	-7.58297	61	EU885432
Pc	600	ZAM	Chf	Chifunde (North Luangwa)	32.43357	-11.86058	62	GQ148698
Pc	409	ZAM	LuS	South Luangwa Nat. Park	31.63793	-13.26840	63	EU885435
Pc	617	ZAM	Mfu	Mfuwe (Flatdogs)	31.77836	-13.10128	64	GQ148703
Pc	615	ZAM	LS2	between Luambe & Mfuwe	32.09203	-12.75000	65	GQ148700
Pc	611	ZAM	LS1	between Luambe & Mfuwe	32.12490	-12.66336	66	GQ148701
Pc	612	ZAM	LS1	between Luambe & Mfuwe	32.12490	-12.66336	66	KU871132
Pc	616	ZAM	LS3	between Luambe & Mfuwe	32.01766	-12.88562	66	GQ148702
Pc	570	ZAM	Lua	Luambe Nat. Park	32.14550	-12.45780	67	EU885437
Pc	604	ZAM	Lua	Luambe Nat. Park	32.14550	-12.45780	67	GQ148699
Pc	155	MLW	LC1	JB, Liwonde Nat. Park	35.30171	-14.85281	68	KU871118
Pc	1639	MLW	Bu2	Bua River, SE Nkhotakota GR	34.15217	-12.83150	69	KU871200
Pc	1640	MLW	Bu2	Bua River, SE Nkhotakota GR	34.15217	-12.83150	69	KU871201
Pc	1649	MLW	Nkh	Western Nkhotakota	34.10867	-13.00967	69	KU871202
Pc	1651	MLW	Nkh	Western Nkhotakota	34.10867	-13.00967	69	KU871203
Pc	411	ZAM	LuS	South Luangwa Nat. Park	31.63793	-13.26840	69	EU885436
Pc	1637	MLW	Sf2	Senga Hills Forest Reserve	34.61853	-13.70677	69	KU871199
Pc	1631	MLW	Sf1	Senga Hills Forest Reserve	34.62217	-13.70967	69	KU871196
Pc	1632	MLW	Sf1	Senga Hills Forest Reserve	34.62217	-13.70967	69	KU871197
Pc	1636	MLW	Sf1	Senga Hills Forest Reserve	34.62217	-13.70967	69	KU871198

7.4.7 Inverted Intergeneric Introgression Between Critically Endangerd Kipunjis and Yellow Baboons in Two Disjunct Populations

BIOLOGY LETTERS

rsbl.royalsocietypublishing.org

Research

Cite this article: Zinner D, Chuma IS, Knauf S, Roos C. 2018 Inverted intergeneric introgression between critically endangered kipunjis and yellow baboons in two disjunct populations. *Biol. Lett.* **14**: 20170729.
http://dx.doi.org/10.1098/rsbl.2017.0729

Received: 27 November 2017
Accepted: 20 December 2017

Subject Areas:
evolution

Keywords:
primates, hybridization, mitochondrial capture, nuclear swamping, divergence-age

Author for correspondence:
Dietmar Zinner
e-mail: dzinner@dpz.eu

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3967377.

THE ROYAL SOCIETY PUBLISHING

Evolutionary biology

Inverted intergeneric introgression between critically endangered kipunjis and yellow baboons in two disjunct populations

Dietmar Zinner[1], Idrissa S. Chuma[2,3,4], Sascha Knauf[2] and Christian Roos[3]

[1]Cognitive Ethology Laboratory, [2]Work Group Neglected Tropical Diseases, and [3]Primate Genetics Laboratory, German Primate Center, Göttingen, Germany
[4]Sokoine University of Agriculture, Morogoro, Tanzania

DZ, 0000-0003-3967-8014; SK, 0000-0001-5744-4946; CR, 0000-0003-0190-4266

Intergeneric hybridization and introgression was reported from one of two populations of the recently discovered kipunji (*Rungwecebus kipunji*), a critically endangered African monkey species of southern Tanzania. Kipunjis of the introgressed population (from Mount Rungwe) carry a mitochondrial DNA (mtDNA) haplotype closely related to those of parapatric yellow baboons (*Papio cynocephalus*), whereas the second kipunji population, in the Udzungwa Mountains, carries the original kipunji mtDNA haplotypes, which diverged from the baboon lineage about 3 million years ago. Interestingly, in our study of yellow baboons in Tanzania, we found that baboons from the southeastern boundary of the Udzungwa Mountains carry mtDNA haplotypes closely related to the original kipunji haplotype, whereas baboons from the northern boundary, as expected, carry mtDNA haplotypes of the northern yellow baboon clade. These findings provide evidence for a case of inverted intergeneric admixture in primates: (i) a baboon mtDNA haplotype introgressed the Mount Rungwe kipunji population by mitochondrial capture and (ii) an Udzungwa Mountains kipunji mtDNA haplotype introgressed a small subpopulation of yellow baboons by either mitochondrial capture or nuclear swamping. The baboon–kipunji example therefore constitutes an interesting system for further studies of the effects of genetic admixture on fitness and speciation.

1. Introduction

Advances in molecular genetics have tremendously influenced our understanding of the impact of hybridization and introgression on the evolution of animal taxa [1,2]. Introgression is a long known phenomenon [3,4] and is defined as the infiltration of genetic material of one species into another through repeated backcrossing of hybrids to one or both parental species [5]. As a result of introgression certain genetic or phenotypical characters of species 'A' can be found in species 'B'. Strong uni-parental backcrossing can lead to mitochondrial or Y-chromosomal capture so that species 'A' might carry mitochondria or the Y-chromosome of species 'B' [1,6]. First indications for introgression are incongruent phylogenies derived from different genetic markers or from genetic and phenotypic characters [1,7].

A remarkable case of intergeneric introgression in primates was reported from the kipunji (*Rungwecebus kipunji*), a critically endangered African monkey, discovered in 2003 in southern Tanzania [8]. Kipunjis occur in two isolated populations: in the forests of Mount Rungwe and the adjacent Livingstone Forest and about 350 km to the northeast in the Ndundulu Forest of the

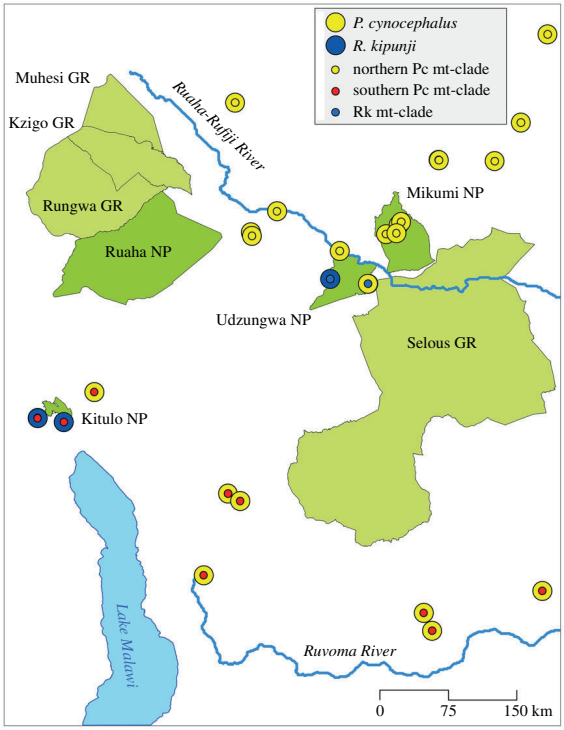


Figure 1. Map of southern Tanzania with yellow baboon (yellow circle) and kipunji (blue circle) sampling locations and geographical occurrence of mtDNA haplotypes of northern (small yellow circles) and southern (small red circles) yellow baboons and kipunjis (small blue circles). Protected areas: GR, game reserve; NP, national park, Pc, *Papio cynocephalus*; Rk, *Rungwecebus kipunji*. Details on samples can be found in the electronic supplementary material, table S1.

Udzungwa Mountains [9]. Analyses of mitochondrial DNA (mtDNA) of the Mount Rungwe population revealed that the kipunji mtDNA sequences are nested within the geographically proximate southern yellow baboon (*Papio cynocephalus*) clade [6,10,11]. In contrast, the population of the Udzungwa Mountains carry mitochondria concordant with a position as sister taxon to baboons [12]. This pattern was interpreted as a result of introgression with mitochondrial capture in the Mount Rungwe population. In contrast the Ndundulu haplotype from the Udzungwa Mountains was considered to represent the true, non-introgressed kipunji haplotype [12]. Here we report on the mtDNA phylogenetic relationships among the Udzungwa kipunjis and sym- and parapatric yellow baboons and a case of inverted intergeneric introgression.

2. Material and methods

During a Tanzania-wide screening for infection of non-human primates with the pathogenic bacterium *Treponema pallidum*, we collected tissue samples from 17 yellow baboons in the Udzungwa Mountain National Park (see electronic supplementary material) (figure 1). After DNA extraction, we amplified and sequenced the Brown region [13] of the mtDNA genome using methods outlined in Newman *et al*. [14] and Zinner *et al*. [15]. Obtained haplotypes were aligned against orthologous sequences derived from GenBank of other baboons, kipunjis, and related taxa that were used as outgroups (electronic supplementary material, table S1). We reconstructed maximum-likelihood (ML) and Bayesian trees and estimated divergence ages. Full details of all analyses and an ethical note are provided in the electronic supplementary material.

3. Results and discussion

As in earlier studies on baboon phylogeny using mtDNA markers [14–19], baboons segregate into various strongly supported clades (ML bootstrap support (BS): greater than 90%, Bayesian posterior probabilities (PP): greater than 0.90). However, the branching pattern among them is not well resolved. Moreover, these clades follow a geographical pattern and their distribution is not always concordant with the boundaries of the six species now generally recognized. As in previous studies, yellow, olive (*Papio anubis*), and chacma baboons (*Papio ursinus*) show para- or polyphylies, which were regarded as results of ancient introgressive hybridization events [15,20].

In our tree reconstruction, as previously found [10–12], the mtDNA lineage of Mount Rungwe kipunjis clusters

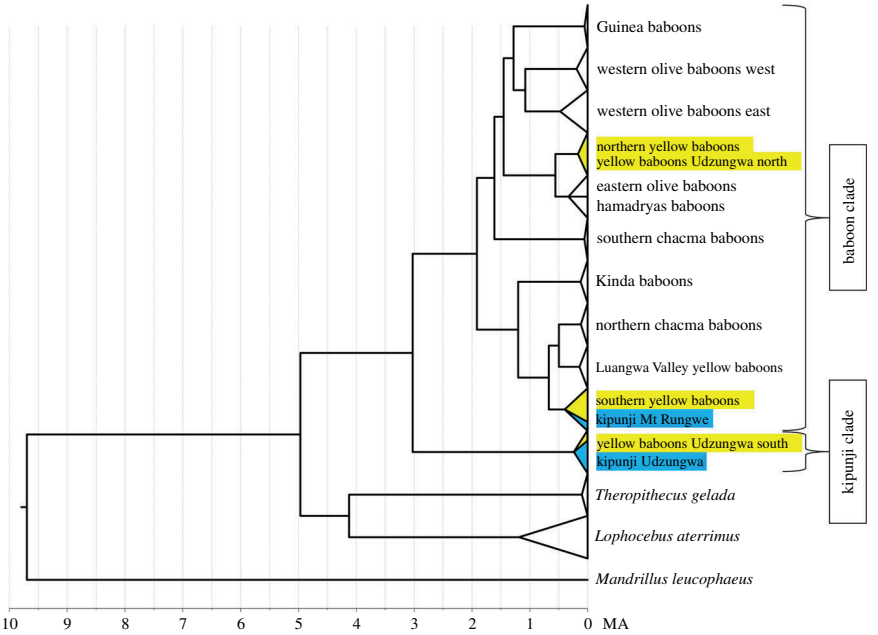


Figure 2. Chronogram showing phylogenetic relationships and divergence times among various baboon and kipunji mtDNA lineages. The mtDNA lineage of the Mount Rungwe kipunji clusters with southern yellow baboons, while yellow baboons from southern Udzungwa cluster with the Udzungwa kipunji. In contrast, yellow baboons from northern Udzungwa cluster with their conspecifics of the northern yellow baboon mtDNA clade (see also electronic supplementary material, figure S1 and table S2). (Online version in colour.)

with southern yellow baboons and the kipunji mtDNA lineage from Udzungwa constitutes a sister clade to the *Papio* clade, with a divergence time of 3.03 million years ago (Ma) (95% highest posterior density (HPD): 2.06–4.09 Ma) (figure 2; electronic supplementary material, table S2). Also expected is the grouping of yellow baboons from northern Udzungwa with conspecifics of the northern yellow baboon clade. However, yellow baboons from southern Udzungwa do not cluster either with southern or with northern yellow baboons. Instead, they represent a closely related sister clade to the mtDNA lineage of Udzungwa kipunjis. We estimated the split of the Mount Rungwe kipunji mtDNA lineage from its nearest southern yellow baboon lineage at 0.05 Ma (0.00–0.13 Ma), while the yellow baboon mtDNA clade from southern Udzungwa diverged from the Udzungwa kipunji lineage at 0.24 Ma (0.07–0.44 Ma) (electronic supplementary material, table S2).

The phylogenetic relationships of the two kipunji populations with respective para- or sympatric yellow baboons suggest three introgression scenarios. Alternative explanations for the observed pattern such as incomplete lineage sorting (ILS) are highly unlikely, because in small populations, as in kipunjis [21], lineage sorting should be relatively fast and the distribution pattern of haplotypes follows a geographical structure, while ILS should be random in respect to geography [15,22]. The first scenario assumes the mtDNA lineage found in Udzungwa kipunjis and baboons represents an ancient *Papio* mtDNA lineage, which would mean that this is the oldest *Papio* mtDNA lineage discovered so far and also that the Udzungwa kipunjis captured baboon mitochondria, similar to the scenario envisioned for the Mount Rungwe kipunji population [11]. The two other scenarios assume that the Udzungwa kipunji mtDNA lineage indeed is the original kipunji mtDNA lineage [12], which would mean that the southern Udzungwa yellow baboons either captured mitochondria from female kipunjis (scenario 2) where theoretically a single event of a female kipunji breeding successfully with a male baboon is sufficient or that baboon males introgressed a subpopulation (e.g. a small isolated group) of the Udzungwa kipunji population over generations leading to nuclear swamping (scenario 3). In scenario 3, male-mediated gene flow from baboons into kipunjis would alter the nuclear gene pool of the kipunjis, up to completely converting the kipunji phenotype into the baboon phenotype, while leaving the ancestral kipunji mtDNA sequences in place. We believe the first scenario is unlikely, because if the Udzungwa kipunji mtDNA lineage represented an old *Papio* lineage, we would expect to find it more frequently in the baboon population of southern Tanzania, which, however, was not the case [19]. It would be interesting to investigate whether this lineage can be also found in baboons of the adjacent Selous Game Reserve. To test the hypothetical scenarios, we would need to analyse whole nuclear genomes inferring amount and frequency of gene flow between kipunjis and baboons. It

would be interesting to see which genes were exchanged, in which direction and whether similar mosaic genome patterns can be detected in both kipunji populations. Also, given that the baboon population carrying the kipunji mtDNA seems to be rather small, it might be interesting to examine whether these baboons show any indications of fitness loss related to genetic admixture, as compared with unadmixed conspecifics.

The kipunji–baboon case is, to our knowledge, the first reported case of inverted intergeneric introgression in primates, and maybe mammals overall. However, inverted introgression among species of the same genus has been reported e.g. in *Lepus*. Brown hares of northern Spain (*L. europaeus*) carry mtDNA lineages of mountain hares (*L. timidus*), whereas mountain hares in northern Russia carry mtDNA lineages of brown hares [23–26]. In the case of Spanish hares, ancient introgression is obvious, because the mountain hare went extinct in the region after the last glacial period, whereas introgressive hybridization is still ongoing in Sweden and Russia. Ongoing hybridization between *Rungwecebus* and *Papio*, however, is not reported.

Ethics. No animals were captured specifically for this study. The study was in line with the Veterinary Act of 2003 and Tanzania Wildlife Research Institute's (TAWIRI) Guideline for Conducting Wildlife Research (2001). Further information is available in the electronic supplementary material.

Data accessibility. Detailed methods are available as electronic supplementary material. Newly generated haplotypes have been deposited in GenBank (accession numbers MG569923–MG569944).

Authors' contributions. D.Z. and C.R. conceived the study. I.S.C. collected samples; S.K. did laboratory work; C.R. did the phylogenetic analyses. All authors discussed the results, collaborated in writing the paper, approved the final version of the manuscript and agreed to be held accountable for the content therein.

Competing interests. The authors declare no competing interests.

Funding. This study was supported by grants of the German Research Foundation (DFG): KN1097/3-1 (to S.K.) and RO3055/2-1 (to C.R.).

Acknowledgements. We are grateful to Julius D. Keyyu (Tanzania Wildlife Research Institute, TAWIRI) and Inyasi Y. Lejora (Tanzania National Parks, TANAPA) for their support and help with the logistics and advice. We thank TAWIRI, TANAPA and the Commission for Science and Technology, Tanzania for necessary permissions. Hassan Mohamed Nguluma, Christina Kibwe and the other Udzungwa National Park staff are thanked for supporting sampling. We thank Simone Lüert and Christiane Schwarz for help with the laboratory work.

References

1. Abbott RJ, Barton NH, Good JM. 2016 Genomics of hybridization and its evolutionary consequences. *Mol. Ecol.* **25**, 2325–2332. (doi:10.1111/mec.13685)
2. Tung J, Barreiro LB. 2017 The contribution of admixture to primate evolution. *Curr. Opin. Genet. Dev.* **47**, 61–68. (doi:10.1016/j.gde.2017.08.010)
3. Anderson E. 1949 *Introgressive hybridization*. New York, NY: Wiley.
4. Heiser CB. 1973 Introgression re-examined. *Bot. Rev.* **39**, 347–366. (doi:10.1007/BF02859160)
5. Anderson E, Hubricht L. 1938 Hybridization in Tradescantia. III. The evidence for introgressive hybridization. *Am. J. Bot.* **25**, 396–402. (doi:10.2307/2436413)
6. Zinner D, Arnold ML, Roos C. 2011 The strange blood: natural hybridization in primates. *Evol. Anthropol.* **20**, 96–103. (doi:10.1002/evan.20301)
7. Funk DJ, Omland KE. 2003 Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.* **34**, 397–423. (doi:10.1146/annurev.ecolsys.34.011802.132421)
8. Davenport TRB, Jones T. 2008 *Rungwecebus kipunji. The IUCN red list of threatened species 2008*, e.T136791A4340286. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T136791A4340286.en (accessed 18 December 2017).
9. Jones T, Ehardt CL, Butynski TM, Davenport TRB, Mpunga NE, Machaga SJ, De Luca DW. 2005 The highland mangabey *Lophocebus kipunji*: a new species of African monkey. *Science* **308**, 1161–1164. (doi:10.1126/science.1109191)
10. Burrell AS, Jolly CJ, Tosi AJ, Disotell TR. 2009 Mitochondrial evidence for the hybrid origin of the kipunji, *Rungwecebus kipunji* (Primates: Papionini). *Mol. Phylogenet. Evol.* **51**, 340–348. (doi:10.1016/j.ympev.2009.02.004)
11. Zinner D, Arnold ML, Roos C. 2009 Is the new primate genus *Rungwecebus* a baboon? *PLoS ONE* **4**, e4859. (doi:10.1371/journal.pone.0004859)
12. Roberts TE, Davenport TRB, Hildebrandt KBP, Jones T, Stanley WT, Sargis EJ, Olson LE. 2010 The biogeography of introgression in the critically endangered African monkey *Rungwecebus kipunji*. *Biol. Lett.* **6**, 233–237. (doi:10.1098/rsbl.2009.0741)
13. Brown WM, Prager EM, Wang A, Wilson AC. 1982 Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J. Mol. Evol.* **18**, 225–239. (doi:10.1007/BF01734101)
14. Newman TK, Jolly CJ, Rogers J. 2004 Mitochondrial phylogeny and systematics of baboons (*Papio*). *Am. J. Phys. Anthropol.* **124**, 17–27. (doi:10.1002/ajpa.10340)
15. Zinner D, Groeneveld LF, Keller C, Roos C. 2009 Mitochondrial phylogeography of baboons (*Papio* spp.)—indication for introgressive hybridization? *BMC Evol. Biol.* **9**, 83. (doi:10.1186/1471-2148-9-83)
16. Wildman DE, Bergman TJ, al-Aghbari A, Sterner KN, Newman TK, Phillips-Conroy JE, Jolly CJ, Disotell TR. 2004 Mitochondrial evidence for the origin of hamadryas baboons. *Mol. Phylogenet. Evol.* **32**, 287–296. (doi:10.1016/j.ympev.2003.12.014)
17. Zinner D, Buba U, Nash S, Roos C. 2011 Pan-African voyagers. The phylogeography of baboons. In *Primates of Gashaka* (eds V Sommer, C Ross), pp. 267–306. New York, NY: Springer.
18. Zinner D, Wertheimer J, Liedigk R, Groeneveld LF, Roos C. 2013 Baboon phylogeny as inferred from complete mitochondrial genomes. *Am. J. Phys. Anthropol.* **150**, 133–140. (doi:10.1002/ajpa.22185)
19. Zinner D, Keller C, Nyahongo JW, Butynski TM, de Jong YA, Pozzi L, Knauf S, Liedigk R, Roos C. 2015 Distribution of mitochondrial clades and morphotypes of baboons *Papio* spp. (Primates: Cercopithecidae) in eastern Africa. *J. East Afr. Nat. Hist.* **104**, 143–168. (doi:10.2982/028.104.0111)
20. Keller C, Roos C, Groeneveld LF, Fischer J, Zinner D. 2010 Introgressive hybridization in southern African baboons shapes patterns of mtDNA variation. *Am. J. Phys. Anthropol.* **142**, 125–136. (doi:10.1002/ajpa.21209)
21. Davenport TRB, De Luca DW, Jones T, Mpunga NE, Machaga SJ, Kitegile A, Phillipps GP. 2008 The critically endangered kipunji *Rungwecebus kipunji* of southern Tanzania: first census and conservation status assessment. *Oryx* **42**, 352–359. (doi:10.1017/S0030605308000422)
22. Avise JC. 2004 *Molecular markers, natural history, and evolution*. Sunderland, MA: Sinauer Associates.
23. Alves PC, Melo-Ferreira J, Freitas H, Boursot P. 2008 The ubiquitous mountain hare mitochondria: multiple introgressive hybridization in hares, genus *Lepus*. *Phil. Trans. R. Soc. B* **363**, 2831–2839. (doi:10.1098/rstb.2008.0053)
24. Marques JP, Farelo L, Vilela J, Vanderpool D, Alves PC, Good JM, Boursot P, Melo-Ferreira J. 2017 Range expansion underlies historical introgressive hybridization in the Iberian hare. *Sci. Rep.* **7**, 40788. (doi:10.1038/srep40788)
25. Melo-Ferreira J, Farelo L, Freitas H, Suchentrunk F, Boursot P, Alves PC. 2014 Home-loving boreal hare mitochondria survived several invasions in Iberia: the relative roles of recurrent hybridisation and allele surfing. *Heredity* **112**, 265–273. (doi:10.1038/hdy.2013.102)
26. Thulin CG, Fang M, Averianov AO. 2006 Introgression from *Lepus europaeus* to *L. timidus* in Russia revealed by mitochondrial single nucleotide polymorphisms and nuclear microsatellites. *Hereditas* **143**, 68–76. (doi:10.1111/j.2006.0018-0661.01952.x)

7.4.8 Serosurvey of *Treponema pallidum* Infection Among Children With Skin Ulcers in the Tarangire-Manyara Ecosystem, Northern Tanzania

Lubinza et al. BMC Infectious Diseases (2020) 20:392
https://doi.org/10.1186/s12879-020-05105-4

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

Serosurvey of *Treponema pallidum* infection among children with skin ulcers in the Tarangire-Manyara ecosystem, northern Tanzania

Clara K. C. Lubinza[1,2], Simone Lueert[3,4], Luisa K. Hallmaier-Wacker[3,4], Esther Ngadaya[1], Idrissa S. Chuma[2], Rudovick R. Kazwala[2], Sayoki G. M. Mfinanga[1,5,6], Klaus Failing[7], Christian Roos[4,8] and Sascha Knauf[3,9*]

Abstract

Background: The first yaws eradication campaign reduced the prevalence of yaws by 95%. In recent years, however, yaws has reemerged and is currently subject to a second, ongoing eradication campaign. Yet, the epidemiological status of Tanzania and 75 other countries with a known history of human yaws is currently unknown. Contrary to the situation in humans in Tanzania, recent infection of nonhuman primates (NHPs) with the yaws bacterium *Treponema pallidum* subsp. *pertenue* (*TPE*) have been reported. In this study, we consider a One Health approach to investigate yaws and describe skin ulcers and corresponding *T. pallidum* serology results among children living in the Tarangire-Manyara ecosystem, an area with increasing wildlife-human interaction in northern Tanzania.

Methods: To investigate human yaws in Tanzania, we conducted a cross-sectional study to screen and interview skin-ulcerated children aged 6 to 15 years, who live in close proximity to two national parks with high numbers of naturally *TPE*-infected monkeys. Serum samples from children with skin ulcers were tested for antibodies against the bacterium using a treponemal (*Treponema pallidum* Particle Agglutination assay) and a non-treponemal (Rapid Plasma Reagin) test.

Results: A total of 186 children aged between 6 and 15 years (boys: 10.7 ± 2.1 (mean ± SD), $N = 132$; girls: 10.9 ± 2.0 (mean ± SD), $N = 54$) were enrolled. Seven children were sampled at health care facilities and 179 at primary schools. 38 children (20.4%) reported active participation in bushmeat hunting and consumption and 26 (13.9%) reported at least one physical contact with a NHP. None of the lesions seen were pathognomonic for yaws. Two children tested positive for treponemal antibodies (1.2%) in the treponemal test, but remained negative in the non-treponemal test.

(Continued on next page)

* Correspondence: sknauf@dpz.eu
[3]Work Group Neglected Tropical Diseases, Infection Biology Unit, Deutsches Primatezentrum GmbH, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Goettingen, Germany
[9]Division Microbiology and Animal Hygiene, Department for Animal Sciences, Georg-August-University, Burkhardtweg 2, 37077 Goettingen, Germany
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusions: We found no serological evidence of yaws among children in the Tarangire-Manyara ecosystem. Nevertheless, the close genetic relationship of human and NHPs infecting *TPE* strains should lead to contact prevention with infected NHPs. Further research investigations are warranted to study the causes and possible prevention measures of spontaneous chronic ulcers among children in rural Tanzania and to certify that the country is free from human yaws.

Keywords: *Treponema pallidum* subsp. *pertenue*, Yaws, Tropics, Africa, Nonhuman primates, Serology, Syphilis, Treponematoses

Background

The bacterium *Treponema pallidum* subsp. *pertenue* (*TPE*) causes a multi-stage disease called yaws, which is commonly found in tropical areas with high rainfall and humidity. Children below the age of 15, living in rural areas with poor hygienic conditions are predominantly affected [1]. In 1948, there were between 50 to 150 million yaws cases globally [1]. African countries such as Ghana, Ivory Coast, Cameroon and the Democratic Republic of the Congo reported more than 100,000 cases annually [1]. The first yaws eradication campaign reduced the prevalence of yaws by 95%. In recent years, however, yaws has reemerged and is currently subject to a second, ongoing eradication campaign. The epidemiological status of Tanzania and 75 other countries with a known history of human yaws is currently unknown [2]. In the 1950s, there were an estimated 61,800 yaws cases in Tanganyika and 5400 cases in Zanzibar [3], the two countries that united to form Tanzania in 1964. Available reports show that the disease was common in Tanzania. High infection rates were for example reported along Lake Tanganyika in the West [3], along Lake Nyasa and Songea in the South [4], and in Tanga, Newala and Lindi in the East [5]. In North Tanzania, few cases of yaws were reported especially among the Maasai and Hadza communities [6]. Records provided by the Tanganyika Director of Medical Services showed that in 1930 and 1950 an estimated 137,000 and 52,400 yaws patients respectively received treatment across the country [3]. Like in many other yaws-endemic countries, there was a dearth of scientific studies to evaluate the epidemiology of yaws. As a result, the majority of reported prevalence data from Tanzania were derived from health facilities and 'colonial traders' records. There is a possibility that yaws statistics are biased due to limitations in documentation in most health facilities, limited access to primary health care for all patients, and false-positive reports due to cross-reacting antibodies to syphilis. The last national yaws report to the WHO in 1978 describes only 77 human yaws cases [7]. Recent studies that evaluate the successful elimination of human yaws in Tanzania do not exist. However, skin ulcers among children are still commonly seen in rural areas of Tanzania [8].

Human yaws initially presents with a skin ulcer that occurs at the invasion site. The skin lesion pathognomonic to the primary stage of yaws (mother yaws) is typically painless. Sometimes lesions are itchy with raised dark margins, an erythematous moist center and they are commonly covered with crust [9]. Other lesions typical in the primary stage are small yellow skin bumps and solitary erythematous papules of two to five centimeters in diameter and squamous macules. The primary skin ulcer is highly infectious and it can take weeks or months to heal spontaneously, leaving behind a hypopigmented or pressed down lesion demarcated by a darker border [9]. After the primary stage lesion has healed, the secondary stage ensues with typical lesions including painful digital swellings and palmar and plantar cracks or discolorations [10]. If untreated these lesions will heal spontaneously after weeks or months and the patient likely enters a latent stage, where evidence of infection can only be obtained through serological tests. Relapses of secondary manifestations have been reported after 5 to 10 years (as reviewed in [10]). If untreated, ~ 10% of patients will proceed to the tertiary stage, which can result in subcutaneous gummatous nodules and chronic periostitis [10]. Transmission occurs through direct contact with infectious lesions.

While data on human yaws are currently not available from Tanzania, infection of nonhuman primates (NHPs) with *TPE* have been reported [11]. The disease is widespread in NHPs and involves at least four species, olive baboons (*Papio anubis*), yellow baboons (*Papio cynocephalus*), vervet monkeys (*Chlorocebus pygerythrus*), and blue monkeys (*Cercopithecus mitis*) [12]. Due to the zoonotic potential of *TPE* strains of NHP origin [13] there is the urgent need to apply a One Health approach to countries that report NHP infection but lack current information on human yaws. The One Health approach dictates that human and animal health are interconnected and should be studied in conjunction [14].

We hypothesized that *TPE* infection is present in children living at the NHP-human interface in rural Tanzania and predicted that a proportion of children with skin ulcers have antibodies against *T. pallidum* (*TP*). We present a description of the skin ulcers in

Tanzanian children and report the corresponding serological results. Our study was designed to investigate the presence, but not to estimate the prevalence of yaws disease in Tanzanian children with a required accuracy.

Methods

Study design, setting and enrollment of participants

We used a cross-sectional study design to screen for *TP* infection among children who live in close proximity to wildlife areas in Tanzania where NHPs are infected with *TPE* [12, 15, 16]. In this first study we selected two areas in the vicinity to Lake Manyara National Park (LMNP) and Tarangire National Park (TNP), respectively (Additional File 1). Both areas are located in a region that reports increasing wildlife-human conflicts [17]. Since *TPE* infection is generally more common in children of primary school age [18], the study involved children between the age of 6 to 15 years from 13 primary schools and three primary health care facilities (Fig. 1). Enrollment took place from November 2017 to February 2018. Further details on the different institutions are provided in Additional File 2. After a short session of health education on yaws, conducted at the primary schools, we asked children to discuss their skin ulcers at home and to self-report their skin ulcers in the presence of their parents or legal guardians on the following day. Only children whose parents or legal guardians came to school and submitted their written parental consent and who in addition to the parental consent expressed their verbal assent to participate, were enrolled and examined by the study team to confirm the self-reported skin ulcers. Children who were brought to health facilities for outpatient services and who had skin ulcers were asked by the attending clinician in the presence of their parent

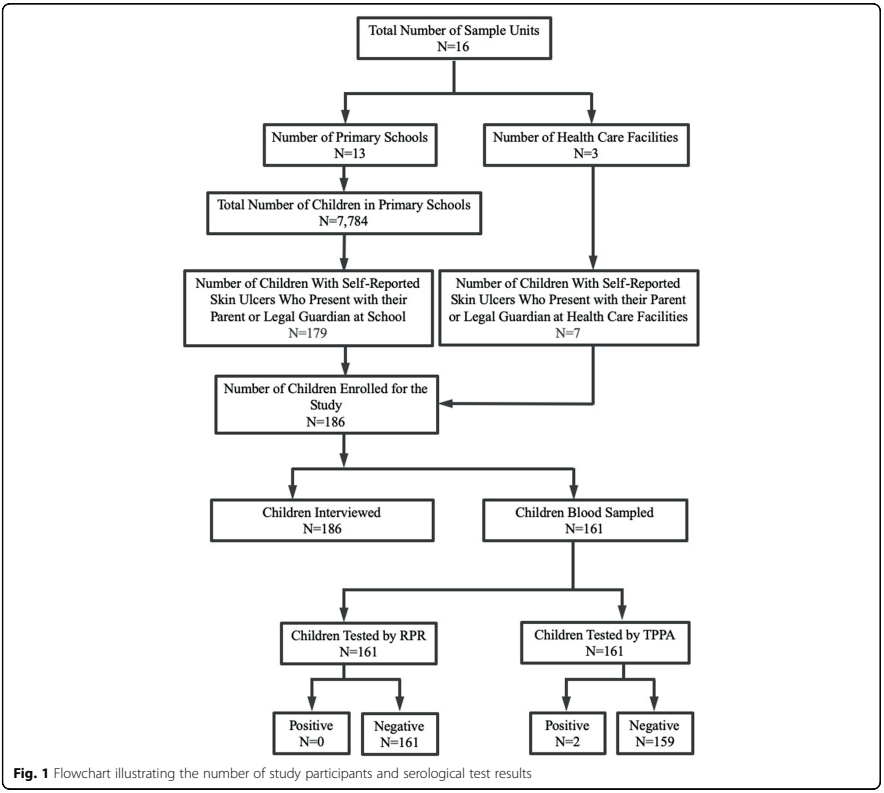


Fig. 1 Flowchart illustrating the number of study participants and serological test results

or legal guardian whether or not they were willing to participate in the study. In case the child and parent/legal guardian agreed, the research team was called to visit the health care facility and the same standards (written consent from parents or a legal guardian and the children's verbal assent) were applied before children were allowed to enroll into the study. Caution was taken to avoid double enrollment by asking the child and parent or legal guardian about previous participation. Skin ulcers of children at schools and health care facilities were cleaned and dressed. In a primary school setting, we advised the parents or guardians to present their child to a clinician at a nearby health facility to receive further treatment for nonhealing wounds. Figure 1 provides a flowchart summarizing the key features of the study.

Sample size

From previous studies conducted in Ghana, one of the current yaws endemic countries, the prevalence of yaws among primary school children (with and without skin ulcers) was estimated at 2.5% [19]. In a first approach, we used this value as a design prevalence to calculate the required sample size that was needed to rule out the existence of yaws infection with an error probability of 5% in the subpopulation of children with skin ulcers in Tanzania. We calculated the sample size using the 1-Stage Freedom analysis tool implemented into Epitool software [20, 21], setting the parameters as follows: population size was set to infinity, an error probability of 0.05 was used and the required test sensitivity and specificity of the Serodia *Treponema Pallidum* Particle Agglutination Assay (Serodia TPPA, FujireBio INC, Japan), which was used to screen for the presence of anti-*T. pallidum* antibodies, was set to 100% each, based on data provided by the manufacturer [22]. Under these conditions, a required sample size of at least 119 children resulted. If the design prevalence was strengthened to 2.0%, the minimum sample size amounted to 149 children. In both cases, design prevalence at 2.5 and 2.0%, the actual sample size in the study exceeded the calculated values.

Study procedures

Children were interviewed by a Tanzanian trained study physician (C.K.C.L.) using a standardized questionnaire that included questions on demographic data (age, sex, education, residence and details on the household), duration and possible causes of their skin ulcer(s), example given an initial injury as well as any history of pain and changes in clinical manifestations. Interviews were conducted in Swahili. A recent study in Central Africa reported frequent physical contact between NHPs and humans [23]. Hence, particular attention was given to information on skin ulcers in other members of the household and the child's lifetime history of physical contact with NHPs and their body fluids or excrements. We furthermore asked the children whether their family keeps livestock.

After the completion of the interview, up to five ml whole blood was collected under aseptic conditions using a closed blood collection system (Vacutainer, Becton Dickinson Medical, USA). Blood samples were subsequently centrifuged at 2000 xg for 15 min at ambient temperature. Aliquots of 1.8 ml serum were collected in cryovial tubes and were kept in liquid nitrogen until they arrived in the laboratory where they were stored at − 80 °C.

Skin ulcers were photographed and documented as shown in the Yaws Recognition Booklet for Communities published by the World Health Organization [24]. Photographs included an additional length standard and sample ID. While the information obtained from the questionnaires was paired with the respective individuals' blood sample, any personal details that could be used to identify the child were anonymized. All sampling procedures were performed by a trained and licensed physician.

Serology

Work steps with potentially infectious material were performed under a BSL-2 laminar flow bench. We used treponemal and nontreponemal tests to test for the presence of anti-*T. pallidum* antibodies.

Treponemal test

We ran all samples in duplicates using the Serodia TPPA following the manufacturer's protocol. The test detects serum antibodies directed against *TP*. We considered the assay valid when the positive control was positive at a dilution of 1:320, the sample diluent reacted negatively with both sensitized and unsensitized particles, and each specimen reacted negatively with unsensitized particles at the final dilution of 1:40. Two researchers interpreted the results independently following the manufacturer's interpretation guidelines. Samples that reacted with both sensitized and unsensitized particles were reanalyzed following an absorption procedure described in the assay's manual.

Non-treponemal test

Serum samples were tested using the Rapid Plasma Reagin test (RPR, Bio–Rad, France), following the manufacturer's instructions. Non-treponemal tests (NTTs) detect antibodies directed against lipoidal antigens from both treponemes and damaged host cells [25]. Test results were interpreted according to the manufacturer's interpretation guidance by two independent researchers. We considered the assay valid, when the positive control (human serum containing antibodies against-*T. pallidum*) showed

agglutination and the negative control (rabbit serum in phosphate buffer) of the test kit showed no agglutination. We categorically excluded samples where the no-template control (NTC) did not remain negative.

Statistical analysis

Statistical analysis was performed utilizing Prism 8 (GraphPad Software Inc). Comparison of categorical variables such as bushmeat hunting and contact to NHPs was done by using 2×2 contingency tables and a two-tailed Fisher's exact test. Proportions were tested at a significance level of 0.05. For prevalence estimation, the upper-limited one-sided 95% confidence intervals (CIs) were calculated using the Wilson/Brown method.

Results

Socio-demographic data

A total of 186 children aged between 6 and 15 years (boys: 10.7 ± 2.1 (mean ± SD), $N = 132$; girls: 10.9 ± 2.0 (mean ± SD), $N = 54$; Fig. 1, Additional File 3) were sampled at their respective schools in the Tarangire-Manyara ecosystem. A mean of 11.6 (±8.6 SD) children was sampled at each of the 16 locations (Additional File 2) where the study was conducted. Children attended different classes at primary school and all lived in villages that were close to wildlife protected areas. The median proximity of the sampling location to the nearest border of a national park that harbors *TPE* infected NHPs was 15.6 km (25th–75th percentile 3.8–31.9, $N = 16$ locations; Distance and Area Measure Software version 1.2.4). Of all 186 children with skin ulcers, 38 children (20.4%) reported active participation in bushmeat hunting and consumption, 26 (13.9%) reported to have had at least one physical contact with a NHP, and 84 children (45.1%) lived in households that keep livestock. Among children with skin ulcers, those who participate in bushmeat hunting and consumption were significantly more often exposed to physical contact with a NHP ($N = 186$, odds ratio 4.4 [two-sided 95% CI 1.9–10.2]; $p = 0.0013$; 2-tailed Fisher's exact test). Thirty children (16.1%) reported that other family members of their household had similar skin lesions.

Skin ulcers

In our subpopulation of skin ulcerated children, we had more boys ($N = 132$, 71%) than girls ($N = 54$, 29%). The ulcers appeared in various shapes, ranging in size between approximately one to five centimeters (Fig. 2, Additional File 4). Pathognomonic yaws lesions were not seen in our group of children. The frequency distribution of the skin ulcer location from this study is illustrated in Additional File 3. Unfortunately, not all data were available from all ulcers. Some children, for example, were unable to answer the question on how long

Fig. 2 Representative skin ulcers in primary school children from Tarangire-Manyara ecosystem. **a**, **b**, **c**, **d**, and **f** are on lower limbs, (**e**) is on the neck. The reported duration of the skin lesions was (**a**): 60 days, (**b**): 7 days, (**c**): 7 days, (**d**): 14 days, (**e**): 14 days, and (**f**): 14 days. A representative figure for one of the bigger range ulcers (> 5 cm) has been included in the Additional Materials (Additional File 4)

they experienced the skin ulcer or technical caveats in combination with time management prevented the study team from documenting ulcer characteristics. In the following, we consistently report data that were available to us. 24.7% ($N = 46/186$) of the children reported that their ulcers are painless. The median duration of ulcers at examination was 14 days (25th–75th percentile 7–30, $N = 183$ ulcers). Twelve (6.4%) children reported the presence of their skin ulcer for 1 year or more. On the cause of the ulcers, 58.1% ($N = 108/186$) of the children reported that their ulcers originated from injuries acquired during sports or daily activities at home or school. However, 41.9% ($N = 78/186$) reported that their skin lesions started as papules of unknown cause and later developed into skin ulcers. Majority of the ulcers were circular in shape (69.9%, $N = 79/113$), had regular margins (85.0%, $N = 96/113$) and sloping edges (97.3%, $N = 110/113$). Most of them (58.4%, $N = 66/113$) were healing, with health granulating tissue on the floor and serous discharge or no discharge. However, 37.2% ($N = 42/113$) were spreading, covered with slough or scabs, with purulent or serosanguinous discharge. The remaining ulcers (4.4%, $N = 5/113$) were callous, with chronic unhealthy granulation tissue.

Serology

The *T. pallidum* specific Serodia TPPA test generated two positives out of a total of 161 tested sera [1.2%; one-sided upper 95% confidence limit = 3.86%]). The samples came from a seven and a thirteen-year-old boy at Mbugwe Primary School in the school district of Babati. No samples, including the two TPPA positives, were positive in the RPR assay. The latter finding, leads to a one-sided upper 95% confidence limit of 1.84% for the prevalence of yaws [one-sided upper 95% confidence limit = 1.84%].

Discussion

In Tanzania, human yaws was last heard of in 1978 where 77 cases were reported to the WHO [7]. Since then, human yaws seemed to have disappeared most likely due to the effect of mass-treatment of yaws patients and their contacts during the first yaws eradication campaign in the 1950s to mid-1960s. However, *TPE* strains have been frequently described in NHPs in Tanzania and are genetically highly similar to human yaws causing strains [11]. Given the close genetic similarity of *TPE* strains of human and NHP origin [11], transmission may occur between NHPs and humans with the highest risk at the NHP-human interface in areas that have a high burden of infected primates. We therefore strategically selected two known hotspots of NHP-*TPE* infection in Tanzania, LMNP and TNP, to screen children for possible yaws infection. Overall, school children are the most affected age group in human yaws [18], which guided our study design. Although the presence of anti-*T. pallidum* antibodies does not allow for the discrimination of any of the three diseases caused by the subspecies *pallidum* (syphilis), *pertenue* (yaws), or *endemicum* (bejel) [10] it is still a very useful tool for mass-screening of yaws as outlined in the Morges Strategy [26]. Although our investigations in Tanzania are ongoing, our findings are supportive of the notion that human yaws is currently not a major health concern in the Tarangire-Manyara ecosystem. Additionally, if NHP to human transmission of yaws occurs, these are most likely very rare events. The WHO is targeting to eradicate yaws globally by the year 2030 [27]. So far only India has been certified free from yaws [28]. In Tanzania, yaws was historically reported from almost the whole country; in the north among Maasai, Hadzabe and nearby communities, in the south in Songea, in the west along Lake Tanganyika, and in the east in Tanga, Newala and Lindi. This study was conducted in the north among Maasai and non-Maasai rural communities. We found skin ulcers to be a common occurrence but there was no serological evidence of yaws. Research is warranted in the remaining areas to identify if any hotspots of infection remain, and if possible certify the county free of yaws.

Only 1.2% of children had antibodies directed against *T. pallidum*. This is similar to findings by Klouman et al., in rural Kilimanjaro, northern Tanzania where the seroprevalence of antibodies against *T. pallidum* in school-aged children was 6.4% among girls and 1.1% among boys [29]. Klouman et al. argue that the most common route of transmission among girls is most likely sexual abuse due to higher prevalence of *TP* antibodies among girls compared to boys, and an additional lack of association between parents' and children's sero status. In our study, children who had *TP* antibodies were both boys, 7 and 13 years old. They did not have non-treponemal antibodies when tested by RPR, suggesting that most likely they have been in contact with syphilis and were treated with antibiotics. Less likely, but still debatable is the possibility that the *TP*-seropositive children had a history of yaws and their skin ulcers are of different aetiology.

While our serological results indicate that children in the Tarangire-Manyara ecosystem are most likely not infected with *TPE*, future research should include broad range tests, such as metagenomics, to diagnose the causes of the frequent skin ulcers in children. These efforts should also include the collection of paired serum and lesions swab samples to increase diagnostic sensitivity and specificity. With an error probability of 0.05, we can conclude that if human yaws is present in children with skin ulcers in the

Tarangire-Manyara ecosystem, its true prevalence was below 3.86% (based on TPPA results) and 1.84% (based on the results of the RPR assay), respectively. Moreover, since the study population is given by children with skin ulcers, which are an important indicator for active yaws infection, the actual prevalence of yaws in the whole population of children, might be even lower in numbers.

The impression that the majority of the skin ulcers are a result of tissue trauma that became infected is also supported by the finding that most ulcers were located at the lower extremities (Additional File 3) and showed the typical morphology of the skin lesions. Most children only had a single chronic skin lesion which was atypical for yaws and not comparable to lesions reported from yaws endemic areas such as Vanuatu [24]. The majority of the ulcers were teeming with pus, painful, foul smelling, swollen or had delayed healing. It became obvious that most of the children were at risk of skin infections such as pyodermas, fungal infections and scabies due to poor hygienic conditions [8]. Skin lesions were neither protected from the environment nor regularly cleaned. In sum, this added to morbidity on children as painful ulcers restricted their participation in daily activities at school and at home. Additionally, unchecked infection could lead to debilitating conditions such as osteomyelitis that carries a risk of permanent disability. Some of the documented ulcers in this study resemble morphological characteristics seen in skin ulcers caused by *Haemophilus ducreyi*. The bacterium is regionally known to mimic yaws-infection in children and has been identified as one of the major differential diagnosis for human yaws infection [30]. Future studies that aim to investigate the causes of skin lesions in children in Tanzania should therefore specifically test for the presence of *H. ducreyi*. Yet, treponemes are a genus of bacteria that infects humans and animals. Many species of the genus are currently described exclusively on the 16S RNA gene level and it must be assumed that a substantial number of *Treponema* species is yet undiscovered. Children in Tanzania are frequently exposed to livestock and their excrements (skin ulcer contamination risk). Since livestock is known to harbor a number of *Treponema* species that colonize the gastro-intestinal-tract or cause hoof diseases [31], we include the information of livestock contact into our survey. We considered the possibility that closely, not yet discovered *Treponema* spp. could challenge the interpretation of test results. This, however, was not an issue in the current study.

There are several limitations associated with the present study, some of which have already been addressed. Based on the variable prevalence rates documented for latent yaws in endemic yaws communities, more studies with an unbiased study design should be encouraged to equally include children with and without skin ulcers as well as their contacts. Sample size is often a limiting factor associated with restricted resources. Moreover, we acknowledge that the Serodia TPPA specificity and sensitivity values that were used to calculate the sample size can be discussed critically since the outcome of interest in this study is based on the treponemal (Serodia TPPA) and non-treponemal test (RPR). Last, future studies would benefit from validation of the standardized questionnaire that is used for interviews.

Conclusions

The presence of chronic spontaneous ulcers among children in Tanzania points towards an unnoticed or a neglected health concern in rural communities, which is sustained by low standards of living. We found no serological evidence of yaws infection in children in Tarangire-Manyara ecosystem. Further studies that elucidate the cause and preventive measures for these ulcers are of public health importance. Furthermore, for a sustainable eradication of human yaws, it will be essential to understand whether NHP infecting *TPE* strains are epidemiologically connected to human infection. Feasible interspecies transmission routes between humans and NHPs exist, e.g., through direct contact during bushmeat butchering [23] or indirect contact through vectors such as necrophagous flies [32]. However, the proof of the existence or nonexistence of transmission events between humans and NHPs requires more data from other ecosystems endemic for NHP *TPE* infection. Our results demonstrate a significant relationship between history of hunting and increased likelihood of coming in physical contact with NHPs ($p = 0.0013$). Although we have found no evidence for *TPE* infection in the children in this study, the close genetic relationship of human and NHP infecting *TPE* strains should lead to contact prevention with infected NHPs e.g., during bushmeat hunting as it puts people at risk to theoretically acquire *TPE* infection. Funders are encouraged to specifically support epidemiological studies that are needed to investigate disease prevalence and which help governments and health care institutions to gain data that allow to certify countries free of yaws. At the same time and based on the recent yaws cases in the Philippines, which are the first reported cases in the Philippines since the 1970s [33], the importance of a robust yaws surveillance cannot be stressed enough.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12879-020-05105-4.

Additional File 1. Map showing Lake Manyara and Tarangire National Park (yellow areas) and the 13 primary schools (red dots) and an additional three primary health care facilities (red cross) where children were samples. The corresponding GPS data can be found in the Additional File 2. The map was constructed using QGIS 3.10.2-A Coruña with open access map source Bing Aerial©Microsoft (http://ecn.t3.tiles.virtualearth.net/tiles/a {q}.jpeg?g = 1).

Additional File 2. Sampling areas and the number of children enrolled from each sampling location.

Additional File 3. (A) Population pyramidic graph of the age distribution in children that were enrolled in this study. The total number of girls (green bars) was 54 compared to 132 boys (blue bars). **(B)** Distribution and frequency of 196 skin ulcer locations in 186 children sampled in this study. Picture source human silhouette (modified, Pixabay [Internet]. Available from: https://pixabay.com/de/illustrations/junge-menschliche-silhouette-kinder-2676579/).

Additional File 4. Representative skin ulcers greater than 3 cm. (**A**) lower limb, 14-year-old girl (**B**) lower limb, 13-year-old boy (**C**) lower limb, 14-year-old girl (**D**) neck, 14-year-old girl. The reported duration of the skin lesions was (**A**): 366 days, (**B**): 14 days, (**C**): 30 days and (**D**): unknown.

Abbreviations

LMNP: Lake Manyara National Park; NHPs: Nonhuman primates; NTT: Non-treponemal test; RPR: Rapid Plasma Reagin; SUA: Sokoine University of Agriculture; TNP: Tarangire National Park; *TP*: *Treponema pallidum*; *TPE*: *Treponema pallidum* subsp. *pertenue*; TPPA: *Treponema pallidum* Particle Agglutination; WHO: World Health Organization

Acknowledgements

We thank the Government of the United Republic of Tanzania (URT) for permission and logistical support that was needed. Institutions include the Ministry for Education and Vocational Training (MoEVT), the Commission for Science and Technology (COSTECH), Tanzania Wildlife Research Institute (TAWIRI), Tanzania National Parks (TANAPA) and Sokoine University of Agriculture (SUA). We are extremely grateful for the support received from the Monduli and Babati districts officials, primary school teachers, health care workers and other staff. Most importantly, our gratitude goes to all people and staff-members working with government and partner institutions in Germany and Tanzania who tirelessly strived to facilitate smooth undertaking of the project at all levels from planning, execution of fieldwork as well as laboratory analyses. Sheila A. Lukehart (University of Washington) and David Šmajs (Masaryk University) are thanked for technical support and scientific advice. Esmeralda M. Francis (SUA) is thanked for her support during field activities.

Authors' contributions

CKCL, SGMM, CR, and SK designed the study. Field work was performed by CKCL, ISC, CR, and SK. Laboratory work took place at Sokoine University of Agriculture and the German Primate Center and was performed by CKCL, SL, LHW, CR, and SK. Data were analyzed by CKCL, SGMM, KF, CR, and SK. All authors (CKCL, SL, LHW, EN, ISC, RRK, SGMM, KF, CR, and SK) contributed to the manuscript preparation. The authors read and approved the final manuscript.

Author's information

CKCL is a medical doctor and a master student at Sokoine University of Agriculture. Her research focuses on epidemiology and characterization of human *TP* infection.

Funding

This study was funded by the German Research Foundation (DFG) grant KN1097/3–1 [SK], KN1097/3–2 [SK], RO3055/2–1 [RO] and RO3055/2–2 [CR]. The funding body had no role in the design of the study, collection, analysis, interpretation or reporting of the findings. We acknowledge support by the Open Access Publication Funds of the Georg-August-University, Goettingen, Germany.

Availability of data and materials

The raw data are available upon reasonable request from the corresponding author.

Ethics approval and consent to participate

Ethical approval was granted by the National Research Ethics Committee at the National Institute for Medical Research (NIMR) in Tanzania (Ref No. NIMR/HQ/R.8a/Vol. IX/1562) and the University of Goettingen in Germany (8/812An). Further permission to conduct the study was obtained from the regional and district authorities in the respective study sites. Parents or legal guardians provided a written informed consent for their children to participate into the study. Assent to participate was also sought from the respective child. Sample collection and documentation was done according to the Helsinki Declaration as last revised by the 64th World Medical Association general assembly in 2013 [34]. The study has been reported according to the STROBE checklist for cross-sectional studies.

Consent for publication

Written informed consent for publication was obtained from parents or legal guardians of the participants.

Competing interests

Authors declare that they have no competing interests.

Author details

[1]National Institute for Medical Research, Muhimbili Medical Research Centre, P.O. Box 3436, Dar es Salaam, Tanzania. [2]College of Veterinary Medicine and Biomedical Sciences, Sokoine University of Agriculture, P.O. Box 3021, Morogoro, Tanzania. [3]Work Group Neglected Tropical Diseases, Infection Biology Unit, Deutsches Primatezentrum GmbH, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Goettingen, Germany. [4]Primate Genetics Laboratory, Deutsches Primatenzentrum GmbH, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Goettingen, Germany. [5]Muhimbili University of Health and Allied Sciences, P.O. Box 65001, Dar es Salaam, Tanzania. [6]School of Life Sciences and Bioengineering, Nelson Mandela African Institute of Science and Technology, Arusha, Tanzania. [7]Unit for Biomathematics and Data Processing, Faculty of Veterinary Medicine, Justus Liebig-University-Giessen, Giessen, Germany. [8]Gene Bank of Primates, Deutsches Primatenzentrum GmbH, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Goettingen, Germany. [9]Division Microbiology and Animal Hygiene, Department for Animal Sciences, Georg-August-University, Burkhardtweg 2, 37077 Goettingen, Germany.

Received: 21 November 2019 Accepted: 18 May 2020
Published online: 03 June 2020

References

1. Mitjà O, Kazadi WM, Asiedu KB, Agana N. Epidemiology of yaws: an update. Clin Epidemiol. 2014;6:119–28.
2. Marks M. Yaws: towards the WHO eradication target. Trans R Soc Trop Med Hyg. 2016;110:319–20.
3. Hackett CJ. Extent and nature of the yaws problem in Africa. Bull World Health Organ. 1953;8(1–3):127–82.
4. Turshen M. The impact of colonialism on health and health services in Tanzania. Int J Health Serv. 1977;7(1):7–35.
5. Howard R. The importance of tertiary yaws. J Trop Med Hyg. 1915;18(3):25–7.
6. Bennett FJ, Barnicot NA, Woodburn JC, Pereira MS, Henderson BE. Studies on viral, bacterial, rickettsial and treponemal diseases in the Hadza of Tanzania and a note on injuries. Hum Biol VO - 45. 1973;45(2):243.
7. World Health Organization. Surveillance on treponematoses. In: Weekly epidemiological records; 1984. p. 377–80.
8. Gibbs SAM. Skin disease and socioeconomic conditions in rural Africa: Tanzania. Int J Dermatol. 1996;35(9):633–9.
9. Mitjà O, Asiedu K, Mabey D. Yaws. Lancet. 2013;381(9868):763–73.
10. Giacani L, Lukehart SA. The endemic treponematoses. Clin Microbiol Rev. 2014;27(1):89–115.
11. Knauf S, Gogarten JF, Schuenemann VJ, De Nys HM, Düx A, Strouhal M, et al. Nonhuman primates across sub-Saharan Africa are infected with the

yaws bacterium *Treponema pallidum* subsp. *pertenue*. Emerg Microbes Infect. 2018;7(1):1–4.
12. Chuma IS, Batamuzi EK, Collins DA, Fyumagwa RD, Hallmaier-Wacker LK, Kazwala RR, et al. Widespread *Treponema pallidum* infection in nonhuman primates, Tanzania. Emerg Infect Dis. 2018;24(6):1002–9.
13. Smith JL, David NJ, Indgin S, Israel CW, Levine BM, Justice J, et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Br J Vener Dis. 1971;47(4):226–51.
14. Rabinowitz PM, Kock R, Kachani M, Kunkel R, Thomas J, Gilbert J, et al. Toward proof of concept of a one health approach to disease prediction and control. Emerg Infect Dis. 2013;19(12):e130265.
15. Harper KN, Fyumagwa RD, Hoare R, Wambura PN, Coppenhaver DH, Sapolsky RM, et al. *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS One. 2012 Dec 20;7(12):e50882.
16. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IAV, Nordhoff M, et al. *Treponema* infection associated with genital ulceration in wild baboons. Vet Pathol. 2012;49(2):292–303.
17. Kaswamila A. Human-wildlife conflicts in Monduli district, Tanzania. Int J Biodivers Sci Manag. 2009;5(4):199–207.
18. Handley BL, Butcher R, Taoaba R, Roberts C, Cama A, Müller A, et al. Absence of serological evidence of exposure to *Treponema pallidum* among children suggests yaws is no longer endemic in Kiribati. Am J Trop Med Hyg. 2019;100(4):940–2 tpmd180799.
19. Mitjà O, Marks M, Konan DJP, Ayelo G, Gonzalez-Beiras C, Boua B, et al. Global epidemiology of yaws: a systematic review. Lancet Glob Heal. 2015;3(6):e324–31.
20. Cameron (1999): Survey toolbox for livestock diseases - a practical manual and software package for active surveillance of livestock diseases in developing countries. Australian Centre for International Agricultural Research, Canberra, Australia. Available from: http://epitools.ausvet.com.au.
21. Cameron AR, Baldock FC. A new probability formula for surveys to substantiate freedom from disease. Prev Vet Med. 1998;34(1):1–17.
22. Knauf S, Dahlmann F, Batamuzi EK, Frischmann S, Liu H. Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Negl Trop Dis. 2015;9(3):e0003637.
23. Narat V, Kampo M, Heyer T, Rupp S, Ambata P, Njouom R, et al. Using physical contact heterogeneity and frequency to characterize dynamics of human exposure to nonhuman primate bodily fluids in Central Africa. PLoS Negl Trop Dis. 2018;12(12):e0006976.
24. World Health Organization. Yaws: recognition booklet for communities. Reprinted with changes, 2014: World Health Organization; 2012. Available from: http://www.who.int/iris/handle/10665/75360.
25. Larsen SA, Steiner BM, Rudolph AH. Laboratory diagnosis and interpretation of tests for syphilis. Clin Microbiol Rev. 1995;8(1):1–21.
26. Organization WH. Eradication of yaws – the Morges strategy. Wkly Epidemiol Rec Relev épidémiologique Hebd. 2012;87(20):189–94.
27. World Health Organization. NTD Roadmap 2021–2030: World Health Organization; 2019. Available from: https://www.who.int/neglected_diseases/news/NTD-Roadmap-targets-2021-2030.pdf?ua=1.
28. Jain SK, Thomas TG, Bora D, Venkatesh S. Eradicating yaws from India: a summary. World Health. 2014;3:4.
29. Klouman E, Masenga EJ, Sam NE. Serological markers for treponemal infection in children in rural Kilimanjaro, Tanzania: evidence of syphilis or non-venereal treponematoses? Sex Transm Infect. 1997;73(6):522–7.
30. Mitjà O, Lukehart S, Pokowas G, Moses P, Kapa A, Godornes C, et al. *Haemophilus ducreyi* as a cause of skin ulcers in children from a yaws-endemic area of Papua New Guinea: a prospective cohort study. Lancet Glob Health. 2014;2:e235–41.
31. Klitgaard K, Strube ML, Isbrand A, Jensen TK, Nielsen MW. Microbiota analysis of environmental slurry and its potential role as a reservoir of bovine digital dermatitis pathogens. Appl Environ Microbiol. 2017;83(11):e00244–17.
32. Knauf S, Raphael J, Mitjà O, Lejora IAV, Chuma IS, Batamuzi EK, et al. Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. EBioMedicine. 2016;11:85–90.
33. Dofitas BL, Kalim SP, Toledo CB, Richardus JH. Yaws in the Philippines: first reported cases since the 1970s. Inf Dis Pov. 2020;9(1):1–16.
34. World Medical Association (WMA). WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. WMA Gen Assem. 2013. Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.

Publisher's Note

8. General Discussion on the Yaws Case in Nonhuman Primates

In chapters 2 to 6, I outlined the basic principles of wildlife disease ecology, the concept of One Health and reviewed the principles that functional disease reservoirs must fulfil. This was followed by an overview of molecular methods that are used to investigate diseases as an integral part of natural ecosystems. This, in combination with the important aspects of disease control, elimination and eradication, has been applied to the investigation of *Treponema* infection in NHPs in chapter 7.

The endemic treponematoses, which comprise yaws (*T. p*. subsp. *pertenue*), bejel (*T. p.* subsp. *endemicum*) and pinta (*Treponema carateum*) have been subject to eradication efforts since the early 1950s (Asiedu et al., 2014). Within 12 years of intense treatment efforts, the global prevalence of human treponematoses dropped by 95%. Until today, the success of this first eradication campaign marks one of the greatest public health achievements in the history of the World Health Organization (Asiedu et al., 2014). Unfortunately, yaws was not eradicated and the remaining yaws cases have let to re-emerging yaws in West Africa, Southern Asia and the Pacific region (Asiedu et al., 2014). Today, human yaws is endemic in 14 countries and an additional 84 countries have a history of human yaws but lack recent epidemiological data (Marks, 2016; World Health Organization, 2018). The currently ongoing second eradication campaign has been conceptualised on the basis of two assumptions: (A) yaws is officially believed to be an exclusively human pathogen (Enserink, 2018; Laursen, 2018) and (B) effective mass-treatment can be achieved in reasonable time through a single orally administered dose of azithromycin (World Health Organization, 2012). Eradication of the yaws disease has originally been anticipated until 2020 (World Health Organization, 2012) and is now projected until 2030 (http://www.who.int/neglected_diseases/news/NTD-Roadmap-targets-2021-2030.pdf?ua=1). On the short-term, the success of human yaws eradication is tied on multiple variables (Marks et al., 2015), such as the rapid growth of macrolide-resistance after a single treatment round (Mitjà et al., 2018) and the ability to detect the last yaws cases in the end phase of eradication (Dyson et al., 2018). In the long-term, sustainability, which equals to an infinite global zero case scenario (chapter 6), will rely on the existence of a functional nonhuman disease reservoir (Hallmaier-Wacker et al., 2017; Knauf et al., 2018; Marks et al., 2015).

Treponema pallidum is maintained in NHPs in Africa

Different African NHP taxa have been tested positive for *T. pallidum* in the wild (Chuma et al., 2018; Harper et al., 2013). The six decades-long demonstration of wild NHPs infected with the bacterium, argues for a well-established host-pathogen relationship and ongoing transmission cycles, in particular, because *T. pallidum*, unlike the sit-and-wait pathogens, is not capable of surviving outside of its host (Fraser et al., 1998). Although there are no experiments conducted that provide evidence for the lack of metabolic functionality in *TPE* strains of NHP origin, the high genetic similarity between human and NHP infecting *T. pallidum* strains (Knauf et al., 2018) suggests an equally restricted metabolic capacity. As a result, *T. pallidum* has a parasitic relationship to its NHP host as its viability largely depends on several host nutritional substrates (Fraser et al., 1998).

There is a clear spatial overlay of areas that have a past and current history of human yaws and areas where infected wild NHPs have been reported (spatial connectivity, chapter 4) (Knauf et al., 2013). This overlay, on the one hand, argues that human and NHP infecting strains of *T. pallidum* share the same ecological requirements in terms of climatic conditions, as well as indicates the potential for interspecies transmission. Similar to human yaws infection, there is a paucity of information on the current extent of *TPE* infection in NHPs. As outlined in chapter 5.3, there are validated commercially available serological tests (chapter 5.1) for the use in baboons on the market (Knauf et al., 2015b), which makes serological investigations a valuable tool to investigate the current epidemiology of the yaws bacterium in NHPs on a temporal and spatial level (chapter 5.3). Even in the absence of clinical lesions many NHPs in Africa test positive for anti-*T. pallidum* antibodies (Chuma et al., 2018; Harper et al., 2012; Knauf et al., 2011; Knauf et al., 2015a; Knauf et al., 2018; Zimmerman et al., 2019). Since wild NHPs do not get treatment against yaws infection and host clearance of infection is unlikely, even in the NHP host. Seropositive animals should, therefore, be considered to be in the latency stage of infection (Chuma et al., 2018; Zimmerman et al., 2019).

There is clear evidence that all NHP infecting *T. pallidum* strains belong to the subsp. *pertenue* (Knauf et al., 2018). However, despite the major improvements in molecular pathogen detection, we are still challenged to build a convincing body of evidence for a disease reservoir in NHPs, when the pathogen (here treponemes) is only identified on the molecular level without isolation or purification to demonstrate their functional involvement in human infection (Hallmaier-Wacker et al., 2017; Smolinski et al., 2003). The high genetic similarity (Knauf et al., 2018; Strouhal et al., 2018) in combination with the historical report of human infection with the monkey isolate strain Fribourg-Blanc (Smith et al., 1971), however, underline the functional similarity of *TPE* strains of human and NHP origin (genetic and functional similarity, chapter 4). Thus, at least the baboon infecting Lake Manyara National Park (LMNP)-1 and the Fribourg-Blanc strain should be considered potentially zoonotic. New possibilities to culture *T. pallidum* (Edmondson et al., 2018) promise new avenues to further investigate the functional and immunological differences between *TPE* strains that infect humans and NHPs.

Asian macaques (*Macaca* sp.) have been used as animal models for human treponematoses including syphilis (Castellani, 1907; Marra et al., 1998; Nichols, 1910; Sepetjian et al., 1969; Turner and Hollander, 1957), which renders macaques susceptible for the yaws bacterium. Yet, despite intense efforts, only a few pet monkeys in a defined area on Sulawesi have been tested positive for antibodies against *T. pallidum* (Klegarth et al., 2017) indicating that yaws infection is not a problem in wild NHPs in Asia. The reason why macaques in areas with a high human yaws burden are not infected, whereas the infection in African NHPs is widespread (Chuma et al., 2018) remains puzzling and contributes to the hypothesis that interspecies transmission events are rare (Chuma et al., 2019) and that human and NHP infecting *T. pallidum* strains are therefore epidemiologically disconnected (Lubinza et al., 2020). From the evolutionary perspective, macaques could have left the African continent before yaws evolved from an ancestor that has rapidly spread across the different primate taxa (Knauf et al., 2018). However, to answer this question more genome data of *T. pallidum* strains are required to generate a molecular timescale for the evolution of *T. pallidum* in primates.

Feasible transmission routes exist

As outlined in chapter 4, a disease reservoir can only function when feasible transmission routes exist (Hallmaier-Wacker et al., 2017). A newly designed MLST system for *TPE* strains in four different Tanzanian NHP species proofed to be a useful tool to gain more insight into how *T. pallidum* is maintained in the different NHP populations (Chuma et al., 2019). There is a remarkable temporal and geographic stability of the strains that infect monkeys in Tanzania (Chuma et al., 2019), similar to what has been described in human infecting strains (Grillová et al., 2018).

The biology of treponemes dictates direct skin-to-skin contact as the main source of infection (Giacani and Lukehart, 2014). The clinical manifestations seen in East African NHPs are mostly found at the genital (Chuma et al., 2018), which is highly suggestive for a sexual transmission mode. Behavioural studies like the one conducted at LMNP (chapter 7.4.4) therefore provide valuable insight into disease transmission dynamics within NHP societies (Paciência et al., 2019). It is necessary to not only focus on the pathogen when studying wildlife diseases but also on host factors that influence the transmission of a pathogen (chapter 2.4). *TPE* infection in baboons seems to impose a selection pressure that induces individual behavioural modifications that alter mate choice and reduce promiscuity (Paciência et al., 2019). These data are of great value for the epidemiological modelling of *T. pallidum* infection and will help to make existing disease models (Nunn et al., 2014) more applicable and predictive.

Host genetics particularly helpful when a sexual transmission pathway is expected. Through comparison of mitochondrial clades among the different host species, sexual interactions can be traced (Kopp et al., 2006; Roos et al., 2011; Zinner et al., 2015, 2012, 2009b, 2009a). The finding of inverted intergeneric introgression between two NHP taxa in Tanzania shows that sexual transmission can also occur across species and even genera (Zinner et al., 2018) and is not restricted to hybridization zones of closely related subspecies. As described in chapter 5.4 this kind of investigation generated important baseline data (in this case mitochondrial genomes (Roos et al., 2018; Zinner et al., 2018, 2015)) that can help to investigate feasible transmission pathways. Wildlife disease investigations should, therefore, incorporate pathogen- and host-related investigations to get a sophisticated understanding of the biology of the pathogen within a given ecosystem.

In addition to the transmission via direct skin contact, early experiments with viable treponemes have shown that flies may play a role in the transmission of the yaws bacterium (Castellani, 1907; Kumm, 1935a; Kumm, 1935b; Kumm and Turner, 1936; Lamborn, 1936; Satchell and Harrison, 1953). The studies conducted in Africa (Knauf et al., 2016) and Papua New Guinea (Houinei et al., 2017) as well as in Cote d'Ivoire (Gogarten et al., 2019) have repeatedly shown that necrophagous flies in areas of NHP and human infection carry the DNA of the spirochete and therefore further support the early findings. However, experiments under controlled laboratory conditions are warranted to investigate the functional aspect of vector transmission (chapter 4), which is currently lacking in all of the recent studies. Although fly transmission is not expected to be a major transmission pathway for *TPE* (Knauf et al., 2016), the existing data support the potential for interspecies transmission.

Despite all the work done to investigate *T. pallidum* infection in NHPs, the One Health concept (chapter 3) requires further investigations into human health, in the same areas where

NHPs have been reported to be infected. Studies such as the one published my Lubinza et al. (Lubinza et al., 2020) provide important data on disease transmission and will in the best case, help health care organizations to certify countries with known yaws history free of yaws.

In conclusion, there is compelling evidence that the yaws bacterium persists in the NHP host for quite a long time. Although there are some indications that interspecies transmission is possible, further research is warranted to investigate NHPs as a source for human yaws infection. Current research data suggest that NHP to human infection, if existent, must be extremely rare. However, sustainable yaws eradication requires an infinite global zero-case scenario. While it is clear, based on the presented data in chapter 7, that *TPE* cannot be eradicated, the second and ongoing eradication campaign might be successful to eradicate yaws infection in humans. However, long-term post-eradication surveillance programs must be installed to detect possible, though extremely rare transmission events between NHPs and humans.

9. Summary

The large number of human diseases with zoonotic origin as well as our increasing understanding of the ecological context in which the disease-causing pathogens subsist, has led to the conceptualization of the One Health model, where the health of humans is inevitably linked to the health of animals and the environment. Although wildlife is generally not seen as a source of pathogen threat, it plays a critical role as an indirect source for candidate pathogens that can have a major impact on human and livestock health. Microorganisms, including those capable of causing host damage (pathogens), are an integral part of natural ecosystems. In this thesis, I use *Treponema pallidum* subsp. *pertenue* (*TPE*), the yaws-bacterium that infects humans and nonhuman primates (NHPs; 7.2.2 Knauf et al. 2013. Emerging Infectious Diseases), as an example to investigate a disease reservoir in the context of natural ecosystems, one health and wildlife health. The identification of NHPs as a disease reservoir for human infection is of fundamental importance to the success of the currently ongoing yaws eradication (7.2.1 Marks et al. 2015. Lancet Infectious Diseases).

To diagnose and characterize *TPE* in NHPs, molecular tests were established (5.2.2.1 Hallmaier-Wacker et al. 2019. Frontiers in Microbiology), validated and standardized (5.3.1 Knauf et al. 2015. PLoS Neglected Tropical Diseases; 5.3.2 Knauf et al. 2018. PLoS Neglected Tropical Diseases; 5.3.3 Becherer et al. 2019 Emerging Infectious Diseases; 5.3.4 Hallmaier-Wacker et al. 2018. Scientific Reports). Subsequently, these tools were applied to investigate *TPE* in the context of natural ecosystems. Yet, the two essential requirements for a functional disease reservoir are pathogen maintenance, as demonstrated by genetic and functional similarity, and the existence of a feasible transmission route (4.1 Hallmaier-Wacker et al. 2017. Emerging Microbes & Infections). Neither laboratory experiments nor isolated cases alone are informative to describe the functionality of NHPs as a disease reservoir for human yaws infection.

Several publications that are part of this thesis indicate that *T. pallidum* is geographically widespread and endemic in African NHPs (7.3.1 Knauf et al. 2015. PLoS One; 7.3.2 Zimmerman et al. 2019. Emerging Infectious Diseases; 7.3.3 Chuma et al. 2018. Emerging Infectious Diseases). Understanding evolutionary processes and mechanisms of

virulence are vital to the identification of pathogen candidates that have the highest potential to impact population health. In the study presented in chapter 7.3.4 (Knauf et al., 2018. Emerging Microbes & Infections), me and my team have been able to demonstrate that NHPs across sub-Saharan Africa are infected with the same bacterium that causes human yaws. Based on whole genome sequences, syphilis (subsp. *pallidum*) and yaws (subsp. *pertenue*) causing strains of *T. pallidum* of human and NHP origin form reciprocally monophyletic groups and NHP-infecting strains all cluster with human-infecting *TPE* strains. The genome of the baboon strain LMNP-1 from Tanzania has the same genome structure as human infecting *TPE* strains and the suspected human-pathogenic strain Fribourg-Blanc, which was isolated from a baboon in West Africa (7.3.4 (Knauf et al. 2018. Emerging Microbes & Infections; 7.3.5 Strouhal et al. 2018. PLoS Neglected Tropical Diseases). Until this study, *TPE* was believed to be an exclusive human pathogen. There is, however, evidence for zoonotic potential in all of the documented *TPE* strains of NHP origin. Interestingly, our survey conducted on treponemal infections in free-ranging and captive macaques in Asia (7.3.6 Klegarth et al. 2017. Emerging Infectious Diseases) further supports a One Health approach to eradicate human yaws, as pet monkeys might function as an additional disease reservoir. However, Asian wild macaques seem to be free of the disease, even in areas that have a high-human yaws prevalence. While this finding warrants further investigations, it also supports the perception that even in Africa, the interspecies transmission of *TPE* is a rare event (7.4.1 Chuma et al. 2019. Scientific Reports). Several feasible transmission routes have been discussed, including necrophagous flies as a vector (7.4.2 Knauf et al. 2016. EBioMedicine; 7.4.3 Houinei et al. 2017. PLoS Neglected Tropical Diseases). Sexual transmission, however, seems to be the most likely route of infection in East African NHPs (7.4.4 Paciência et al. 2019. Science Advances; 7.4.7 Zinner et al. 2018. Biology Letters). A striking finding of our work is the discovery that *TPE* imposed a selection pressure in NHPs that results in individual behavioural modifications and thus altered mate choice (7.4.4 Paciência et al. 2019. Science Advances). These kinds of evolutionary processes have rarely been witnesses in long-lived vertebrates and help us to understand how pathogens shape mating systems.

Based on my research, there is compelling evidence that the yaws bacterium persists in the NHP host for quite a long time. Our data suggest that NHP-to-human infection, if existent, must be extremely rare and that long-term post-eradication surveillance programs are needed to detect possible transmission events between NHPs and humans (7.4.8 Lubinza et al. 2020. BMC Infectious Diseases). The *Treponema* case shows, that an ecological and evolutionary perspective on pathogens is a requirement to adequately manage diseases and to conduct a risk assessment of future transmission events. This, however, requires baseline data and is not restricted to the simple presence or absence of microbes that have the potential to cause diseases. Baseline data must include temporal changes in prevalence and virulence, which can be age or strain depended or based on changes in the host community composition or size. The decision to control, eliminate or eradicate a disease depends on the technical and biological feasibility including the existence of a nonhuman reservoir as well as the benefit in excess of the costs.

9.1 Zusammenfassung

Die hohe Anzahl an humanen Krankheitserregern mit zoonotischem Ursprung, als auch das zunehmende Verständnis über die Lebensräume, in denen Krankheitserreger vorkommen, hat zu dem One Health Konzept geführt, bei dem die Gesundheit des Menschen untrennbar mit der Gesundheit von Tier und Umwelt verbunden ist. Wildtiere werden allgemein nicht als direkte Ansteckungsquelle für Krankheiten gesehen. Indirekt spielen sie jedoch eine Rolle als Ursprung von Erregern, die einen erheblichen Einfluss auf die Gesundheit von Menschen und Nutztieren haben. Mikroorganismen, inklusive solcher, die im Stande sind, ihren Wirt zu schädigen (Pathogene), sind ein integraler Bestandteil von natürlichen Ökosystemen. In dieser Habilitationsschrift nutze ich *Treponema pallidum* subsp. *pertenue (TPE*), das Frambösie-Bakterium welches Menschen und nichtmenschliche Primaten (NMPen) infiziert (7.2.2 Knauf et al. 2013. Emerging Infectious Diseases) als ein Beispiel, wie man ein Krankheitsreservoir im Kontext von natürlichen Ökosystemen, One Health und der Wildtiergesundheit untersucht. Der Nachweis eines Krankheitsreservoirs für humane Infektionen in NMPen ist für den Erfolg der aktuellen Frambösie-Ausrottungsbemühungen fundamental wichtig (7.2.1 Marks et al. 2015. Lancet Infectious Diseases).

Um *TPE* in NMPen zu diagnostizieren und zu charakterisieren (5.2.2.1 Hallmaier-Wacker et al. 2019. Frontiers in Microbiology), wurden molekulare Tests etabliert, validiert und standardisiert (5.3.1 Knauf et al. 2015. PLoS Neglected Tropical Diseases; 5.3.2 Knauf et al. 2018. PLoS Neglected Tropical Diseases; 5.3.3 Becherer et al. 2019 Emerging Infectious Diseases; 5.3.4 Hallmaier-Wacker et al. 2018. Scientific Reports). Im Folgenden wurden diese Methoden angewendet, um *TPE* im Zusammenhang mit natürlichen Ökosystem zu erforschen. Zwei wesentliche Voraussetzungen für die Anerkennung eines funktionalen Erregerreservoirs ist die Erhaltung des Pathogens, die durch eine hohe genetische und funktionale Ähnlichkeit nachgewiesen wird, als auch die Existenz plausibler Übertragungswege (4.1 Hallmaier-Wacker et al. 2017. Emerging Microbes & Infections). Weder Laborexperimente, noch isolierte Fallberichte allein sind informativ genug, um die Funktionalität von NMPen als Erregerreservoirs für die humane Frambösie zu beschreiben.

Zahlreiche Veröffentlichungen, die Bestandteil dieser Habilitationsschrift sind, zeigen, dass *T. pallidum* geographisch weitverbreitet ist und endemisch in afrikanischen NMPen vorkommt (7.3.1 Knauf et al. 2015. PLoS One; 7.3.2 Zimmerman et al. 2019. Emerging Infectious Diseases; 7.3.3 Chuma et al. 2018. Emerging Infectious Diseases). Das Verständnis von evolutionären Prozessen und Mechanismen der Virulenz ist vital für das Ansprechen von Krankheitserregerkandidaten, die das höchste Potential haben, die Population zu schädigen. In der Studie, die im Kapitel 7.3.4 (Knauf et al., 2018. Emerging Microbes & Infections) präsentiert wird, zeige ich mit meinem Team, das NMPen in Sub-Sahara Afrika mit dem gleichen Bakterium infiziert sind, das auch die humane Frambösie verursacht. Basierend auf vollständigen Genomen, bilden *T. pallidum* Stämme, die Syphilis (subsp. *pallidum*) und Frambösie (subsp. *pertenue*) verursachen, reziprokale monophyletische Gruppen und alle NMPen infizierende Stämme clustern mit *TPE* Stämmen die Human-Infektionen verursachen. Das Genom des Pavianstammes LMNP-1 aus Tansania hat die gleiche Genomstruktur wie humane *TPE* Stämme und der mutmaßlich humanpathogene Stamm Fribourg-Blanc, der von einem Pavian in West Afrika isoliert wurde (7.3.4 Knauf et al. 2018. Emerging Microbes & Infections; 7.3.5 Strouhal et al. 2018. PLoS Neglected Tropical Diseases). Bis zu dieser Studie

wurde *TPE* nur als ein Pathogen des Menschen angesehen. Es gibt jedoch Hinweise auf ein zoonotisches Potential in allen dokumentierten *TPE* Stämmen, die NMPen befallen. Interessanterweise unterstützt unsere Studie über Treponematosen in freilebenden und in menschlicher Obhut gehaltener Makaken in Asien (7.3.6 Klegarth et al. 2017. Emerging Infectious Diseases) ebenfalls einen One Health Ansatz, um die Frambösie des Menschen auszurotten, da Affen, die als Haustiere gehalten werden, als ein zusätzliches Krankheitsreservoir fungieren könnten. Freilebende asiatische Makaken scheinen jedoch frei von der Erkrankung zu sein, sogar in Gebieten mit einer hohen Prävalenz an humaner Frambösie. Dieser Ergebnisse rechtfertigen einerseits weitere Untersuchungen, sie stützen aber auch die Wahrnehmung, dass auch in Afrika die artübergreifende Transmission von *TPE* ein seltenes Ereignis ist (7.4.1 Chuma et al. 2019. Scientific Reports). Zahlreiche mögliche Übertragungswege wurden diskutiert, einschließlich der Übertragung durch nekrophage Fliegen als Vektor (7.4.2 Knauf et al. 2016. EBioMedicine; 7.4.3 Houinei et al. 2017. PLoS Neglected Tropical Diseases). Die sexuelle Übertragung ist jedoch der wahrscheinlichste Infektionsweg in ostafrikanischen NMPen (7.4.4 Paciência et al. 2019. Science Advances; 7.4.7 Zinner et al. 2018. Biology Letters). Ein herausragendes Resultat unserer Forschung war die Tatsache, dass *TPE* einen Selektionsdruck auf NMPen ausübt, der individuelle Verhaltensmodifikationen verursacht, die in der Folge das Paarungsverhalten verändern (7.4.4 Paciência et al. 2019. Science Advances). Solche evolutionären Prozesse wurden selten in langlebigen Vertebraten beobachtet und helfen uns zu verstehen, wie Pathogene das Paarungsverhalten bestimmen.

Meine Forschung zeigt, dass es überzeugende Belege dafür gibt, dass das Frambösiebakterium schon sehr lang in den NMPen persistiert. Unsere Daten zeigen, dass eine NMPen-zu-Mensch Infektion, sollte diese stattfinden, ein sehr seltenes Ereignis sein muss und dass Langzeit-post-Eradikationsstudien zur Überwachung der Frambösiefreiheit benötigt werden, um ein mögliches Übertragungsereignis zwischen NMP und Mensch zu erfassen (7.4.8 Lubinza et al. 2020. BMC Infectious Diseases).

Das *Treponema*-Beispiel zeigt, dass eine ökologisch-evolutionäre Sichtweise auf Krankheitserreger notwendig ist, um Krankheiten adäquat zu managen und das Risiko von zukünftigen Übertragungen einschätzen zu können. Dies wiederum bedarf Basisdaten, die nicht darauf beschränkt sind, nur das Vorhandensein- oder Nichtvorhandensein von Mikroorganismen mit dem Potential, Krankheiten zu verursachen, zu dokumentieren. Basisdaten müssen zeitliche Veränderungen in der Prävalenz und Virulenz beinhalten, die durch das Alter des Wirts, den Stamm oder Veränderungen in der Zusammensetzung der Wirtsgemeinschaft beeinflusst sein können. Die Entscheidung darüber, Erkrankungen zu kontrollieren, zu eliminieren oder auszurotten, hängt von den technischen und biologischen Möglichkeiten ab, sowie von der Existenz eines nicht menschlichen Erregerreservoirs und dem Kosten-Nutzen-Faktor.

10. Acknowledgments

My work would have not been possible without the tremendous support of my wife **Yvonne** and my daughter **Ella Siv Aina**. Thanks for your endless patience and motivation throughout the years. I hope, 2021 was the last year, where I had to say “No worries, next year everything will get easier”. Nothing is as valuable as you, Richard and Darwin.

A. Wehrend's (Justus-Liebig-University, Germany) altruistic support, guidance and motivation had a major impact on my career development. Thanks for helping me to stay on track over the last years. I am also thankful for the support from **M. Lierz** and **M. Kramer** (Justus-Liebig-University, Germany), who enabled me to work on my habilitation at the Faculty of Veterinary Medicine at Justus-Liebig-University of Giessen.

U. Siebert (Hanover Veterinary School, Germany) is thanked for her support and research collaboration as well as for providing me the opportunity to teach veterinary students at the Hanover Veterinary School, Germany.

H. Bostedt (Justus-Liebig-University, Germany) trained and motivated me as a student and was also a constant source of collegial and professional advice throughout the years since I left the University of Giessen.

I owe **Simone Lüert** and **Luisa K. Hallmaier-Wacker** (Deutsches Primatenzentrum GmbH, Germany) a special depth of gratitude. I enjoyed every moment of working together as a team, developing ideas and discovering the unknowns of the spirochete world.

My students **I. S. Chuma** (Tanzania National Parks, Tanzania), **C. K. C. Lubinza** (National Institute of Medical Research, Tanzania), **Linda Hisgen** (Deutsches Primatenzentrum GmbH, Germany) and **F. M. D. Paciência** are thanked for their inspiration and work over the last years. It was and is a pleasure to supervise you.

C. Roos (Deutsches Primatenzentrum GmbH, Germany) generously supported my work since I entered the German Primate Center. Thanks for teaching me genetics. What else can I say than 'Rock n' Roll' buddy. A special thank goes also to **D. Zinner** (Deutsches Primatenzentrum GmbH, Germany) for the years-long fruitful collaboration and his expertise in behavioural research.

The German Primate Center (Deutsches Primatenzentrum GmbH, Germany), and in particular **S. Pöhlmann** is thanked for his support and for hosting my work group in the Department of Infection Biology. **S. Sahin**, **B. Burchard** and **R. Teepe** are thanked for their professional support in coordinating and managing my research projects, as well as **S. Treue** and **M. Lankheit**, are thanked for the long and trustful working relationship. Moreover, the Primate Genetics Department, in particular, **L. Walter** and **C. Schwarz** are thanked for continuous help and sharing of laboratory infrastructure.

My long-term collaborators deserve a special depth of gratitude. My current knowledge and research outcome would not be the same without their dedicated guidance, mentoring, training, discussions on my research findings and loyalty. I thank **S. Lukehart** and her lab members at the University of Washington, USA, **D. Šmajs** and his lab members at Masaryk University, Czech Republic, **J. D. Keyyu** and his team at the Tanzania Wildlife Research Institute, Tanzania, **I. A. V. Lejora** and his team at Tanzania National Parks, Tanzania, in particular **N.**

Myonga and **R. Kaitila** as well as the wonderful **park rangers** at Lake Manyara National Park, who dedicate their lives for the protection of Tanzania's natural resources and my students, **R. R. Kazwala** (Sokoine University of Agriculture, Tanzania), **S. G. M. Mfinanga** and **E. Ngadaya** both at the National Institute for Medical Research, Tanzania, **A. Sylverken** and her team at the Kumasi Center for Collaborative Research, Ghana, as well as the team around **C. Drosten**, **M. Müller**, **V. Cormann**, **F. J. Drexler** at the Charité University Medical Center, Germany, **K. J. Petrželková** (Masaryk University, Czech Republic), **T. Bohm** (African Parks, Republic of the Congo), and my colleagues from Down Under **R. J. Vaughan-Higgins** (Murdock University, Australia), **N. Buller** (Department of Primary Industries and Regional Development, Australia), **Carlo Pacioni** (Department of Environmental, Land, Water and Planning, Australia), **Ben Adler** (Monash University, Australia), **Anna Meredith** and **Pam Whiteley** (University of Melbourne, Australia).

Research Projects require a substantial amount of funding. I am particularly thankful for the financial support received by the German Research Foundation (**Deutsche Forschungsgemeinschaft**, Germany), which had a huge impact on developing my independent research track and the development of my career. Special credits must be given to **B. Schilling** and **A. Strecker** for their administrative support.

Thanks to my friends **A. A. Mwago**, **E. S. Kihwele** (Tanzania National Parks, Tanzania), and **A. Lawrenz** (Der Grüne Zoo Wuppertal, Germany) for their friendship and support throughout the years.

11. References

Abt, M. C., Pamer, E. G. (2014): Commensal bacteria mediated defenses against pathogens. Current Opinion in Immunology 29, 16–22. doi:10.1016/j.coi.2014.03.003.

Adler, B., Sattler, C., Adler, H. (2017): Herpesviruses and their host cells: A successful liaison. Trends in Microbiology 25, 229–241. doi:10.1016/j.tim.2016.11.009.

Akira, S., Takeda, K. (2004): Toll-like receptor signalling. Nature Reviews Immunology 4, 499–511. doi:10.1038/nri1391.

Alizon, S., de Roode, J. C., Michalakis, Y. (2013): Multiple infections and the evolution of virulence. Ecology Letters 16, 556–567. doi:10.1111/ele.12076.

Anand, S., Radhakrishna, S. (2017): Investigating trends in human-wildlife conflict: is conflict escalation real or imagined? Journal of Asia-Pacific Biodiversity 10, 154–161. doi:10.1016/j.japb.2017.02.003.

Anderson, R. M., May, R. M. (1981): The population dynamics of microparasites and their invertebrate hosts. Philosophical Transactions of the Royal Society B 291, 451–524. doi:10.1098/rstb.1981.0005.

Antal, G. M., Lukehart, S. A., Meheus, A. Z. (2002): The endemic treponematoses. Microbes and Infection 4, 83–94. doi:10.1016/S1286-4579(01)01513-1.

Anthony, S., Epstein, J., Murray, K., Navarrete-Macias, I., Zambrana-Torrelio, C., Solovyov, A., Ojeda-Flores, R., Arrigo, N. C., Islam, A., Khan, S. A., Hosseini, P., Bogich, T. L., Olival, K. J., Sanchez-Leon, M. D., Kareh, W. B., Goldstein, T., Luby, S. P., Morse, S. S., Mazet, J. A. K., Daszak, P., Lipkin, W. I. (2013): A strategy to estimate unknown viral diversity in mammals. MBio 4. e00598-13. doi:10.1128/mBio.00598-13.

Artois, M., Bengis, R., Delahay, R. J., Duchêne, M. J., Duff, P. J., Ferroglio, E., Gortazar, C., Hutchings, M. R., Kock, R. A., Leighton, F. A., Mörner, T., Smith, G. C. (2009): Wildlife disease surveillance and monitoring. In: R. J. Delahay, G. C. Smith, M. R. Hutchings, Management of Disease in Wild Mammals. Springer, 187–213. doi:10.1007/978-4-431-77134-0_10.

Ashford, R. (2003): When is a reservoir not a reservoir? Emerging Infectious Diseases 9, 1495.

Ashton, P., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., Wain, J., O'Grady, J. (2015): MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. Nature Biotechnology 33, 296–300. doi:10.1038/nbt.3103.

Asiedu, K., Fitzpatrick, C., Jannin, J. (2014): Eradication of yaws: historical efforts and achieving WHO's 2020 target. PLoS Neglected Tropical Diseases 8, e3016. doi:10.1371/journal.pntd.0003016.

Baer, R., Bankier, A. T., Biggin, M., D., Deininger P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C., Tuffnell, P. S., Barell, B. G. (1984): DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature 310, 207–211. doi:10.1038/310207a0.

Barasona, J. A., Gallardo, C., Cadenas-Fernández, E., Jurado, C., Rivera, B., Rodríguez-Bertos, A., Arias, M., Sanchez-Vizcaino, J. M. (2019): First oral vaccination of Eurasian wild boar against African swine fever virus genotype II. Frontiers in Veterinary Science 6, 137. doi:10.3389/fvets.2019.00137.

Barnes, H. E., Liu, G., Weston, C. Q., King, P., Pham, L. K., Waltz, S., Helzer, K. T., Day, L., Sphar, D., Yamamoto, R. T., Forsyth, R. A. (2014): Selective microbial genomic DNA isolation using restriction endonucleases. PLoS One 9, e109061. doi:10.1371/journal.pone.0109061.

Barrett, S. (2004): Eradication versus control: the economics of global infectious disease policies. Bulletin of the World Health Organization 82, 683–688.

Barrett, S. (2003): Global disease eradication. Journal of the European Economic Association 1, 591–600. doi:10.1162/154247603322391224.

Bedford, T., Malik, H. S. (2016): Did a single amino acid change make ebola virus more virulent? Cell 167, 892–894. doi:10.1016/j.cell.2016.10.032.

Belkum, V.A., Tassios, P., Dijkshoorn, L., Haeggman, S., Cookson, B., Fry, N., Fussing, V., Green, J., Feil, E., Gerner-Smidt, P., Brisse, S., Struelens, M. (2007): Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clinical Microbiology and Infection 13, 1–46. doi:10.1111/j.1469-0691.2007.01786.x.

Benelli, G., Beier, J. C. (2017): Current vector control challenges in the fight against malaria. Acta Tropica 174, 91–96. doi:10.1016/j.actatropica.2017.06.028.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J, Carter, R. J., Cheetham, R. K., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., Irving, L. J., Karbelashvili, M. S., Kirk, S. M., Li, H., Liu, X., Maisinger, K. S., Murray, L. J., Obradovic, B., Ost, T., Parkinson, M. L., Pratt, M. R., Rasolonjatovo, I. M. J., Reed, M. T., Rigatti, R., Rodighiero, C., Ross, M. T., Sabot, A., Sankar, S. V., Scally, A., Schroth, G. P., Smith, M. E., Smith, V. P., Spiridou, A., Torrance, P. E., Tzonev, S. S., Vermaas, E. H., Walter, K., Wu, X., Zhang, L., Alam, M. D., Anastasi, C., Aniebo, I. C., Bailey, D. M. D., Bancarz, I. R., Banerjee, S., Barbour, S. G., Baybayan, P. A., Benoit, V. A., Benson, K. F., Bevis, C., Black, P. J., Boodhun, A., Brennan, J. S., Bridgham, J. A., Brown, R. C., Brown, A. A., Buermann, D. H., Bundu, A. A., Burrows, J. C., Carter, N. P., Castillo, N., Catenazzi, M. C. E., Chang, S., Cooley, R. N., Crake, N. R., Dada, O. O., Diakoumakos, K. D., Dominguez-Fernandez, B., Earnshaw, D. J., Egbujor, U. C., Elmore, D. W., Etchin, S. S., Ewan, M. R., Fedurco, M., Fraser, L. J., Fuentes Fajardo, K. V., Furey, W. S., George, D., Gietzen, K. J., Goddard, C. P., Golda, G. S., Granieri, P. A., Green, D. E., Gustafson, D. L., Hansen, N. F., Harnish, K., Haudenschild, C. D., Heyer, N. I., Hims, M. M., Ho, J. T., Horgan, A. M., Hoschler, K., Hurwitz, S., Ivanov, D. V., Johnson, M. Q., James, T., Jones, T. A. H., Kang, G. D., Kerelska, T. H., Kersey, A. D., Khrebtukova, I., Kindwall, A. P., Kingsbury, Z., Kokko-Gonzales, P. I., Kumar, A., Laurent, M. A., Lawley, C. T., Lee, S. E., Lee, X., Liao, A. K., Loch, J. A., Lok, M., Luo, S., Mammen, R. M., Martin, J. W., McCauley, P. G., McNitt, P., Mehta, P., Moon, K. W., Mullens, J. W., Newington, T., Ning, Z., Ng, B. L., Novo, S. M., O'Neill, M. J., Osborne, M. A., Osnowski, A., Ostadan, O., Paraschos, L. L., Pickering, L., Pike, A. C., Pike, A. C., Pinkard, D. C., Pliskin, D. P., Podhasky, J., Quijano, V. J., Raczy, C., Rae, V. H., Rawlings, S. R., Rodriguez, A. C., Roe, P. M., Rogers, J., Rogert Bacigalupo, M. C., Romanov, N., Romieu, A., Roth, R. K., Rourke, N. J., Ruediger, S. T., Rusman, E., Sanches-Kuiper, R. M., Schenker, M. R., Seoane, J. M., Shaw, R. J., Shiver, M. K., Short, S. W., Sizto, N. L., Sluis, J. P., Smith, M. A., Sohna Sohna, J. E., Spence, E. J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C. L., Turcatti, G., vandeVondele, S., Verhovsky, Y., Virk, S. M., Wakelin, S.,

Walcott, G. C., Wang, J., Worsley, G. J., Yan, J., Yau, L., Zuerlein, M., Rogers, J., Mullikin, J. C., Hurles, M. E., McCooke, N. J., West, J. S., Oaks, F. L., Lundberg, P. L., Klenerman, D., Durbin, R., Smith, A. J. (2008): Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456, 53–59. doi:10.1038/nature07517.

Biswas, G., Sankara, D. P., Agua-Agum, J., Maiga, A. (2013): Dracunculiasis (guinea worm disease): eradication without a drug or a vaccine. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 368, 20120146. doi:10.1098/rstb.2012.0146.

Blancou, J., Artois, M., Gilot-Fromont, E., Kaden, V., Rossi, S., Smith, G. C., Hutchings, M. R., Chambers, M. A., Houghton, S., Delahay, R. J. (2009): Options for the control of disease 1: targeting the infectious or parasitic agent. In: R. J. Delahay, G. C. Smith, M. R. Hutchings, Management of disease in wild mammals. Springer, 97-120. doi:10.1007/978-4-431-77134-0_6.

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., Abebe, E. (2005): Defining operational taxonomic units using DNA barcode data. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 360, 1935–1943. doi:10.1098/rstb.2005.1725.

Bolzoni, L., Leo, G. A. (2013): Unexpected consequences of culling on the eradication of wildlife diseases: the role of virulence evolution. The American Naturalist 181, 301–313. doi:10.1086/669154.

Bonilla, F., Oettgen, H. (2010): Adaptive immunity. Journal of Allergy and Clinical Immunology 125, 33-40. doi:10.1016/j.jaci.2009.09.017.

Boots, M. (2015): The need for evolutionarily rational disease interventions: Vaccination can select for higher virulence. PLOS Biology 13, e1002236. doi:10.1371/journal.pbio.1002236.

Botten, J., Mirowsky, K., Kusewitt, D., Bharadwaj, M., Yee, J., Ricci, R., Feddersen, R. M., Hjelle, B. (2000): Experimental infection model for Sin Nombre hantavirus in the deer mouse (*Peromyscus maniculatus*). Proceedings of the National Academy of Sciences 97, 10578–10583. doi:10.1073/pnas.180197197.

Brearley, G., Rhodes, J., Bradley, A., Baxter, G., Seabrook, L., Lunney, D., Liu, Y., McAlpine, C. (2013): Wildlife disease prevalence in human-modified landscapes. Biological Reviews 88, 427–442. doi:10.1111/brv.12009.

Brooks, J. P., Edwards, D. J., Harwich, M. D., Rivera, M. C., Fettweis, J. M., Serrano, M. G., Reris, R. A., Sheth, N. U., Huang, B., Girerd, P., Vaginal Microbiome Consortium (additional members), Strauss III, J. F., Jefferson, K. K., Buck, G. A. (2015): The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. BMC Microbiology 15, 66. doi:10.1186/s12866-015-0351-6.

Buck, J., Ripple, W. (2017): Infectious agents trigger trophic cascades. Trends in Ecology & Evolution 32, 681–694. doi:10.1016/j.tree.2017.06.009.

Buermans, H., den Dunnen, J. (2014): Next generation sequencing technology: Advances and applications. Biochimica et Biophysica Acta 1842, 1932–1941. doi:10.1016/j.bbadis.2014.06.015.

Buffie, C. G., Bucci, V., Stein, R. R., McKenney, P.T., Ling, L., Gobourne, A., No, D., Liu, H., Kinnebrew, M., Viale, A., Littmann, E., van den Brink, M. R. M., Jenq, R. R., Taur, Y., Sander, C., Cross, J. R., Toussaint, N. C., Xavier, J. B., Pamer, E. G. (2015): Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. Nature 517, 205–208. doi:10.1038/nature13828.

Burmeister, A., Lenski, R., Meyer, J. (2016): Host coevolution alters the adaptive landscape of a virus. Proceedings of the Royal Society B: Biological Sciences 283. doi:10.1098/rspb.2016.1528.

Bustamante-Marin, X. M., Ostrowski, L. E. (2017): Cilia and mucociliary clearance. Cold Spring Harbor Perspectives in Biology 9. doi:10.1101/cshperspect.a028241.

Buyon, L., Slaven, R., Emerson, P. M., King, J., Debrah, O., Aboe, A., Ruiz-Tiben, E. (2018): Achieving the endgame: Integrated NTD case searches. PLoS Neglected Tropical Diseases 12, e0006623. doi:10.1371/journal.pntd.0006623.

Callaway, E. (2016): Dogs thwart effort to eradicate Guinea worm. Nature 529, 10–11. doi:10.1038/529010a.

Caplan, A. L. (2009): Is disease eradication ethical? The Lancet 373, 2192–2193. doi:10.1016/S0140-6736(09)61179-X.

Carroll, D., Daszak, P., Wolfe, N. D., Gao, G. F., Morel, C. M., Morzaria, S., Pablos-Mendez, A., Tomori, O., Mazet, J. A. K. (2018): The global virome project. Science 359, 872–874. doi:10.1126/science.aap7463.

Carter, S. P., Delahay, R. J., Smith, G. C., Macdonald, D. W., Riordan, P., Etherington, T. R., Pimley, E. R., Walker, N. J., Cheeseman, C. L. (2007): Culling-induced social perturbation in Eurasian badgers *Meles meles* and the management of TB in cattle: an analysis of a critical problem in applied ecology. Proceedings Biological Sciences 274, 2769–2777. doi:10.1098/rspb.2007.0998.

Casadevall, A., Pirofski, L. (1999): Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. Infection and Immunity 67, 3703–3713.

Castellani, A. (1907): Experimental investigations on framboesia tropica (yaws). The Journal of Hygiene (London) 7, 558–569.

Caulfield, T., Sipp, D., Murry, C. E., Daley, G. Q., Kimmelman, J. (2016): Scientific community. Confronting stem cell hype. Science (New York, NY) 352, 776–777. doi:10.1126/science.aaf4620.

Chen, Y., Zhang, W., Knabel, S. J. (2007): Multi-virulence-locus sequence typing identifies single nucleotide polymorphisms which differentiate epidemic clones and outbreak strains of Listeria monocytogenes. Journal of Clinical Microbiology 45, 835–846. doi:10.1128/JCM.01575-06.

Cheng, X. (1995): DNA modification by methyltransferases. Current Opinion in Structural Biology 5, 4–10.

Chiou, K. L., Bergey, C. M. (2018): Methylation-based enrichment facilitates low-cost, noninvasive genomic scale sequencing of populations from feces. Scientific Reports 8, 1975. doi:10.1038/s41598-018-20427-9.

Chirag, S., Fomda, B. A., Khan, A., Malik, A. A., Lone, G. N., Khan, B. A., Zahoor, D. (2015): Detection of hydatid-specific antibodies in the serum and urine for the diagnosis of cystic echinococcosis in patients from the Kashmir Valley, India. Journal of Helminthology 89, 232–237. doi:10.1017/S0022149X13000837.

Chuma, I. S., Batamuzi, E. K., Collins, D., Fyumagwa, R. D., Hallmaier-Wacker, L. K., Kazwala, R. R., Keyyu, J. D., Lejora, I. A., Lipende, I. F., Lüert, S., Paciencia, F. M. D., Piel, A., Stewart, F. A., Zinner, D., Roos, C., Knauf, S. (2018): Widespread *Treponema pallidum* infection in nonhuman primates, Tanzania. Emerging Infectious Diseases 24, 1002–1009. doi:10.3201/eid2406.180037.

Chuma, I. S., Roos, C., Atickem, A., Bohm, T., Collins, D. A., Grillová, L., Hallmaier-Wacker, L. K., Kazwala, R. R., Keyyu, J. D., Lüert, S., Maloueki, U., Oppelt, J., Petrželková, K. J., Piel, A., Stewart, F. A., Šmajs, D., Knauf, S. (2019): Strain diversity of *Treponema pallidum* subsp. *pertenue* suggests rare interspecies transmission in African nonhuman primates. Scientific Reports 9, 14243. doi:10.1038/s41598-019-50779-9.

Cincotta, R., Wisnewski, J., Engelman, R. (2000): Human population in the biodiversity hotspots. Nature 404, 990–992. doi:10.1038/35010105.

Cohan, F. M. (2002): What are bacterial species? Annual Review of Microbiology 56, 457–487. doi:10.1146/annurev.micro.56.012302.160634.

Cohan, F. M., Kopac, S. (2017): Genetics and evolution of infectious diseases (Second Edition). Sciencedirect, 25–49. doi:10.1016/B978-0-12-799942-5.00002-0.

Collins, F. S., Hamburg, M.A. (2013): First FDA authorization for next-generation sequencer. The New England Journal of Medicine 369, 2369–2371. doi:10.1056/NEJMp1314561.

Dass, S. A., Vasudevan, A., Dutta, D., Soh, L. J., Polsky, R., Vyas, A. (2011): Protozoan parasite *Toxoplasma gondii* manipulates mate choice in rats by enhancing attractiveness of males. PloS One 6, e27229. doi:10.1371/journal.pone.0027229.

Daszak, P., Cunningham, A., Hyatt, A. (2000): Emerging infectious diseases of wildlife--threats to biodiversity and human health. Science 287, 443-449. doi:10.1126/science.287.5452.443.

DeCandia, A. L., Dobson, A. P., von Holdt, B. M. (2018): Toward an integrative molecular approach to wildlife disease. Conservation Biology: The Journal of the Society for Conservation Biology 32, 798-807. doi:10.1111/cobi.13083.

Dejnirattisai, W., Supasa, P., Wongwiwat, W., Rouvinski, A., Barba-Spaeth, G., Duangchinda, T., Sakuntabhai, A., Cao-Lormeau, V. M., Malasit, P., Rey, F. A., Mongkolsapaya, J., Screaton, G. R. (2016): Dengue virus sero-cross-reactivity drives antibody-dependent enhancement of infection with zika virus. Nature Immunology 17, 1102–1108. doi:10.1038/ni.3515.

Delahay, R., Smith, G. C., Hutchings, M. R. (2009): The science of wildlife disease management. In: R. J. Delahay, G. C. Smith, M. R. Hutchings, Management of disease in wild mammals. Springer, 1–8. doi:10.1007/978-4-431-77134-0_1.

Dennison, N. J., Jupatanakul, N., Dimopoulos, G. (2014): The mosquito microbiota influences vector competence for human pathogens. Current Opinion in Insect Science 3, 6–13. doi:10.1016/j.cois.2014.07.004.

Destoumieux-Garzón, D., Mavingui, P., Boetsch, G., Boissier, J., Darriet, F., Duboz, P., Fritsch, C., Giraudoux, P., Le Roux, F., Morand, S., Paillard, C., Pontier, D., Sueur, C., Voituron, Y. (2018): The One Health concept: 10 years old and a long road ahead. Frontiers in Veterinary Science 5, 14. doi:10.3389/fvets.2018.00014.

Dittmann, C., Han, H. M., Grabenbauer, M., Laue, M. (2015): Dormant *Bacillus* spores protect their DNA in crystalline nucleoids against environmental stress. Journal of Structural Biology 191, 156–164. doi:10.1016/j.jsb.2015.06.019.

Djouaka, R., Zeukeng, F., Bigoga, J., Kakou-Ngazoa, S. E., Akoton, R., Tchigossou, G., Coulibaly, D. N., Tchebe, S. J. E., Aboubacar, S., Nguepdjo, C. N., Tossou, E., Adeoti, R., Nsonga, T. M. N., Akpo, Y., Djegbe, I., Tamo, M., Mbacham, W. F., Ablordey, A. (2018): Domestic animals infected with *Mycobacterium ulcerans* – Implications for transmission to

humans. PLOS Neglected Tropical Diseases 12, e0006572. doi:10.1371/journal.pntd.0006572.

Dobson, A., Foufopoulos, J. (2001): Emerging infectious pathogens of wildlife. Philosophical Transactions of The Royal Society B: Biological Sciences 356, 1001–1012. doi:10.1098/rstb.2001.0900.

Dowdle, W. (1998): The principles of disease elimination and eradication. Bulletin of the World Health Organization 76, 22–25.

Drexler, J. F., Corman, V. M., Drosten, C. (2014): Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. Antiviral Research 101, 45–56. doi:10.1016/j.antiviral.2013.10.013.

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A. M., Berger, A., Burguiere, A. M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Müller, S., Rickerts, V., Stürmer, M., Vieth, S., Klenk, H. D., Osterhaus, A., Schmitz, H., Doerr, H. W. (2003): Identification of a novel coronavirus in patients with severe acute respiratory syndrome. The New England Journal of Medicine 348, 1967–1976. doi:10.1056/NEJMoa030747.

Dyson, L., Marks, M., Crook, O., Sokana, O., Solomon, A., Bishop, A., Mabey, D. C. W., Hollingsworth, T. D. (2018): Targeted treatment of yaws with household contact tracing: How much do we miss? American Journal of Epidemiology 187, 837–844. doi:10.1093/aje/kwx305.

Eberhard, M., Cleveland, C., Zirimwabagabo, H., Yabsley, M., Ouakou, P., Ruiz-Tiben, E. (2016): Guinea worm (*Dracunculus medinensis*) infection in a wild-caught frog, Chad. Emerging Infectious Diseases 22, 1961–1962. doi:10.3201/eid2211.161332.

Eberhard, M., Ruiz-Tiben, E., Hopkins, D., Farrell, C., Toe, F., Weiss, A., Withers Jr. P. C., Jenks, M. H., Thiele, E. A., Cotton, J. A., Hance, Z., Holroyd, N., Cama, V. A., Tahir, M. A., Mounda, T. (2014): The peculiar epidemiology of dracunculiasis in Chad. The American Journal of Tropical Medicine and Hygiene 90, 61–70. doi:10.4269/ajtmh.13-0554.

Edgar, R. C. (2018a): Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. PeerJ 6, e4652. doi:10.7717/peerj.4652.

Edgar, R. C. (2018b): Updating the 97% identity threshold for 16S ribosomal RNA OTUs. Bioinformatics (Oxford, England) 34, 2371–2375. doi:10.1093/bioinformatics/bty113.

Edmondson, D. G., Hu, B., Norris, S. J. (2018): Long-term *in vitro* culture of the syphilis spirochete *Treponema pallidum* subsp. *pallidum*. MBio 9, e01153-18. doi:10.1128/mBio.01153-18.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., de Winter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J., Turner, S. (2009): Real-time DNA sequencing from single polymerase molecules. Science 323, 133–138. doi: 10.1126/science.1162986.

El-Metwally, S., Ouda, O. M., Helmy, M. (2014): First- and next-generation sequencing methods. In: S. El-Metwally, O. M. Ouda, M. Helmy, Next generation sequencing technologies and challenges in sequence assembly. Springer, 29–36.

Enserink, M. (2018): On a remote Pacific island, this doctor has revived a 60-year quest to eradicate a disfiguring disease. Science. doi:10.1126/science.aau8267.

Ewald, P. W. (1998): The evolution of virulence and emerging diseases. Journal of Urban Health: Bulletin of the New York Academy of Medicine 75, 480–491. doi:10.1007/BF02427686.

Ewald, P. W. (1996): Guarding against the most dangerous emerging pathogens. Emerging Infectious Diseases 2, 245–257. doi:10.3201/eid0204.960401.

Ewald, P. W. (1995): The evolution of virulence: a unifying link between parasitology and ecology. Journal of Parasitology 81, 659–669.

Ewald, P. W. (1991a): Waterborne transmission and the evolution of virulence among gastrointestinal bacteria. Epidemiology and Infection 106, 83–119. doi:10.1017/s0950268800056478.

Ewald, P. W., 1994. Evolution of infectious diseases. Oxford University Press. ISBN: 0-19-506058-X.

Ewald, P. W. (1991b): Transmission modes and the evolution of virulence. Human Nature 2, 1–30. doi:10.1007/BF02692179.

Ewald, P. W. (1987): Pathogens-induced cycling of outbreak insect populations. In: P. Barbosa, J. C. Schultz, Insect outbreaks. Academic Press, 269–286.

Ewald, P. W. (1983): Host-parasite relations, vectors, and the evolution of disease severity. Annual Review of Ecology and Systematics 14, 465–485.

Feehery, G. R., Yigit, E., Oyola, S. O., Langhorst, B. W., Schmidt, V. T., Stewart, F. J., Dimalanta, E. T., Amaral-Zettler, L. A., Davis, T., Quail, M. A., Pradhan, S. (2013): A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. PLoS ONE 8, e76096. doi:10.1371/journal.pone.0076096.

Fenner, F. (1982): Global eradication of smallpox. Clinical Infectious Diseases 4, 916–930. doi:10.1093/clinids/4.5.916.

Fenner, F., Henderson, D., Arita, I., Jezek, Z., Ladnyi, I. (1988): Smallpox and its eradication. 1st ed. World Health Organization. ISBN: 9241561106.

Flegr, J. (2007): Effects of *Toxoplasma* on human behavior. Schizophrenia Bulletin 33, 757–760. doi:10.1093/schbul/sbl074.

Foley, S. L., Lynne, A. M., Nayak, R. (2009): Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 9, 430–440. doi:10.1016/j.meegid.2009.03.004.

Frank, S. (1996): Models of parasite virulence. The Quarterly Review of Biology 71, 37–78.

Fraser, C., Alm, E. J., Polz, M. F., Spratt, B. G., Hanage, W. P. (2009): The bacterial species challenge: making sense of genetic and ecological diversity. Science (New York, NY) 323, 741–746. doi:10.1126/science.1159388.

Fraser, C., Norris, S., Weinstock, G., White, O., Sutton, G., Dodson, R., Gwinn, M., Hickey, E. K., Clayton, R., Ketchum, K. A., Sodergren, E., Hardham, J. M., McLeod, M. P., Salzberg, S., Peterson, J., Khalak, H., Richardson, D., Howell, J. K., Chidambaram, M., Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M. D., Fujii, C., Garland, S., Hatch, B., Horst, K., Roberts, K., Sandusky, M., Weidman, J., Smith, H. O., Venter, J. C. (1998): Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science 281, 375–388. doi:10.1126/science.281.5375.375.

Fribourg-Blanc, A., Mollaret, H. H. (1969): Natural treponematosis of the African primate. Primates in Medical Research 3, 113–121.

Fribourg-Blanc, A., Niel, G., Mollaret, H. H. (1963): Serologic and microscopic confirmation of treponemosis in Guinea baboons. Bulletin de la Societe de Pathologie Exotique et de ses Filiales 56, 474–485.

Galán-Puchades, T. M. (2016): Dogs and Guinea worm eradication. The Lancet Infectious Diseases 16, 770. doi:10.1016/S1473-3099(16)30080-9.

Galán-Puchades, T. M. (2017): The Guinea worm: A zoonotic parasite of dogs. Trends in

Parasitology 34, 3-4. doi:10.1016/j.pt.2017.10.003.

Gardner, I., Hietala, S., Boyce, W. (1996): Validity of using serological tests for diagnosis of diseases in wild animals. Revue Scientifique et Technique (International Office of Epizootics) 15, 323–335.

George, S., Levecke, B., Kattula, D., Velusamy, V., Roy, S., Geldhof, P., Sarkar, R., Kang, G. (2016): Molecular identification of hookworm isolates in humans, dogs and soil in a tribal area in Tamil Nadu, India. PLoS Neglected Tropical Diseases 10, e0004891. doi:10.1371/journal.pntd.0004891.

Gerland, P., Raftery, A. E., Sevčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G. K., Wilmoth, J. (2014): World population stabilization unlikely this century. Science (New York, NY) 346, 234–237. doi:10.1126/science.1257469.

Gerson, C., Sabater, J., Scuri, M., Torbati, A., Coffey, R., Abraham, J., Lauredo, I., Forteza, R., Wanner, A., Salathe, M., Abraham, W. M. (2000): The lactoperoxidase system functions in bacterial clearance of airways. American Journal of Respiratory Cell and Molecular Biology 22, 665–671. doi:10.1165/ajrcmb.22.6.3980.

Giacani, L., Lukehart, S. (2014): The endemic treponematoses. Clinical Microbiology Reviews 27, 89–115. doi:10.1128/CMR.00070-13.

Gill, S. R., Pop, M., DeBoy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S, Gordon, J. I., Relman, D. A., Fraser-Liggett, C. M., Nelson, K. E. (2006): Metagenomic analysis of the human distal gut microbiome. Science 312, 1355–1359. doi:10.1126/science.1124234.

Glaeser, S., Kampfer, P. (2015): Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. Systematic and Applied Microbiology 38, 237–245. doi:10.1016/j.syapm.2015.03.007.

Godfroid, J. (2017): Brucellosis in livestock and wildlife: zoonotic diseases without pandemic potential in need of innovative one health approaches. Archives of Public Health (Archives Belges de Sante Publique) 75, 34. doi:10.1186/s13690-017-0207-7.

Godornes, C., Giacani, L., Barry, A., Mitja, O., Lukehart, S. (2017): Development of a multilocus sequence typing (MLST) scheme for *Treponema pallidum* subsp. *pertenue*: Application to yaws in Lihir Island, Papua New Guinea. PLoS Neglected Tropical Diseases 11, e0006113. doi:10.1371/journal.pntd.0006113.

Gogarten, J. F., Düx, A., Mubemba, B., Pléh, K., Hoffmann, C., Mielke, A., Müller-Tiburtius, J., Sachse, A., Wittig, R. M., Calvignac-Spencer, S., Leedertz, F. H. (2019): Tropical rainforest flies carrying pathogens form stable associations with social non-human

primates. Molecular Ecology. doi:10.1111/mec.15145.

Gomes-Solecki, M., Santecchia, I., Werts, C. (2017): Animal models of leptospirosis: Of mice and hamsters. Frontiers in Immunology 8, 58. doi:10.3389/fimmu.2017.00058.

Gómez, P., Ashby, B., Buckling, A. (2014): Population mixing promotes arms race host–parasite coevolution. Proceedings of the Royal Society of London B: Biological Sciences 282, 20142297–20142297. doi:10.1098/rspb.2014.2297.

Goodenough, A. E., Harrell, A. N., Keating, R. L., Rolfe, R. N., Stubbs, H., MacTavish, L., Hart, A. G. (2017): Managing grassland for wildlife: the effects of rotational burning on tick presence and abundance in African savannah habitat. Wildlife Biology 2017. doi:10.2981/wlb.00318.

Gortazar, C., Diez-Delgado, I., Barasona, J., Vicente, J., Fuente, J., Boadella, M. (2015): The wild side of disease control at the wildlife-livestock-human interface: A review. Frontiers in Veterinary Science 1, 27. doi:10.3389/fvets.2014.00027.

Grillová, L., Bawa, T., Mikalová, L., Gayet-Ageron, A., Nieselt, K., Strouhal, M., Sednaoui, P., Ferry, T., Cavassini, M., Lautenschlager, S., Dutly, F., Pla-Diaz, M., Krützen, M., Gonzalez-Candelas, F., Bagheri, H. C., Smajs, D., Arora, N., Bosshard, P. P. (2018): Molecular characterization of *Treponema pallidum* subsp. *pallidum* in Switzerland and France with a new multilocus sequence typing scheme. PloS One 13, e0200773. doi:10.1371/journal.pone.0200773.

Hallmaier-Wacker, L. K., Lueert, S., Roos, C., Knauf, S. (2018): The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. Scientific Reports 8, 6292. doi:10.1038/s41598-018-24573-y.

Hallmaier-Wacker, L. K., Munster, V. J., Knauf, S. (2017): Disease reservoirs: from conceptual frameworks to applicable criteria. Emerging Microbes & Infections 6, e79. doi:10.1038/emi.2017.65.

Harper, K. N., Fyumagwa, R. D., Hoare, R., Wambura, P. N., Coppenhaver, D. H., Sapolsky, R. M., Alberts, S. C., Tung, J., Rogers, J., Kilewo, M., Batamuzi, E. K., Leendertz, F. H., Armelagos, G. J., Knauf, S. (2012): *Treponema pallidum* infection in the wild baboons of East Africa: distribution and genetic characterization of the strains responsible. PLoS One 7, e50882. doi:10.1371/journal.pone.0050882.

Harper, K. N., Knauf, S. (2013): *Treponema pallidum* Infection in Primates: Clinical Manifestations, Epidemiology, and Evolution of a Stealthy Pathogen. In: J. F. Brinkworth, K. Pechenkina, Primates, pathogens, and evolution. Springer, 189–219. doi:10.1007/978-1-4614-7181-3_7.

Harrison, A., Newey, S., Gilbert, L., Haydon, D. T., Thirgood, S. (2010): Culling wildlife hosts to control disease: mountain hares, red grouse and louping ill virus. Journal of Applied Ecology 47, 926–930. doi:10.1111/j.1365-2664.2010.01834.x.

Hassell, J. M., Begon, M., Ward, M. J., Fèvre, E. M. (2017): Urbanization and disease emergence: Dynamics at the wildlife–livestock–human interface. Trends in Ecology & Evolution 32, 55–67. doi:10.1016/j.tree.2016.09.012.

Hattman, S. (2005): DNA-[adenine] methylation in lower eukaryotes. Biochemistry 70, 550–558.

Haydon, D., Cleaveland, S., Taylor, L., Laurenson, M. (2002): Identifying reservoirs of infection: a conceptual and practical challenge. Emerging Infectious Diseases 8, 1468–1473. doi:10.3201/eid0812.010317.

Haydon, D., Randall, D., Matthews, L., Knobel, D. (2006): Low-coverage vaccination strategies for the conservation of endangered species. Nature 443, 692–695. doi:10.1038/nature05177.

Hayman, D. T., Luis, A. D., Restif, O., Baker, K. S., Fooks, A. R., Leach, C., Horton, D. L., Suu-Ire, R., Cunningham, A. A., Wood, J. L. N., Webb, C. T. (2018): Maternal antibody and the maintenance of a lyssavirus in populations of seasonally breeding African bats. PloS One 13, e0198563. doi:10.1371/journal.pone.0198563.

Heather, J. M., Chain, B. (2016): The sequence of sequencers: The history of sequencing DNA. Genomics 107, 1–8. doi:10.1016/j.ygeno.2015.11.003.

Heil, M. (2016): Host manipulation by parasites: Cases, patterns, and remaining doubts. Frontiers in Ecology and Evolution 4, 80. doi:10.3389/fevo.2016.00080.

Hirsch, V., Dapolito, G., Goeken, R., Campbell, B. (1995): Phylogeny and natural history of the primate lentiviruses, SIV and HIV. Current Opinion in Genetics & Development 5, 798–806.

Hoeprich, P. D. (1989): Host-parasite relationships and the pathogenesis of infectious disease. In: P. D. Hoeprich, M. C. Jordan, Infectious diseases, 4th edition. J.B. Lippincott Company, 41-53.

Hoffmann, C., Zimmermann, F., Biek, R., Kuehl, H., Nowak, K., Mundry, R., Agbor, A., Angedakin, S., Arandjelovic, M., Blankenburg, A., Brazolla, G., Corogenes, K., Couacy-Hymann, E., Deschner, T., Dieguez, P., Dierks, K., Düx, A., Dupke, S., Eshuis, H., Formenty, P., Yuh, Y. G., Goedmakers, A., Gogarten, J. F., Granjon, A. C., McGraw, S., Grunow, R., Hart, J., Jones, S., Junker, J., Kiang, J., Langergraber, K., Lapuente, J., Lee, K., Leendertz, S. A., Leguillon, F., Leinert, V., Löhrich, T., Marrocoli, S., Mätz-Rensing, K.,

Meier, A., Merkel, K., Metzger, S., Murai, M., Niedorf, S., De Nys, H., Sachse, A., van Schijndel, J., Thiesen, U., Ton, E., Wu, D., Wieler, L. H., Boesch, C., Klee, S. R., Wittig, R. M., Calvignac-Spencer, S., Leedertz, F. H. (2017): Persistent anthrax as a major driver of wildlife mortality in a tropical rainforest. Nature 548, 82–86. doi:10.1038/nature23309.

Hopkins, D. R. (2013): Disease eradication. The New England Journal of Medicine 368, 54–63. doi:10.1056/NEJMra1200391.

Houinei, W., Godornes, C., Kapa, A., Knauf, S., Mooring, E., Gonzalez-Beiras, C., Watup, R., Paru, R., Advent, P., Bieb, S., Sanz, S., Bassat, Q., Spinola, S. M., Lukehart, S. A., Mitja, O. (2017): *Haemophilus ducreyi* DNA is detectable on the skin of asymptomatic children, flies and fomites in villages of Papua New Guinea. PLoS Neglected Tropical Diseases 11, e0004958. doi:10.1371/journal.pntd.0004958.

Hudson, P., Dobson, A., Lafferty, K. (2006): Is a healthy ecosystem one that is rich in parasites? Trends in Ecology and Evolultion 21, 381–385. doi:10.1016/j.tree.2006.04.007.

Huff, J., Barry, P. (2003): B-virus (Cercopithecine herpesvirus 1) infection in humans and macaques: potential for zoonotic disease. Emerging Infectious Diseases 9, 246–250. doi:10.3201/eid0902.020272.

Hunkapiller, T., Kaiser, R., Koop, B., Hood, L. (1991): Large-scale and automated DNA sequence determination. Science (New York, NY) 254, 59–67. doi:10.1126/science.1925562.

Illumina (2017): An introduction to next-generation sequencing technology. Illumina Inc. Pub. No. 7709-2012-008-B.

Ip, C. L., Loose, M., Tyson, J. R., de Cesare, M., Brown, B.L., Jain, M., Leggett, R. M., Eccless, D. A., Zalunin, V., Urban, J. M., Piazza, P., Bowden, R. J., Paten, B., Mwaigwisya, S., Batty, E. M., Simpson, J. T., Snutch, T. P., Birney, E., Buck, D., Goodwin, S., Jansen, H. J., O'Grady, J., Olsen, H. E. (2015): MinION analysis and reference consortium: Phase 1 data release and analysis. F1000Research 4, 1075. doi:10.12688/f1000research.7201.1.

Iritani, R., Visher, E., Boots, M. (2019): The evolution of stage-specific virulence: Differential selection of parasites in juveniles. Evolution Letters 3, 162–172. doi:10.1002/evl3.105.

Janeway, C. A., Medzhitov, R. (2002): Innate immune recognition. Annual Review of Immunology 20, 197–216. doi:10.1146/annurev.immunol.20.083001.084359.

Janowicz, A., Massis, F., Ancora, M., Cammà, C., Patavino, C., Battisti, A., Prior, K., Harmsen, D., Scholz, H., Zilli, K., Sacchini, L., Di Giannatale, E., Garofolo, G. (2018): Core genome multilocus sequence typing and single nucleotide polymorphism analysis in the epidemiology of *Brucella melitensis* infections. Journal of Clinical Microbiology 56.

doi:10.1128/JCM.00517-18.

Jeltsch, A. (2002): Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem: A European Journal of Chemical Biology 3, 274–293.

Johnson, P. T., Calhoun, D. M., Riepe, T., vitt-Galles, T., Koprivnikar, J. (2019): Community disassembly and disease: realistic—but not randomized—biodiversity losses enhance parasite transmission. Proceedings of the Royal Society B: Biological Sciences 286, 20190260. doi:10.1098/rspb.2019.0260.

Johnson, S. K., Fitza, M. A., Lerner, D. A., Calhoun, D. M., Beldon, M. A., Chan, E. T., Johnson, P. T. J. (2018): Risky business: linking *Toxoplasma gondii* infection and entrepreneurship behaviours across individuals and countries. Proceedings of the Royal Society B 285, 20180822. doi:10.1098/rspb.2018.0822.

Johnson, S. S., Zaikova, E., Goerlitz, D. S., Bai, Y., Tighe, S. W. (2017): Real-time DNA sequencing in the Antarctic dry valleys using the Oxford Nanopore sequencer. Journal of Biomolecular Techniques 28, 2–7. doi:10.7171/jbt.17-2801-009.

Jolley, K. A., Bliss, C. M., Bennett, J. S., Bratcher, H. B., Brehony, C., Colles, F. M., Wimalarathna, H., Harrison, O. B., Sheppard, S. K., Cody, A. J., Maiden, M. C. J. (2012): Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. Microbiology 158, 1005–1015. doi:10.1099/mic.0.055459-0.

Jolley, K. A., Maiden, M. C. (2014): Using multilocus sequence typing to study bacterial variation: prospects in the genomic era. Future Microbiology 9, 623–630. doi:10.2217/fmb.14.24.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., Daszak, P. (2008): Global trends in emerging infectious diseases. Nature 451, 990. doi:10.1038/nature06536.

Kahn, L. H., Kaplan, B., Monath, T. P., Steele, J. H. (2008): Teaching “one medicine, one health”. The American Journal of Medicine 121, 169–170. doi:10.1016/j.amjmed.2007.09.023.

Karesh, W., Dobson, A., Lloyd-Smith, J., Lubroth, J., Dixon, M., Bennett, M., Aldrich, S., Harrington, T., Formenty, P., Loh, E. H., Machalaba, C. C., Thomas, M. J., Heymann, D. L. (2012): Ecology of zoonoses: natural and unnatural histories. Lancet 380, 1936–1945. doi:10.1016/S0140-6736(12)61678-X.

Katz, D. L., Elmore, J. G., Wild, D. M., Lucan, S. C., 2014. Basic epidemiologic concepts and principles. In: D. L. Katz, J. G. Elmore, D. M. Wild, S. C. Lucan, Jekel’s epidemiology,

biostatistics, preventive medicine and public health, 4th edition. Elsevier Saunders, 3-15.

Keesing, F., Belden, L., Daszak, P., Dobson, A., Harvell, C., Holt, R., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., Ostfeld, R. S. (2010): Impacts of biodiversity on the emergence and transmission of infectious diseases. Nature 468, 647–652. doi:10.1038/nature09575.

Kingry, L. C., Rowe, L. A., Respicio-Kingry, L. B., Beard, C. B., Schriefer, M. E., Petersen, J. M. (2016): Whole genome multilocus sequence typing as an epidemiologic tool for *Yersinia pestis*. Diagnostic Microbiology and Infectious Disease 84, 275–280. doi:10.1016/j.diagmicrobio.2015.12.003.

Klegarth, A., Ezeonwu, C., Rompis, A., Lee, B., Aggimarangsee, N., Chalise, M., Cortes, J., Feeroz, M., Molini, B. J., Godornes, B. C., Marks, M., Schillaci, M., Engel, G., Knauf, S., Lukehart, S. A., Jones-Engel, L. (2017): Survey of treponemal infections in free-ranging and captive macaques, 1999-2012. Emerging Infectious Diseases 23, 816–819. doi:10.3201/eid2305.161838.

Knauf, S., Barnett, U., Maciej, P., Klapproth, M., Ndao, I., Frischmann, S., Fischer, J., Zinner, D., Liu, H. (2015a): High prevalence of antibodies against the bacterium *Treponema pallidum* in Senegalese Guinea baboons (*Papio papio*). PLoS One 10, e0143100. doi:10.1371/journal.pone.0143100.

Knauf, S., Batamuzi, E., Mlengeya, T., Kilewo, M., Lejora, I., Nordhoff, M., Ehlers, B., Harper, K. N., Fyumagwa, R., Hoare, R., Failing, K., Wehrend, A., Kaup, F. J, Leendertz, F. H., Mätz-Rensing, K. (2011): *Treponema* infection associated with genital ulceration in wild baboons. Veterinary Pathology 49, 292–303. doi:10.1177/0300985811402839.

Knauf, S., Dahlmann, F., Batamuzi, E. K., Frischmann, S., Liu, H. (2015b): Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Neglected Tropical Diseases 9, e0003637. doi:10.1371/journal.pntd.0003637.

Knauf, S., Gogarten, J. F., Schuenemann, V. J., Nys, H. M., Düx, A., Strouhal, M., Mikalova, L., Bos, K. I., Armstrong, R., Batamuzi, E. K., Chuma, I. S., Davoust, B., Diatta, G., Fyumagwa, R. D., Kazwala, R. R., Keyyu, J. D., Lejora, I. A. V., Levasseur, A., Liu, H., Mayhew, M. A., Mediannikov, O., Raoult, D., Wittig, R. M., Roos, C., Leendertz, F. H., Smajs, D., Nieselt, K., Krause, J., Calvignac-Spencer, S. (2018): Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. Emerging Microbes & Infections 7, 157. doi:10.1038/s41426-018-0156-4.

Knauf, S., Liu, H., Harper, K. N. (2013): Treponemal infection in nonhuman primates as possible reservoir for human yaws. Emerging Infectious Diseases 19, 2058–2060. doi:10.3201/eid1912.130863.

Knauf, S., Raphael, J., Mitja, O., Lejora, I. A., Chuma, I. S., Batamuzi, E. K., Keyyu, J. D., Fyumagwa, R., Lüert, S., Godornes, C., Liu, H., Schwarz, C., Smajs, D., Grange, P., Zinner, D., Roos, C., Lukehart, S. A. (2016): Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. EBioMedicine 11, 85–90. doi:10.1016/j.ebiom.2016.07.033.

Kobayashi, S. D., Malachowa, N., DeLeo, F. R. (2018): Neutrophils and bacterial immune evasion. Journal of Innate Immunity 10, 432–441. doi:10.1159/000487756.

Kock, R. (2014): Drivers of disease emergence and spread: is wildlife to blame? The Onderstepoort Journal of Veterinary Research 81, E1-4. doi:10.4102/ojvr.v81i2.739.

Koenderman, L., Buurman, W., Daha, M. R. (2014): The innate immune response. Immunology Letters 162, 95–102. doi:10.1016/j.imlet.2014.10.010.

Konstantinidis, K. T., Ramette, A., Tiedje, J. M. (2006): The bacterial species definition in the genomic era. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 361, 1929–1940. doi:10.1098/rstb.2006.1920.

Kopp, G., da Silva, F. M., Fischer, J., Brito, J., Regnaut, S., Roos, C., Zinner, D. (2006): The influence of social systems on patterns of mitochondrial DNA variation in baboons. International Journal of Primatology 35, 210–225. doi:10.1007/s10764-013-9725-5.

Kumar, H., Kawai, T., Akira, S. (2009): Pathogen recognition in the innate immune response. The Biochemical Journal 420, 1–16. doi:10.1042/BJ20090272.

Kumm, H. W. (1935a): The natural infection of *Hippelates pallipes* Loew with the spirochaetes of yaws. Transactions of the Royal Society of Tropical Medicine and Hygiene 29, 265–272. doi:10.1016/s0035-9203(35)90089-x.

Kumm, H. W. (1935b): The digestive mechanism of one of the west indian “eye gnats”: *Hippelates pallipes* Loew. Annals of tropical medicine and parasitology 29, 283–302.

Kumm, H. W., Turner, T. B. (1936): The transmission of yaws from man to rabbits by an insect vector, *Hippelates pallipes* Loew. The American Journal of Tropical Medicine and Hygiene s1-16, 245–271.

Kuris, A. M., Hechinger, R. F., Shaw, J. C., Whitney, K. L., Aguirre-Macedo, L., Boch, C. A., Dobson, A. P., Dunham, E. J., Fredensborg, B. L., Huspeni, T. C., Lorda, J., Mababa, L., Mancini, F. T., Mora, A. B., Pickering, M., Talhouk, N. L., Torchin, M. E., Lafferty, K. D. (2008): Ecosystem energetic implications of parasite and free-living biomass in three estuaries. Nature 454, 515–518. doi:10.1038/nature06970.

Lacks, S., Greenberg, B. (1975): A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA. The Journal of Biological Chemistry 250, 4060–4066.

Lafond, R., Lukehart, S. (2006): Biological basis for syphilis. Clinical Microbiology Reviews 19, 29–49. doi:10.1128/CMR.19.1.29-49.2006.

Lamborn, W. (1936): The experimental transmission to man of *Treponema pertenue* by the fly *Musca sorbens*, WD. The American Journal of Tropical Medicine and Hygiene 15, 235–239.

Langton, T. (2019): The badger cull policy is not evidence based. The Veterinary Record 184, 715. doi:10.1136/vr.l4088.

Laurence, M., Hatzis, C., Brash, D. E. (2014): Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. PLoS ONE 9, e97876. doi:10.1371/journal.pone.0097876.

Laursen, L. (2018): Wild primates threaten efforts to wipe out skin disease. Nature 554, 286. doi:10.1038/d41586-018-01703-0.

Laver, T., Harrison, J., O'Neill, P., Moore, K., Farbos, A., Paszkiewicz, K., Studholme, D. J. (2015): Assessing the performance of the Oxford Nanopore Technologies MinION. Biomolecular Detection and Quantification 3, 1–8. doi:10.1016/j.bdq.2015.02.001.

Lee, J. W., McKibbin, W. J. (2004): Estimating the global economic costs of SARS. Institute of Medicine (US) Forum on Microbial Threats: Learning from SARS: Preparing for the Next Disease Outbreak, National Academies Press (US), 1–376.

Lee, K., Izumiya, H., Iyoda, S., Ohnishi, M., Group, E. (2019): Effective surveillance using multilocus variable number tandem repeat analysis and whole genome sequence analyses in enterohemorrhagic *Escherichia coli* O157. Applied and Environmental Microbiology 85, e00728-19. doi:10.1128/AEM.00728-19.

Lei, B., Olival, K. (2014): Contrasting patterns in mammal-bacteria coevolution: *Bartonella* and *Leptospira* in bats and rodents. PLoS Neglected Tropical Diseases 8, e2738. doi:10.1371/journal.pntd.0002738.

Lenski, R., May, R. (1994): The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. Journal of Theoretical Biology 169, 253–265. doi:10.1006/jtbi.1994.1146.

Leulier, F., Parquet, C., Pili-Floury, S., Ryu, J. H. H., Caroff, M., Lee, W. J. J., Mengin-Lecreulx, D., Lemaitre, B. (2003): The *Drosophila* immune system detects bacteria through specific peptidoglycan recognition. Nature Immunology 4, 478–484. doi:10.1038/ni922.

Levin, S., Pimentel, D. (1981): Selection of intermediate rates of increase in parasite-host

systems. The American Naturalist 117, 308-315.

Locey, K. J., Lennon, J. T. (2016): Scaling laws predict global microbial diversity. Proceedings of the National Academy of Sciences 113, 5970-5975. doi:10.1073/pnas.1521291113.

López-García, A., Pineda-Quiroga, C., Atxaerandio, R., Pérez, A., Hernández, I., García-Rodríguez, A., Gonzalez-Recio, O. (2018): Comparison of Mothur and QIIME for the analysis of rumen microbiota composition based on 16S rRNA amplicon sequences. Frontiers in Microbiology 9, 3010. doi:10.3389/fmicb.2018.03010.

Lu, H., Giordano, F., Ning, Z. (2016): Oxford Nanopore MinION sequencing and genome assembly. Genomics, Proteomics & Bioinformatics 14, 265–279. doi:10.1016/j.gpb.2016.05.004.

Lubinza, C. K. C., Lüert, S., Hallmaier-Wacker, L. K., Ngadaya, E., Chuma, I. S., Kazwala, R. R., Mfinanga, S. G. M., Failing, K., Roos, C.; Knauf, S. (2020): Serosurvey of *Treponema pallidum* infection among children with skin ulcers in the Tarangire-Manyara ecosystem, northern Tanzania. BMC Infectious Diseases. Under Review.

Maiden, M., Bygraves, J., Feil, E., Morelli, G., Russell, J., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., Spratt, B. G. (1998): Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America 95, 3140–3145. doi:10.1073/pnas.95.6.3140.

Maiden, M. C. (2006): Multilocus sequence typing of bacteria. Annual Review of Microbiology 60, 561–588. doi:10.1146/annurev.micro.59.030804.121325.

Maiden, M. C., van Rensburg, M. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A., McCarthy, N. D. (2013): MLST revisited: the gene-by-gene approach to bacterial genomics. Nature Reviews Microbiology 11, 728–736. doi:10.1038/nrmicro3093.

Marchesi, J. R., Ravel, J. (2015): The vocabulary of microbiome research: a proposal. Microbiome 3, 31. doi:10.1186/s40168-015-0094-5.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S., B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S., C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R. Tomasz, A., Vogt, T. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P.,

Begley, R. F., Rothberg, J. M. (2005): Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380. doi:10.1038/nature03959.

Marinus, M., Løbner-Olesen, A. (2014): DNA methylation. EcoSal Plus 1. doi:10.1128/ecosalplus.ESP-0003-2013.

Marks, M. (2016): Yaws: towards the WHO eradication target. Transactions of the Royal Society of Tropical Medicine and Hygiene 110, 319–320. doi:10.1093/trstmh/trw032.

Marks, M., Mitja, O., Vestergaard, L., Pillay, A., Knauf, S., Chen, C., Bassat, Q., Martin, D. L., Fegan, D., Taleo, F., Kool, J., Lukehart, S., Emerson, P. M., Solomon, A. W., Ye, T., Ballard, R. C., Mabey, D. C. W., Asiedu, K. B. (2015): Challenges and key research questions for yaws eradication. Lancet Infectious Diseases 15, 1220–1225. doi:10.1016/S1473-3099(15)00136-X.

Marra, C., Castro, C., Kuller, L., Dukes, A., Centurion-Lara, A., Morton, W., Lukehart, S. A. (1998): Mechanisms of clearance of *Treponema pallidum* from the CSF in a nonhuman primate model. Neurology 51, 957–961.

Marra, C., Sahi, S., Tantalo, L., Godornes, C., Reid, T., Behets, F., Rompalo, A., Klausner, J. D., Yin, Y. P., Mulcahy, F., Golden, M. R., Centurion-Lara, A., Lukehart, S. A. (2010): Enhanced molecular typing of treponema pallidum: geographical distribution of strain types and association with neurosyphilis. The Journal of Infectious Diseases 202, 1380–1388. doi:10.1086/656533.

Maxam, A., Gilbert, W. (1977): A new method for sequencing DNA. Proceedings of the National Academy of Sciences of the United States of America 74, 560–564. doi:10.1073/pnas.74.2.560.

May, R., Anderson, R. (1983): Epidemiology and genetics in the coevolution of parasites and hosts. Proceedings of the Royal Society of London Series B, Biological Sciences 219, 281–313. doi:10.1098/rspb.1983.0075.

McNally, L., Brown, S. (2015): Building the microbiome in health and disease: niche construction and social conflict in bacteria. Philosophical Transactions of the Royal Society B: Biological Science 370. doi:10.1098/rstb.2014.0298.

Metcalf, J. C., Farrar, J., Cutts, F. T., Basta, N. E., Graham, A. L., Lessler, J., Ferguson, N. M., Burke, D. S., Grenfell, B. T. (2016): Use of serological surveys to generate key insights into the changing global landscape of infectious disease. The Lancet 388, 728–730. doi:10.1016/S0140-6736(16)30164-7.

Mitjà, O., Godornes, C., Houinei, W., Kapa, A., Paru, R., Abel, H., Gonzales-Beiras, C., Bieb, S. V., Wangi, J., Barry, A. E., Sanz, S., Basset, Q., Lukehart, S. A. (2018): Re-

emergence of yaws after single mass azithromycin treatment followed by targeted treatment: a longitudinal study. The Lancet 391, 1599–1607. doi:10.1016/S0140-6736(18)30204-6.

Murphy K, Travers P, Walport M. (2008a): Innate immunity. In: K. M. Murphy, P. Travers, M. Walport, Janeway's immunobiology, 7th edition. T&F informa, 39–108.

Murphy, K., Travers, P., Walport, M. (2008b): Appendix I: immunologists' toolbox. In: K. M. Murphy, P. Travers, M. Walport, Janeway's immunobiology, 7th edition. Garland Science, 735–782.

Mysara, M., Vandamme, P., Props, R., Kerckhof, F. M. M., Leys, N., Boon, N., Raes, J., Monsieurs, P. (2017): Reconciliation between operational taxonomic units and species boundaries. FEMS Microbiology Ecology 93. doi:10.1093/femsec/fix029.

Mysterud, A., Rolandsen, C. M. (2019): Fencing for wildlife disease control. Journal of Applied Ecology 56, 519–525. doi:10.1111/1365-2664.13301.

Neel, C., Etienne, L., Li, Y., Takehisa, J., Rudicell, R., Bass, I., Moundindo, J., Mebenga, A., Esteban, A., Heuverswyn, F. V., Liegeois, F., Kranzusch, P. J., Walsh, P. D., Sanz, C. M., Morgan, D. B., Ndjango, J. B. N., Plantier, J. C., Locatelli, S., Gonder, M. K., Leendertz, F. H., Boesch, C., Todd, A., Delaporte, E., Mpoudi-Ngole, E., Hahn, B. H., Peeters, M. (2010): Molecular epidemiology of simian immunodeficiency virus infection in wild-living gorillas. Journal of Virology 84, 1464–1476. doi:10.1128/JVI.02129-09.

Nguyen, N., Fashing, P., Boyd, D., Barry, T., Burke, R., Goodale, C., Jones, S. C. Z., Kerby, J. T., Kellogg, B. S., Lee, L. M., Miller, C. M., Nurmi, N. O., Ramsay, M. S., Reynolds, J. D., Stewart, K. M., Turner, T. J., Venkataraman, V. V., Knauf, Y., Roos, C., Knauf, S. (2015): Fitness impacts of tapeworm parasitism on wild gelada monkeys at Guassa, Ethiopia. American Journal of Primatology 77, 579–594. doi:10.1002/ajp.22379.

Nichols, E., Gómez, A. (2014): Dung beetles and fecal helminth transmission: patterns, mechanisms and questions. Parasitology 141, 614–623. doi:10.1017/S0031182013002011.

Nichols, H. J. (1910): Experimental yaws in the monkey and rabbit. The Journal of Experimental Medicine 12, 616–622.

Nicholson, W., Munakata, N., Horneck, G., Melosh, H., Setlow, P. (2000): Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. Microbiology and Molecular Biology Reviews 64, 548–572. doi:10.1128/mmbr.64.3.548-572.2000.

Nunn, C., Scully, E., Kutsukake, N., Ostner, J., Schülke, O., Thrall, P. (2014): Mating competition, promiscuity, and life history traits as predictors of sexually transmitted disease risk in primates. International Journal of Primatology 35, 764-786. doi:10.1007/s10764-014-9781-5.

Nyhus, P. J. (2015): Human–wildlife conflict and coexistence. Annual Review of Environment and Resources 41, 1–29. doi:10.1146/annurev-environ-110615-085634.

Office International des Epizooties (2019): Principles and methods of validation of diagnostic assays for infectious diseases. Manual of diagnostic tests for aquatic animals, World Organisation for Animal Health (OiE). ISBN: 978-92-9044-887-7. https://www.oie.int/standard-setting/aquatic-manual/access-online/.

Overmann, J. (2013): Principles of enrichment, isolation, cultivation, and preservation of prokaryotes. In: E. Rosenberg, E. F. DeLong, E. Stackebrandt, S. Lory, F. Thompson, The Prokaryotes. Springer, New York, NY. doi.org/10.1007/0-387-30741-9_5.

Overmann, J., Scholz, A. (2017): Microbiological research under the nagoya protocol: facts and fiction. Trends in Microbiology 25, 85–88. doi:10.1016/j.tim.2016.11.001.

Paciência, F. M. D., Rushmore, J., Chuma, I. S., Lipende, I. F., Caillaud, D., Knauf, S., Zinner, D. (2019): Mating avoidance in female olive baboons (*Papio anubis*) infected by *Treponema pallidum*. Science Advances. 5, eaaw92724. doi: 10.1126/scladv.aaw9724.

Pavón, A. B., Maiden, M. C. (2009): Multilocus sequence typing. In: Molecular epidemiology of microorganisms. Methods in Molecular Biology (Methods and Protocols) 551, 129–140. Humana Press, Totowa, NJ. doi:10.1007/978-1-60327-999-4_11.

Pérez-Losada, M., Arenas, M., Castro-Nallar, E. (2017): Multilocus sequence typing of pathogens: methods, analyses, and applications. In: M. Tibayrenc, Genetics and evolution of infectious diseases. Elsevier, 383–404.

Pérez-Losada, M., Cabezas, P., Castro-Nallar, E., Crandall, K. A. (2013): Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 16, 38–53. doi:10.1016/j.meegid.2013.01.009.

Pillai, S. D. (1997): Rapid molecular detection of microbial pathogens: breakthroughs and challenges. In: O. R. Kaaden, C. P. Czerny, W. Eichhorn, Viral zoonoses and food of animal origin. Springer, 67–82. doi:10.1007/978-3-7091-6534-8_7.

Pillay, A., Liu, H., Chen, C., Holloway, B., Sturm, A., Steiner, B., Morse, S. A. (1998): Molecular subtyping of *Treponema pallidum* subspecies *pallidum*. Sexually Transmitted Diseases 25, 408–414.

Piwowar, H. A. (2011): Who shares? Who doesn't? Factors associated with openly archiving raw research data. PLoS ONE 6, e18657. doi:10.1371/journal.pone.0018657.

Poirotte, C., Kappeler, P., Ngoubangoye, B., Bourgeois, S., Moussodji, M., Charpentier, M. (2016): Morbid attraction to leopard urine in *Toxoplasma*-infected chimpanzees. Current Biology 26, R98-9. doi:10.1016/j.cub.2015.12.020.

Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., Barrio-Amoros, C. L., Salazar-Valenzuela, D., Prost, S. (2018): Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. GigaScience 7, giy033. doi:10.1093/gigascience/giy033.

Porter, R., Norman, R., Gilbert, L. (2013): An alternative to killing? Treatment of reservoir hosts to control a vector and pathogen in a susceptible species. Parasitology 140, 247–257. doi:10.1017/S0031182012001400.

Prentice, J. C., Fox, N. J., Hutchings, M. R., White, P. C., Davidson, R. S., Marion, G. (2019): When to kill a cull: factors affecting the success of culling wildlife for disease control. Journal of the Royal Society, Interface 16, 20180901. doi:10.1098/rsif.2018.0901.

Radhouani, H., Silva, N., Poeta, P., Torres, C., Correia, S., Igrejas, G. (2014): Potential impact of antimicrobial resistance in wildlife, environment and human health. Frontiers in Microbiology 5, 23. doi:10.3389/fmicb.2014.00023.

Ratel, D., Ravanat, J. L. L., Charles, M. P. P., Platet, N., Breuillaud, L., Lunardi, J., Berger, F., Wion, D. (2006): Undetectable levels of N6-methyl adenine in mouse DNA: Cloning and analysis of PRED28, a gene coding for a putative mammalian DNA adenine methyltransferase. FEBS Letters 580, 3179–3184. doi:10.1016/j.febslet.2006.04.074.

Reed, P. E., Mulangu, S., Cameron, K. N., Ondzie, A. U., Joly, D., Bermejo, M., Rouquet, P., Fabozzi, G., Bailey, M., Shen, Z., Keele, B. F., Hahn, B., Karesh, W. B., Sullivan, N. J. (2014): A new approach for monitoring ebolavirus in wild great apes. PLoS Neglected Tropical Diseases 8, e3143. doi:10.1371/journal.pntd.0003143.

Richard, M., Knauf, S., Lawrence, P., Mather, A. E., Munster, V. J., Muller, M. A., Smith, D., Kuiken, T. (2017): Factors determining human-to-human transmissibility of zoonotic pathogens via contact. Current Opinion in Virology 22, 7–12. doi:10.1016/j.coviro.2016.11.004.

Roche, B., Dobson, A., Guegan, J., Rohani, P. (2012): Linking community and disease ecology: the impact of biodiversity on pathogen transmission. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 2807–2813. doi:10.1098/rstb.2011.0364.

Roche, B., Guégan, J. F. (2011): Ecosystem dynamics, biological diversity and emerging infectious diseases. Comptes Rendus Biologies 334, 385–392. doi:10.1016/j.crvi.2011.02.008.

Roeder, P., Mariner, J., Kock, R. (2013): Rinderpest: the veterinary perspective on eradication. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 20120139. doi:10.1098/rstb.2012.0139.

Roos, C., Chuma, I. S., Collins, A. D., Knauf, S., Zinner, D. (2018): Complete mitochondrial genome of an olive baboon (*Papio anubis*) from Gombe National Park, Tanzania. Mitochondrial DNA Part B 3, 177–178. doi:10.1080/23802359.2018.1437813.

Roos, C., Zinner, D., Kubatko, L., Schwarz, C., Yang, M., Meyer, D., Nash, S. D., Xing, J., Batzer, M. A., Brameier, M., Leendertz, F. H., Ziegler, T., Perwitasari-Farajallah, D., Nadler, T., Walter, L., Osterholz, M. (2011): Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. BMC Evolutionary Biology 11, 77. doi:10.1186/1471-2148-11-77.

Rosselló-Mora, R., Amann, R. (2001): The species concept for prokaryotes. FEMS Microbiology Reviews 25, 39–67. doi:10.1111/j.1574-6976.2001.tb00571.x.

Sachse, S., Straube, E., Lehmann, M., Bauer, M., Russwurm, S., Schmidt, K. H. H. (2009): Truncated human cytidylate-phosphate-deoxyguanylate-binding protein for improved nucleic acid amplification technique-based detection of bacterial species in human samples. Journal of Clinical Microbiology 47, 1050–1057. doi:10.1128/JCM.02242-08.

Sánchez-Romero, M. A., Cota, I., Casadesús, J. (2015): DNA methylation in bacteria: from the methyl group to the methylome. Current Opinion in Microbiology 25, 9–16. doi:10.1016/j.mib.2015.03.004.

Sanger, F., Air, G., Barrell, B., Brown, N., Coulson, A., Fiddes, C., Hutchison III, C. A., Slocombe, P. M., Smith, M. (1977a): Nucleotide sequence of bacteriophage phi X174 DNA. Nature 265, 687–695. doi:10.1038/265687a0.

Sanger, F., Nicklen, S., Coulson, A. (1977b): DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74, 5463–5467.

Satchell, G., Harrison, R. (1953): Experimental observations on the possibility of transmission of yaws by wound-feeding Diptera, in Western Samoa. Transactions of the Royal Society of Tropical Medicine and Hygiene 47, 148–153.

Sawitzky, D. (1997): Transmission, species specificity, and pathogenicity of Aujeszky's disease virus. Archives of Virology Supplementum 13, 201–206.

Schloss, P., Westcott, S., Ryabin, T., Hall, J., Hartmann, M., Hollister, E., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallingen, G. G., Van Horn, D. J., Weber, C. F. (2009): Introducing mothur: open-source, platform-independent,

community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75, 7537–7541. doi:10.1128/AEM.01541-09.

Schloss, P. D. (2016): Application of a database-independent approach to assess the quality of operational taxonomic unit picking methods. MSystems 1, e00027-16. doi:10.1128/mSystems.00027-16.

Sepetjian, M., Guerraz, F., Salussola, D., Thivolet, J., Monier, J. (1969): Contribution to the study of the treponeme isolated from monkeys by A. Fribourg-Blanc. Bulletin of the World Health Organization 40, 141–151.

Shakespeare, M. (2009): Zoonoses. 2nd ed. Pharmaceutical Press.

Shendure, J., Ji, H. (2008): Next-generation DNA sequencing. Nature Biotechnology 26, 1135–1145. doi:10.1038/nbt1486.

Shimoyama, T., Sawada, Y., Sawada, N., Chinda, D., Fukuda, S. (2017): Accuracy of a stick-type kit and enzyme-linked immunosorbent assay in detecting *Helicobacter pylori* antibodies in urine of people living in the Japan sea region of Northern Japan. Japanese Journal of Infectious Diseases 70, 207–209. doi:10.7883/yoken.JJID.2015.642.

Siembieda, J., Kock, R., McCracken, T., Newman, S. (2011): The role of wildlife in transboundary animal diseases. Animal Health Research Reviews 12, 95–111. doi:10.1017/S1466252311000041.

Sleeman, J. M., DeLiberto, T., Ngyuen, N. (2017): Optimization of human, animal, and environmental health by using the One Health approach. Journal of Veterinary Science 18 (S1), 263–268. doi:10.4142/JVS.2017.18.S1.263.

Smith, J., David, N., Indgin, S., Israel, C., Levine, B., Justice, J., McCrary III, J. A., Medina, R., Paez, P., Santana, E., Sarkar, M., Schatz, N. J., Spitzer, M. L., Spitzer, W. O., Walter, E. K. (1971): Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. The British Journal of Venereal Diseases 47, 226–251. doi:10.1136/sti.47.4.226.

Smith, J. L. (2003): The role of gastric acid in preventing foodborne disease and how bacteria overcome acid conditions. Journal of Food Protection 66, 1292–1303.

Smith, K., Goldberg, M., Rosenthal, S., Carlson, L., Chen, J., Chen, C., Ramachandran, S. (2014): Global rise in human infectious disease outbreaks. Journal of The Royal Society Interface 11, 20140950–20140950. doi:10.1098/rsif.2014.0950.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H., Hood, L. E. (1986): Fluorescence detection in automated DNA sequence analysis. Nature 321, 674–679. doi:10.1038/321674a0.

Smolinski, M., Hamburg, M., Lederberg, J. (2003): Microbial threats to health: emergence, detection, and response. National Academic Press.

Staley, J. T. (2006): The bacterial species dilemma and the genomic-phylogenetic species concept. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 361, 1899–1909. doi:10.1098/rstb.2006.1914.

Stallknecht, D. (2007): Impediments to wildlife disease surveillance, research, and diagnostics. Current Topics in Microbiology and Immunology 315. 445–461.

Strouhal, M., Mikalová, L., Haviernik, J., Knauf, S., Bruisten, S., Noordhoek, G. T., Oppelt, J., Cejkova, D., Smajs, D. (2018): Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Indonesia: Modular structure of several treponemal genes. PLoS Neglected Tropical Diseases 12, e0006867. doi:10.1371/journal.pntd.0006867.

Sullivan, M. B. (2015): Viromes, Not Gene Markers, for Studying Double-Stranded DNA Virus Communities. Journal of Virology 89, 2459–2461. doi:10.1128/JVI.03289-14.

Swart, A., Bekker, D. L., Maas, M., de Vries, A., Pijnacker, R., Reusken, C. B., Van der Giessen, J. W. B. (2017): Modelling human Puumala hantavirus infection in relation to bank vole abundance and masting intensity in the Netherlands. Infection Ecology & Epidemiology 7, 1287986. doi:10.1080/20008686.2017.1287986.

Tanner, E., White, A., Acevedo, P., Balseiro, A., Marcos, J., Gortázar, C. (2019): Wolves contribute to disease control in a multi-host system. Scientific Reports 9, 7940. doi:10.1038/s41598-019-44148-9.

Taylor, L., Latham, S., Woolhouse, M. (2001): Risk factors for human disease emergence. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 356, 983–989. doi:10.1098/rstb.2001.0888.

Tibayrenc, M. (2009): Multilocus enzyme electrophoresis for parasites and other pathogens. Methods in Molecular Biology (Clifton, NJ) 551, 13–25. doi:10.1007/978-1-60327-999-4_2.

Tripathi, R., Sharma, P., Chakraborty, P., Varadwaj, P. (2016): Next-generation sequencing revolution through big data analytics. Frontiers in Life Science 9, 119–149. doi:10.1080/21553769.2016.1178180.

Turner, T., Hollander, D. (1957): Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. Monograph Series World Health Organization, 3–266.

Turner, W., Kausrud, K., Krishnappa, Y., Cromsigt, J., Ganz, H., Mapaure, I., Cloete, C. C., Havarua, Z., Küsters, M., Getz, W. M., Stenseth, N. C. (2014): Fatal attraction: vegetation responses to nutrient inputs attract herbivores to infectious anthrax carcass sites. Proceedings of the Royal Society B: Biological Science 281. doi:10.1098/rspb.2014.1785.

Viana, M., Mancy, R., Biek, R., Cleaveland, S., Cross, P., Lloyd-Smith, J., Haydon, D. T. (2014): Assembling evidence for identifying reservoirs of infection. Trends in Ecology and Evolution 29, 270–279. doi:10.1016/j.tree.2014.03.002.

Vicente, J., Apollonio, M., Blanco-Aguiar, J. A., Borowik, T., Brivio, F., Casaer, J., Croft, S., Ericsson, G., Ferroglio, E., Gavier-Widen, D., Gortazar, C., Jansen, P. A., Keuling, Ol., Kowalczyk, R., Petrovic, K., Plhal, R., Podgorski, T., Sange, M., Scandura, M., Schmidt, K., Smith, G. C., Soriguer, R., Thulke, H. H., Zanet, S., Acevedo, P. (2019): Science-based wildlife disease response. Science 364, 943-944. doi:10.1126/science.aax4310.

Vinkers, C. H., Tijdink, J. K., Otte, W. M. (2015): Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. British Medical Journal (Clinical Research Edition) 351, h6467. doi:10.1136/bmj.h6467.

Volden, R., Palmer, T., Byrne, A., Cole, C., Schmitz, R. J., Green, R. E., Vollmers, C. (2018): Improving nanopore read accuracy with the R2D2 method enables the sequencing of highly multiplexed full-length single-cell cDNA. Proceedings of the National Academy of Sciences 115 (39), 9726–9731. doi: 10.1073/pnas.1806447115.

Voo, K., Carlone, D., Jacobsen, B., Flodin, A., Skalnik, D. (2000): Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1. Molecular and Cellular Biology 20, 2108–2121. doi:10.1128/mcb.20.6.2108-2121.2000.

Vovis, G., Lacks, S. (1977): Complementary action of restriction enzymes endo R-DpnI and Endo R-DpnII on bacteriophage f1 DNA. Journal of Molecular Biology 115, 525–538.

Wallis, J., Lee, R. D. (1999): Primate conservation: the prevention of disease transmission. International Journal of Primatology 20, 803–826.

Walther, B. A., Ewald, P. W. (2004): Pathogen survival in the external environment and the evolution of virulence. Biological Reviews of the Cambridge Philosophical Society 79, 849–869.

Watson, J., Crick, F. (1953): The structure of DNA. Cold Spring Harbor Symposia on Quantitative Biology 18, 123–131. doi:10.1101/SQB.1953.018.01.020.

Weinersmith, K., Faulkes, Z. (2014): Parasitic manipulation of hosts' phenotype, or how to make a zombie--an introduction to the symposium. Integrative and Comparative Biology 54,

93–100. doi:10.1093/icb/icu028.

White, P., Lewis, A., Harris, S. (1997): Fertility control as a means of controlling bovine tuberculosis in badger (Meles meles) populations in south-west England: predictions from a spatial stochastic simulation model. Proceedings Biological Sciences 264, 1737–1747. doi:10.1098/rspb.1997.0241.

van de Wijgert, J. (2017): The vaginal microbiome and sexually transmitted infections are interlinked: Consequences for treatment and prevention. PLoS Medicine 14, e1002478. doi:10.1371/journal.pmed.1002478.

Willcox, R., Guthe, T. (1966a): Survival of *T. pallidum* outside the body: I. General observations. Bulletin of the World Health Organization 35, 78–85.

Willcox, R., Guthe, T. (1966b): Survival of *T. pallidum* outside the body: II. Factors favouring survival. Bulletin of the World Health Organization 35, 86–90.

Wobeser, G. (2002): Disease management strategies for wildlife. Revue Scientifique et Technique (International Office of Epizootics) 21, 159–178.

Wolfe, N., Dunavan, C., Diamond, J. (2007): Origins of major human infectious diseases. Nature 447, 279–283. doi:10.1038/nature05775.

Woodroffe, R. (2001): Assessing the risks of intervention: immobilization, radio-collaring and vaccination of African wild dogs. Oryx 35, 234–244. doi:10.1046/j.1365-3008.2001.00186.x.

Woolhouse, M. E., Gowtage-Sequeria, S. (2005): Host range and emerging and reemerging pathogens. Emerging Infectious Diseases 11, 1842–1847. doi:10.3201/eid1112.050997.

World Health Organization (2012): Eradication of yaws--the Morges strategy. Relevé Épidémiologique Hebdomadaire Sect D'hygiène Du Secrétariat De La Société Des Nations Wkly Epidemiological Rec Heal Sect Secr Leag Nations 87, 189–194.

World Health Organization (2018): Eradication of yaws: procedures for verification and certification of interruption of transmission. WHO/HTM/NTD/IDM/2018.02. ISBN: 9789241512909.

Xue, Z., Kable, M. E., Marco, M. L. (2018): Impact of DNA sequencing and analysis methods on 16S rRNA gene bacterial community analysis of dairy products. MSphere 3, e00410-18. doi:10.1128/mSphere.00410-18.

Yeo, J., Killington, R. A., Watson, D. H., Powell, K. L. (1981): Studies on cross-reactive antigens in the herpesviruses. Virology 108, 256–266. doi:10.1016/0042-6822(81)90434-7.

Zhang, J. M., An, J. (2007): Cytokines, inflammation, and pain. International Anesthesiology Clinics 45, 27–37. doi:10.1097/AIA.0b013e318034194e.

Zhang, W., Jayarao, B. M., Knabel, S. J. (2004): Multi-virulence-locus sequence typing of *Listeria monocytogenes*. Applied and Environmental Microbiology 70, 913–920. doi:10.1128/aem.70.2.913-920.2004.

Zimmerman, D. M., Hardgrove, E. H., von Fricken, M. E., Kamau, J., Chai, D., Mutura, S., Kivali, V., Hussein, F., Ambala, P., Surmat, A., Maina, J. G., Knauf, S. (2019): Endemicity of yaws shown by *Treponema pallidum* antibodies in nonhuman primates, Kenya. Emerging Infectious Diseases 25 (11), 2147–2149. doi:10.3201/eid2511.190716.

Zinner, D., Arnold, M., Roos, C. (2009a): Is the new primate genus *Rungwecebus* a baboon? PLoS One 4, e4859. doi:10.1371/journal.pone.0004859.

Zinner, D., Chuma, I. S., Knauf, S., Roos, C. (2018): Inverted intergeneric introgression between critically endangered kipunjis and yellow baboons in two disjunct populations. Biology Letters 14. doi:10.1098/rsbl.2017.0729.

Zinner, D., Groeneveld, L., Keller, C., Roos, C. (2009b): Mitochondrial phylogeography of baboons (*Papio* spp.): indication for introgressive hybridization? British Medical Journal Evolutionary Biology 9, 83. doi:10.1186/1471-2148-9-83.

Zinner, D., Keller, C., Nyahongo, J. W., Butynski, T. M., de Jong, Y. A., Pozzi, L., Knauf, S., Liedigk, R., Roos, C. (2015): Distribution of mitochondrial clades and morphotypes of baboons *Papio* spp. (Primates: Cercopithecidae) in Eastern Africa. Journal of East African Natural History 105, 143–168. doi:10.2982/028.104.0111.

Zinner, D., Wertheimer, J., Liedigk, R., Groeneveld, L., Roos, C. (2012): Baboon phylogeny as inferred from complete mitochondrial genomes. American Journal of Physical Anthropology 150, 133–140. doi:10.1002/ajpa.22185.

Zobaníková, M., Strouhal, M., Mikalová, L., Čejkova, D., Ambrožová, L., Pospíšilová, P., Fulton, L. L., Chen, L., Sodergren, E. Weinstock, G. M., Šmajs, D. (2013): Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian isolate is highly similar to the yaws subspecies. PLoS Neglected Tropical Diseases 7, e2172. doi:10.1371/journal.pntd.0002172.

12. Authorship Contributions

Names of my trainees are underlined; *contributed equally

1. Richard, M., **Knauf, S.**, Lawrenz, P., Mather, A., Munster, V., Müller, M., Smith, D., Kuiken, T. (2016): Factors determining human-to-human transmissibility of zoonotic pathogens via contact. Opinion in Virology. 22, 7-12. doi: 10.1016/j.coviro.2016.11.004. PubMed PMID: 27907884.
Publication Type: Review article
Contribution: Participated in the conceptualization of the paper and co-wrote the manuscript.
Journal Impact Factor: 5.4

2. Hallmaier-Wacker, L.K., Munster, V.J., **Knauf, S.** (2017): Disease reservoirs: From conceptual frameworks to applicable criteria. Emerging Microbes & Infections. 6 (9), e79. doi: 10.1038/emi.2017.65. PubMed PMID: 28874791.
Publication Type: Review article
Contribution: Conceptualized and co-wrote, read, reviewed and approved the manuscript.
Journal Impact Factor: 6.2

3. Hallmaier-Wacker, L.K., Lüert, S., Gronow, S., Spröer, C., Overmann, J., Buller, N., Vaughan-Higgins, R.J., **Knauf, S.** (2019): A metataxonomic tool investigate the diversity of *Treponema*. Frontiers in Microbiology 10, 2094. doi: 10.3389/fmicb.2019.02094
Publication Type: Original article
Contribution: Conceived and designed the study, performed experiments in the laboratory, analyzed data, wrote, read, reviewed and approved the manuscript.
Journal Impact Factor: 4.1

4. **Knauf, S.**, Dahlmann, F., Batamuzi, E.K., Frischmann, S., Hsi, L. (2015): Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. PLoS Neglected Tropical Diseases. 9 (3), e0003637. doi: 10.1371/journal.pntd.0003637. PubMed PMID: 25803295; PubMed Central PMCID: PMC4372418.
Publication Type: Original article
Contribution: Conceived and designed the experiments, performed the experiments including field immobilizations, analysed data, contributed reagents and materials as well as wrote, read, revised and approved the manuscript.
Journal Impact Factor: 4.5

5. **Knauf, S.**, Lüert, S., Šmajs, D., Strouhal, M., Chuma, I. S., Frischmann, S., Bakheit, M. (2018): Gene target selection for Loop-mediated isothermal amplification for rapid discrimination of *Treponema pallidum* subspecies. PLoS Neglected Tropical Diseases. PLoS Neglected Tropical Diseases. 12, e0006396. doi: 10.1371/journal.pntd.0006396.
Publication Type: Original article
Contribution: Conceptualized the study, performed experiments, curated and analysed data, administrated the project, contributed resources, supervised the study, performed the validation process, wrote, read, revised and approved the manuscript.
Journal Impact Factor: 4.5

6. Bechere, L., **Knauf, S**., Marks, M., Lueert, S., Frischmann, S., Borst, N., von Stetten, F., Bieb, S., Adu-Sarkodie, Y., Asiedu, K., Mitja, O., Bakheit, M. (2019) Validation of a multiplex mediator displacement LAMP for the detection of *Treponema pallidum* and *Haemophilus ducreyi* using clinical samples. Emerging Infectious Diseases. 26(2), 282-288. doi: 10.3201/eid2602.190505. PubMed PMID: 31961303.
Publication Type: Original article
Contribution: Conceptualized the study, performed qPCR experiments, curated and analysed data, co-wrote, read, reviewed and approved the manuscript.
Journal Impact Factor: 7.4

7. Hallmaier-Wacker, L.K., Lueert, S., Roos, C., **Knauf, S.** (2018): The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. Scientific Reports. 8, 6292. doi: 10.1038/s41598-018-24573-y.
Publication Type: Original article
Contribution: Designed the study and analysed data, supervised the project, co-wrote, read, reviewed and approved the manuscript.
Journal Impact Factor: 4.5

8. Šmajs, D., Strouhal, M., **Knauf, S.** (2018): Genetics of human and animal uncultivable treponemal pathogens. Journal of Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases. 61, 92-107. doi: 10.1016/j.meegid.2018.03.015. PubMed PMID: 29578082.
Publication Type: Review article
Contribution: Conceptualized the manuscript, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 2.6

9. Marks, M., Mitjà, O., Vestergaard, L.S., Pillay, A., **Knauf, S.**, Chen, C.Y., Bassat, Q., Martin, D.L., Fegan, D., Taleo, F., Kool, J., Lukehart, S., Emerson, P. M., Solomon, A.W., Ye, T., Ballard, R. C., Mabey, D. C. W., Asiedu, K.B. (2015): Challenges and key research questions for yaws eradication. Lancet Infectious Diseases. 15 (10), 1220-1225.
Publication Type: Review article
Contribution: Reviewed the draft manuscript, provided comments, critical review and helped in the revision of the manuscript.
Journal Impact Factor: 27.5

10. **Knauf, S.**, Liu, H., Harper, K. N. (2013): Treponemal infection in nonhuman primates as potential reservoir for human yaws. Emerging Infectious Diseases, 19 (12), 2058-2060.
Publication Type: Review article
Contribution: Conceptualized the review, analysed data, wrote, read, revised and approved the manuscript.
Journal Impact Factor: 7.4

11. **Knauf, S.**, Barnett, U., Maciej, P., Klapproth, M., Ndao, I., Frischmann, S., Fischer, J., Zinner, D., Liu, H. (2015): High prevalence of antibodies against the bacterium *Treponema pallidum* in Senegalese Guinea baboons (*Papio papio*). PLoS One. 10 (11), e0143100. doi: 10.1371/journal.pone.0143100. PubMed PMID: 26588087.
Publication Type: Original article
Contribution: Conceived and designed the experiments, performed the experiments, analysed the data, contributed reagents and materials, wrote, read, revised and approved the manuscript.
Journal Impact Factor: 2.8

12. Zimmerman D. M., Hardgrove, E. H., von Fricken, M. E., Kamau, J., Chai, D., Mutura, S., Kivali, V., Hussein, F., Ambala, P., Surmat, A., Maina, J. G., **Knauf, S.** (2019): Endemicity of yaws shown by *Treponema pallidum* antibodies in nonhuman primates, Kenya. Emerging Infectious Diseases. 25 (11), 2147-2149. doi:10.3201/eid2511.190716.
Publication Type: Original article
Contribution: Conceived and co-designed the experiments, analysed the data, contributed reagents and materials, co-wrote, read, reviewed and approved the manuscript.
Journal Impact Factor: 7.4

13. Chuma, I. S., Batamuzi E. K., Collins, D. A., Fyumagwa, R. D., Hallmaier-Wacker, L. K., Kazwala, R. R., Keyyu, J. D., Lejora, I. A., Lipende, I. F., Lüert, S., Paciência, F. D. M., Piel, A., Stewart, F. A., Zinner, D., Roos, C., **Knauf S.** (2018): Widespread *Treponema pallidum* infection in nonhuman primates, Tanzania. Emerging Infectious Diseases. 24, 1002-1009 doi: 10.3201/eid2406.180037.
Publication Type: Original article
Contribution: Conceptualized the study, supervised field work, performed and supervised laboratory experiments, curated and analysed data, administrated the project, contributed resources, performed the validation process, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 7.4

14. **Knauf, S.***, Gogarten, J. F.*, Schuenemann, V.*, de Nys, H.M.*, Düx, A., Strouhal, M., Mikalová, L., Bos, K. I., Armstrong, R., Batamuzi, E. K., Chuma, I.S., Davoust, B., Diatta, G., Fyumagwa, R., Kazwala, R. R., Keyyu, J. D., Lejora, I. A. V., Levasseur, A., Liu, H., Mayhew, M. A., Mediannikov, O., Raoult, D., Wittig, R. M., Roos, C., Leendertz, F.H., Šmajs, D.*, Nieselt, K.*, Krause, J.*, Calvignac-Spencer, S.* (2018): Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. Emerging Microbes & Infections. 7, 175. doi:10.1038/s41426-018-0156-4.
Publication Type: Original article
Contribution: Collection and processing of samples from East Africa: conceptualized the study, performed and supervised laboratory experiments, curated and analysed data, administrated the project, contributed resources, performed the validation process, wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 6.2

15. Strouhal, M., Mikalová, L., Haviernik, J., **Knauf, S.**, Bruisten, S., Noordhoek, G. T., Oppelt, J., Čejkova, D., Šmajs, D. (2018): Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Indonesia: Modular structure of several treponemal genes. PLoS Neglected Tropical Diseases. 12, e0006867. doi: 10.1371/journal.pntd.0006867.
Publication Type: Original article
Contribution: Analysed the data, contributed resources and materials, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 4.5

16. Klegarth, A. R., Ezeonwu, C. A., Lee, B. P. Y. H., Aggimarangsee, N., Chalise, M., Cortes, J., Feeroz, M., Molini, B. J., Godornes, B. C., Marks, M., Schillaci, M., Engel, G., **Knauf, S.**, Lukehart, S. A., Jones-Engel, L. (2017): Survey of Treponemal Infections in free-ranging and captive macaques, 1999-2012. Emerging Infectious Diseases; 23(5), 816-819. doi: 10.3201/eid2305.161838. PubMed PMID: 28418297; PubMed Central PMCID: PMCPMC5403046.
Publication Type: Original article
Contribution: Conceptualized and designed the experiments, analysed the data, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 7.4

17. Chuma, I. S.*, Roos, C.*, Atickem, A., Bohm, T., Collins, D. A.; Grillová, L., Hallmaier-Wacker, L. K., Kazwala, R. R., Keyyu, J. D., Lüert, S., Maloueki, U., Oppelt, J., Petrželková, K. J., Piel, A., Stewart, F. A., Šmajs, D., **Knauf, S**. (2019): Strain diversity of *Treponema pallidum* subsp. *pertenue* suggests rare interspecies transmission in African nonhuman primates. Scientific Reports 9, 14243. doi:10.1038/s41598-019-50779-9.
Publication Type: Original article
Contribution: Conceived and designed the study, designed the MLST system, performed and supervised experiments, analysed data and supervised data curation, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 4.5

18. **Knauf, S.**, Raphael, J., Mitjá, O., Lejora, I. A. V., Chuma, I. S., Batamuzi, E. K., Keyyu, J. D., Fyumagwa, R., Lüert, S., Godornes, C., Liu, H., Schwarz, C., Grange, P., Šmajs, D., Roos, C.*, Lukehart, S. A.* (2016): Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. EBioMedicine; 11: 85-90. doi:10.1016/j.e.biom.2016.07.033.
Publication Type: Original article
Contribution: Conceived and designed the study, supervised fieldwork, performed laboratory experiments, analysed and curated data, wrote, read, revised and approved the manuscript.
Journal Impact Factor: 6.7

19. Houinei, W., Godornes, C., Kapa, A., **Knauf, S.**, Mooring, E. Q., González-Beiras C., Watup, R., Paru, R., Advent P., Bieb, S., Sanz, S., Bassat, Q., Spinola, S. M., Lukehart, S., Mitjá O. (2017): *Haemophilus ducreyi* DNA is detecable on the skin of asymptomatic children, flies and fomites in villages of Papua New Guinea. PLoS Neglected Tropical Diseases. 11(5),

e0004958. doi: 10.1371/journal.pntd.0004958. PubMed PMID: 28489855; PubMed Central PMCID: PMCPMC5425006.
Publication Type: Original article
Contribution: Co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 4.5

20. Paciência, F. M. D., Rushmore, J., Chuma, I. S., Lipende, I. F., Caillaud, D., **Knauf, S***, Zinner, D.* (2019): Mating Avoidance in Female Olive Baboon (*Papio anubis*) Infected by *Treponema pallidum.* Science Advances. 5(12), eaaw92724. doi: 10.1126/scladv.aaw9724. PubMed PMID. 31840059.
Publication Type: Original Article
Contribution: Conceived and co-designed the study, performed fieldwork including collaring and sampling of baboons, co-supervised fieldwork, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 12.8

21. Roos, C., Chuma, I. S., Collins, D. A., **Knauf, S.**, Zinner, D. (2018): Complete mitochondrial genome of an olive baboon (*Papio anubis*) from Gombe National Park, Tanzania. Mitochondrial DNA Part B. 3, 177-178. doi: 10.1080/23802359.2018.1437813.
Publication Type: Original article
Contribution: Conceived and co-designed experiment, contributed reagents and materials, co-wrote, read, reviewed and revised the manuscript.
Journal Impact Factor: 0.5

22. Zinner, D., Keller, C., Nyahongo, J. W., Butynski, T.M., de Jong, Y. A., Pozzi, L., **Knauf, S.**, Liedigk, R., Roos, C. (2015): Distribution of mitochondrial clades and morphotypes of baboons *Papio* spp. (Primates: Cercopithecidae) in eastern Africa. Journal of East African Natural History. 104 (1&2): 143-168. doi: 10.2982/028.104.0111.
Publication Type: Original article
Contribution: Conceived and co-designed experiment, contributed reagents and materials, performed laboratory experiments, co-wrote, read, reviewed and revised the manuscript.
Journal Impact Factor: 0.3 (RG Journal Impact)

23. Zinner, D., Chuma, I. S., **Knauf, S.**, Roos, C. (2018): Inverted intergeneric introgression between critically endangered kipunjis and yellow baboons in two disjunct populations. Biology Letters. 14: 20170729. doi: 10.1098/rsbl.2017.0729.
Publication Type: Original article
Contribution: Performed laboratory work, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 3.3

24. Lubinza, C. K. C., Lüert, S., Hallmaier-Wacker, L. K., Ngadaya, E., Chuma, I. S., Kazwala, R. R., Mfinanga, S. G. M., Failing, K., Roos, C.; **Knauf, S.** (2020): Serosurvey of *Treponema*

pallidum infection among children with skin ulcers in the Tarangire-Manyara ecosystem, northern Tanzania. BMC Infectious Diseases. 20: 392. doi: 10.1186/s12879-020-05105-4.
Publication Type: Original article
Contribution: Conceptualized the study, supervised field work, performed and supervised laboratory experiments, curated and analysed data, administrated the project, contributed resources, performed the validation process, co-wrote, read, reviewed, revised and approved the manuscript.
Journal Impact Factor: 2.6

édition scientifique
VVB LAUFERSWEILER VERLAG

VVB LAUFERSWEILER VERLAG
STAUFENBERGRING 15
D-35396 GIESSEN

Tel: 0641-5599888 Fax: -5599890
redaktion@doktorverlag.de
www.doktorverlag.de

ISBN: 978-3-8359-6959-9

Photo cover: © Author