Journal of Applied Microbiology

Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

Development of a hierarchical typing approach for *Mycobacterium avium* subsp. *paratuberculosis* (MAP) and characterization of MAP field cultures from Central Germany

A. Fawzy[1,2,3], M. Zschöck[3], C. Ewers[2] and T. Eisenberg[2,3]

1 Department of Medicine and Infectious Diseases, Faculty of Veterinary Medicine, Cairo University, Giza Square, Egypt
2 Institute for Hygiene and Infectious Diseases of Animals, Justus-Liebig-University, Giessen, Germany
3 Hessian State Laboratory, Giessen, Germany

Keywords
epidemiology, Germany, MAP, MIRU-VNTR, SNP, typing.

Correspondence
Ahmad Fawzy, Hessian State Laboratory, Schubertstr. 60, 35392 Giessen, Germany.
E-mail: dr.ahmedfawzy_vet@yahoo.com

2019/1951: received 29 October 2019, revised 30 April 2020 and accepted 13 May 2020

doi:10.1111/jam.14722

Abstract

Aims: Development of a novel hierarchical *Mycobacterium avium* subsp. *paratuberculosis* (MAP) typing approach and characterization of MAP field cultures in Central Germany.

Methods and Results: By combining single nucleotide polymorphisms (SNPs) and mycobacterial interspersed repetitive unit-variable number tandem repeat, we developed a highly discriminating and phylogenetically accurate hierarchical MAP typing approach. Moreover, a novel stepwise workflow was employed to reduce the number of SNP reactions required making the typing approach more affordable. MAP field cultures (n = 142) from dairy herds in Central Germany were classified as cattle type and showed a high level of heterogeneity. Intra-herd multiple genotypes were evident in (13–25%) of the investigated herds.

Conclusions: The hierarchical MAP typing approach proved to be useful in fine discrimination between MAP cultures within limited geographical regions. This could potentially be used in unravelling MAP transmission chains in the respective regions. The observed heterogeneity in some herds is assumed to be due to either multiple introductions through inter-herd trade or intra-herd evolution over time.

Significance and Impact of the Study: Future MAP epidemiological studies will benefit from the advantages of the novel hierarchical typing approach. The SNP number reduction approach employed here could be extrapolated for other analogous pathogens.

Introduction

Mycobacterium avium subsp. *paratuberculosis* (MAP) is a member of the *M. avium* complex and the causative agent of Johne's disease, one of the most economically important diseases in ruminants worldwide (Losinger 2005). Clinical signs are characterized by emaciation and chronic granulomatous enteritis and MAP is suspected to play a role in Crohn's disease in humans (Liverani *et al.* 2014). Basically, two major genetic lineages were described for MAP isolates, namely Cattle (C) type and Sheep (S) type (Biet *et al.* 2012). Typing of the involved pathogen is a prerequisite for understanding the epidemiology of MAP infection and to trace back the sources of infection with an ultimate goal of designing more efficient control programmes (Bannantine *et al.* 2013). It is also an important element for basic microbiological research investigating essential bacterial traits such as pathogenicity, virulence and antibiotic resistance (Li *et al.* 2009).

Researchers have developed and applied a variety of MAP genotyping methods (Bannantine *et al.* 2013). Especially in the last decade, mycobacterial interspersed repetitive unit-variable number tandem repeats (MIRU-VNTR) were among the most commonly used markers for MAP genotyping mainly due to its portability and reproducibility (Castellanos *et al.* 2010; Oakey *et al.* 2014; Rónai *et al.* 2015). MIRU-VNTR loci are specific DNA sequences repeated in a head to tail manner where the repeat copy number varies among bacterial strains. Their application for MAP typing has been used in combination with other markers to increase the overall typing resolution (Thibault *et al.* 2008; Möbius *et al.* 2008; Fritsch *et al.* 2012). However, whole genome sequencing (WGS)-based phylogeny of MAP isolates has recently revealed that MIRU-VNTR analysis may be prone to homoplasy (Ahlstrom *et al.* 2015; Bryant *et al.* 2016). Homoplasy means that observed similarities between individuals are indeed not due to a shared common ancestor, but a result of convergent or reverse evolution usually caused by rapidly mutating markers such as MIRU-VNTR (Comas *et al.* 2009; Bryant *et al.* 2016). Therefore, using MIRU-VNTR in a MAP typing scheme can potentially lead to phylogenetically erroneous typing results resulting in under- or over-estimation of the relatedness between isolates.

To overcome homoplasy, combining genetic markers with different mutation rates in a hierarchical manner has proven to be a highly discriminating and phylogenetically accurate typing scheme in other monomorphic pathogens such as *Bacillus anthracis* and *Mycobacterium tuberculosis* (Keim *et al.* 2004; Comas *et al.* 2009). Respective schemes start with markers of slow mutation rate as single nucleotide polymorphisms (SNPs) to define the phylogenetic roots followed by rapidly mutating ones like VNTRs to define peripheral phylogenetic branches in a high-resolution manner. Only recently, Leao *et al.* (2016) has employed 14 SNPs in a comprehensive MAP typing scheme exploiting the WGS data of a set of 133 out of the 141 genomes generated for isolates originating from 17 countries and eight different hosts in another study (Bryant *et al.* 2016).

In this work, we aimed to employ this newly described SNP assay to develop an analogous hierarchical MAP typing approach. Within the frame of this approach, we also attempted to make the SNP analysis more affordable by decreasing the number of required PCRs; wherein we first carried out MIRU-VNTR followed by MIRU-VNTR genotype-guided stepwise SNP analysis. We further aimed to characterize MAP field cultures to unravel some epidemiological aspects of MAP infection in Central Germany.

Materials and methods

MAP strains/cultures

One MAP reference strain (DSM 44133) and MAP field cultures ($n = 142$, Table S1) were used in this study. Field cultures were obtained from the mycobacteria collections of the Hessian State Laboratory, Giessen, Germany. They were collected in the course of voluntary MAP surveillance programmes in three German federal states (Hesse, Thuringia and Saxony; Eisenberg *et al.* 2013; Donat *et al.* 2016; Hahn *et al.* 2017). The respective cultures were primarily cultivated on Herrold's egg yolk medium (HEYM) agar slants supplemented with mycobactin J (MJ; HEYM-MJ; BD, Heidelberg, Germany) using faecal or environmental samples originating from dairy herds ($n = 81$), where growth was observed within 6–8 weeks. Each herd has provided one or more cultures and the geographical distribution of the herds in the respective states is illustrated in Fig. 1. The term 'cultures' was used instead of 'isolates' since subcultures maintained after primary cultivation were often carried out using several colonies in case of single colonies or colony material in case of a bacterial lawn.

DNA isolation and MAP confirmation

DNA was extracted by mixing a loop full of the colony material harvested from the surface of agar slants in 100 µl distilled water followed by boiling at 100°C for 15 min. The DNA was then processed from the supernatant after two consecutive centrifugation steps each at 20 817 ***g*** for 5 min. No measurement of the DNA concentration was carried out afterwards. DNA extracted from all MAP field cultures were subjected to a MAP-specific commercial qPCR kit (VetMAX^{TM} MAP Real-Time PCR Screening Kit; Thermo Fisher Scientific, Karlsruhe, Germany) according to the instructions of the manufacturer. For every qPCR run, a positive (DNA extract from the MAP reference strain DSM 44133) and a no-template control (NTC) were used. The identity of all MAP field cultures was confirmed by giving specific amplification.

MIRU-VNTR PCRs (INMV typing)

In singleplex PCRs, the eight loci described by Thibault *et al.* (2007) were tested in all MAP strains/cultures. Each PCR reaction (20 µl) contained 10 µl of Hotstar Taq MasterMix (Qiagen, Hilden, Germany), 1 µl of each forward and reverse primer (10 pmol μl^{-1}) as described by Thibault *et al.* (2007), 6 µl DNase free PCR-grade water (Qiagen) and 2 µl of the extracted DNA. For

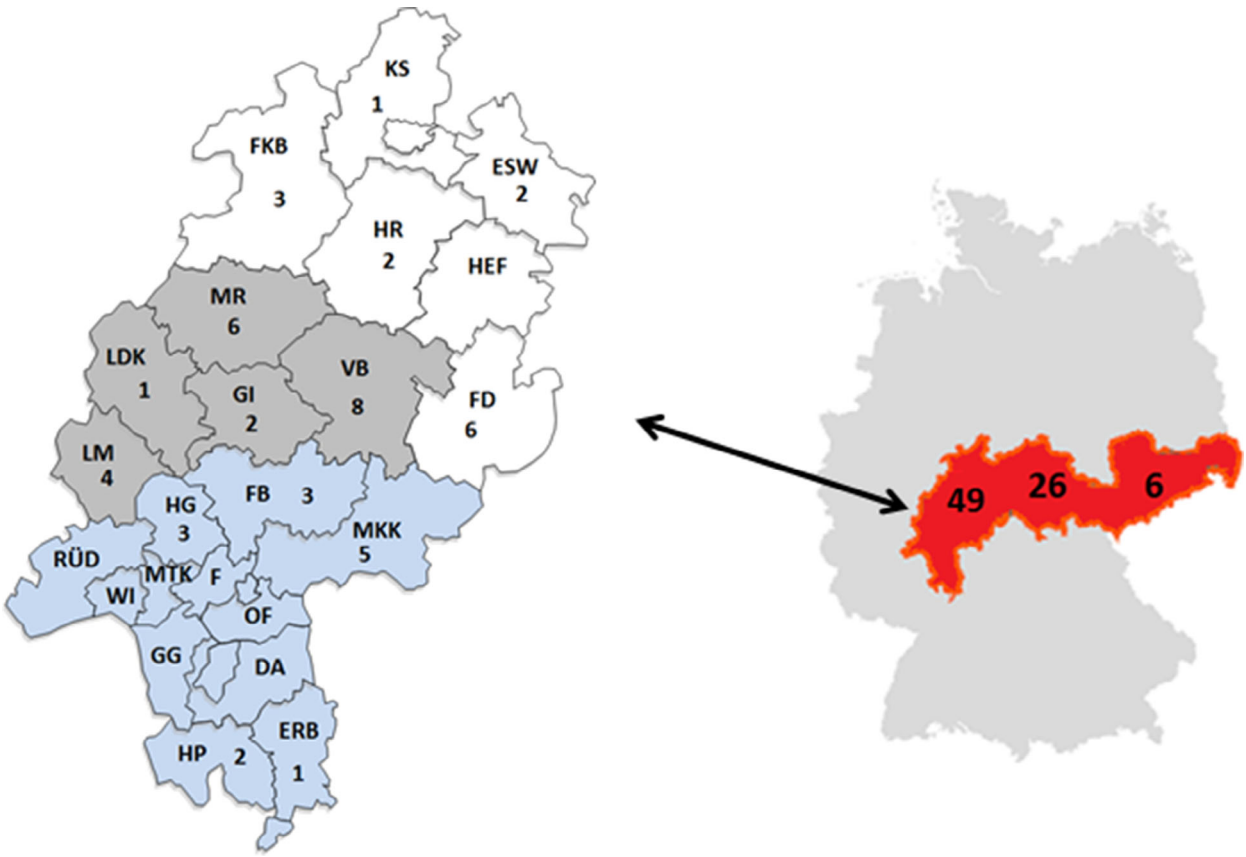

Figure 1 Distribution of dairy herds from which MAP cultures were obtained in three German federal states (Hesse, Thuringia and Saxony). In Hesse, the distributions in the different administrative provinces (North (Kassel), Centre (Giessen) and South (Darmstadt)) and counties (depicted as county abbreviations) are given. No data were available for T and S. KS (Kassel), FKB (Waldeck-Frankenberg), HR (Schwalm-Eder), ESW (Eschwege), HEF (Werra-Meißner), FD (Fulda), MR (Marburg-Biedenkopf), VB (Vogelsberg), LDK (Lahn-Dill), GI (Giessen), LM (Limburg-Weilburg), HG (Hochtaunus), FB (Friedberg), MKK (Main-Kinzig), RÜD (Rheingau-Taunus), WI (Wiesbaden), MTK (Main-Taunus), F (Frankfurt/M), OF (Offenbach), GG (Groß-Gerau), DA (Darmstadt-Dieburg), HP (Bergstrasse) and ERB (Odenwald). [Colour figure can be viewed at wileyonlinelibrary.com]

VNTR loci 3, 10 and 32, 1x Q solution (Qiagen) was added to the reaction according to the manufacturer's instructions to overcome the nonspecific reactions associated with amplification of guanine and cytosine (GC)-rich templates (Mamedov *et al.* 2008). Cycling conditions were as following: 1× (95°C—15 min) 40× (94°C—30 s; 58°C—30 s; 72°C—30 s) 1× (72°C—10 min). All PCR reactions were carried out in a thermocycler (T3000; Biometra, Göttingen, Germany). A NTC was included in every PCR run as a control against contamination. PCR products were stained with ethidium bromide in a 2% agarose gel (100 V for 1 h) and visualized with a gel documentation system (BioDoc-It, UVP, UK). PCR products of VNTR 7 locus were purified, sequenced and analysed as described before (Fawzy *et al.* 2016). The repeat copy numbers of the eight loci were combined into a numerical code and each isolate was assigned to an **IN**RA (National Institute of Agricultural Research, France) **M**IRU-**V**NTR (INMV) type as described before (Thibault *et al.* 2007).

SNP PCRs

In singleplex PCRs, 13 SNPs (snp3842359 and 12 SNPs listed in Table S2) recently described by Leao *et al.* (2016) were tested in this work using the same primers. Instead of using the decision tree described by Leao *et al.* (2016) for SNP analysis, we adopted a novel INMV-guided and stepwise typing approach as described in detail in the results section. PCR master mix and amplification conditions were identical to those described above for PCRs of VNTR loci 3, 10 and 32. PCR products were purified using MicroElute DNA Cycle-Pure Kit (Omega bio-tek, Norcross, GA) according to the instructions of the manufacturer and submitted to Seqlab-Microsynth laboratories (Göttingen, Germany) for sequencing. BLASTN

analysis was carried out using DNA sequences against the MAP strain K10 genome (GenBank accession number AE016958.1) for SNP identification. Moreover, a visual verification of the results was carried out by inspection of the chromatograms to check for double peaks at each SNP position.

Phylogenetic analysis

Phylogenetic analyses for SNPs and MIRU-VNTR (INMV) were carried out with the help of the online database MLVAbank (available at http://mlva.u-psud.fr/mlvav4/genotyping/; accessed on 10.11.2018 (Grissa *et al.* 2008)) using unweighted pair group method with arithmetic mean clustering method. Phylogenetic trees were visualized using the online tool TreeDyn (available at http://phylogeny.lirmm.fr/phylo_cgi/index.cgi; accessed on 19.11.2018 (Chevenet *et al.* 2006)). A composite tree as a basis for a hierarchical typing approach (SNP+INMV) was constructed by combining the two respective phylogenetic trees together.

Discriminatory power calculations

MAP cultures/strains used

In addition to the MAP reference strain DSM 44133, only epidemiologically unrelated cultures were used (underlined in Table S1; $n = 78$). Cultures originating from the same herd and having the same genotype were considered epidemiologically related and were excluded from calculations as were also cultures showing multiple bands at one or more MIRU-VNTR loci.

Discriminatory power

The discriminatory power of the typing panels both separately and combined was evaluated using the discriminatory index (DI) with the help of an online tool available at (http://www.comparingpartitions.info/index.php?link=Tool accessed on 15.10.2018), which uses the following formula (Hunter and Gaston 1988) for calculation:

$$\mathrm{DI} = 1 - 1\,\frac{1}{N(N-1)}\sum_{j=1}^{s} N_j\left(N_j - 1\right)$$

where N is the total number of epidemiologically unrelated strains, S is the total number of genotypic profiles determined by each typing panel or by the combination of individual panels and N_j is the number of strains belonging to the jth profile. DI is a measure of the probability that a certain typing method would discriminate between two epidemiologically unrelated strains of a particular microbial population. The online tool also calculates the confidence intervals of 95% (CI 95%) for DI according to (Grundmann *et al.* 2001) enabling an objective comparison between different typing techniques. Moreover, to figure out if the DI values of different typing panels are statistically different, p-values were also calculated according to the jackknife pseudo-values resampling method (Severiano *et al.* 2011). Three categories of significance values were considered. 1. $P \leq 0{\cdot}05$ is weakly statistically significant, 2. $P \leq 0{\cdot}01$ is statistically significant and 3. $P \leq 0{\cdot}001$ is statistically highly significant.

Results

Development of a hierarchical MAP typing approach (SNP+INMV)

Although INMV typing was carried out in the laboratory first (Table S1) followed by an INMV-guided and stepwise SNP analysis as detailed below, a novel hierarchical MAP typing approach was adopted in this work, in which the order of analysis was reversed. At the first level, cultures were allocated to a phylogenetic group based on SNP analysis (Leao *et al.* 2016). At the second level, INMV typing was used for finer discrimination within each group.

A novel INMV-guided SNP typing approach

Using 14 SNPs, Leao *et al.* (2016) have adopted a decision tree based on four sequential steps to characterize MAP isolates into 14 phylogenetic groups. In an attempt to reduce the number of required SNP reactions, we used INMV typing results to design a novel stepwise SNP typing approach as an alternative to the one adopted by Leao *et al.* (2016). This was based on the assumption that INMV types could indeed carry valuable phylogenetic signals as previously described for VNTR loci of other pathogens such as *Salmonella enterica* serovar Typhi (Octavia and Lan 2009). This means that certain INMV types could be strongly associated with, but not necessarily exclusive for, a specific phylogenetic group. The phylogenetic signals of INMV types were employed as follows:

INMV types can differentiate between MAP types C and S. First, all MAP cultures/strains under study ($n = 143$) were subjected to a SNP analysis (snp3842359) that could differentiate between the two major MAP types C and S (Leao *et al.* 2016). Upon BLASTN analysis of the respective sequences, all cultures displayed an adenine (A) nt at the SNP position, which is characteristic for MAP type C.

Second, we conducted a phylogenetic analysis for a collection of MAP isolates/cultures (types C and S) based on

their INMV results (alleles). Besides INMV types uniquely described in this work ($n = 7$, Fig. 2), this collection also included other INMV types ($n = 40$) reported in two comprehensive MAP genotyping studies (Biet *et al.* 2012; Bryant *et al.* 2016). As illustrated in Fig. 2, the INMV-based phylogenetic tree clearly differentiates between types C and S into two clades. Therefore, we now tentatively propose this approach as a means for differentiating between MAP types C and S. In upcoming studies, researcher(s) should subject each identified INMV type, together with other types described in this work, to phylogenetic analysis as described in the Materials and Methods section. The researcher may then assign the INMV type to one of the two types (C or S) according to the clade it clusters with.

Prioritization of SNP analysis for MAP type C. Based on previously published data (Bryant *et al.* 2016; Leao *et al.* 2016), we used INMV types to prioritize SNP analysis for MAP type C cultures in the form of a workflow (Fig. 3). The basics and components of this approach could be summarized as follows:

i The SNPs described for phylogenetic groups of MAP type C isolates are mutually exclusive, that is, when an isolate possesses a certain SNP, it lacks all other ones (Leao *et al.* 2016). Therefore, if these SNPs would be stepwisely analysed **in a specific order**, the analysis could be stopped once a phylogenetic group-defining SNP is identified.
ii *A prioritized order* of SNPs to be analysed for each MAP type C isolate/culture could be achieved when the phylogenetic signals of the INMV types are used as a guide (Fig. 3).
iii *Phylogenetic signals* inferred from INMV types depend on the current knowledge about the association between INMV types and the phylogenetic groups. Therefore, we collected the respective data published to date (Bryant *et al.* 2016; Leao *et al.* 2016) in one table (Table S2). Based on these data, two phylogenetic signals were defined to tailor the order of SNPs to be analysed, namely *identity match* and *relative abundancy*. Briefly, using the data in Table S2 one should first test the phylogenetic group that have the most similar INMV type and if more than one group possess the same INMV type, the group with the largest number of isolates should be tested first followed by the other groups in a descending manner. The approach is described in detail in Fig. 3.
iv MAP cultures were analysed in this work. Therefore, a heterogeneous MAP population could exist in one culture. If at a given analysed SNP double peaks were identified, the analysis should be extended for all SNPs to figure out the other phylogenetic group(s) coexisting in the respective culture.
v Cultures possessing multiple INMV types/alleles were not analysed according to the workflow described in Fig. 3 and a stepwise SNP analysis was carried out in the following order (bison group, subgroup B and clade 1–clade10).

As a practical application of the proposed workflow, the order of SNPs to be analysed for the INMV types described in this work is presented in Table S3. The number of reactions that were required for detection of the phylogenetic group for each culture is shown in Table S1 (last column on the right). One or two SNP reactions were required to detect the phylogenetic group in the majority of the investigated cultures/strains.

Added value of the novel hierarchical typing approach (SNP+INMV)

As illustrated in Fig. 4, the hierarchical typing approach (SNP+INMV) bypasses the homoplasy associated with INMV typing (e.g. cultures of the INMV 2 type were found to belong to five different SNP-defined phylogenetic groups). Concerning discriminatory power, the combination of both typing panels (SNP+INMV) resulted in a statistically highly significant (P-values $< 0{\cdot}001$ and $0{\cdot}005$) higher DI value ($0{\cdot}938$; CI 95% $0{\cdot}917$–$0{\cdot}960$) compared to SNP ($0{\cdot}829$; CI 95% $0{\cdot}791$–$0{\cdot}867$) and INMV ($0{\cdot}879$; CI 95% $0{\cdot}834$–$0{\cdot}924$) panels respectively.

Characterization of MAP field cultures from Central Germany

MIRU-VNTR analysis (INMV types)

Eighteen different INMV profiles were evident (Table S1). In the online INMV database (available at http://mac-inmv.tours.inra.fr/ accessed on 10.06.2018), there were no matches for six INMV profiles (32422128, 43332128, 22522328, 33332128, 42332118 and 32332318). These profiles have later been included in the database upon request under the INMV types 142, 143, 151, 152, 153 and 214 respectively. A relatively high DI ($0{\cdot}879$; CI (95%) $0{\cdot}834$–$0{\cdot}924$) was evident (calculated based on the epidemiologically unrelated cultures, underlined in Table S1). INMV 2 was the most commonly identified type (28%), followed by one of the new profiles (INMV 143) with a frequency of 17% (Table 1). On the other side, four INMV types (118, 142, 152 and 153) were each represented by only one culture (Table 1). In nine environmental cultures, double alleles were detected at one or two INMV loci (Table S1).

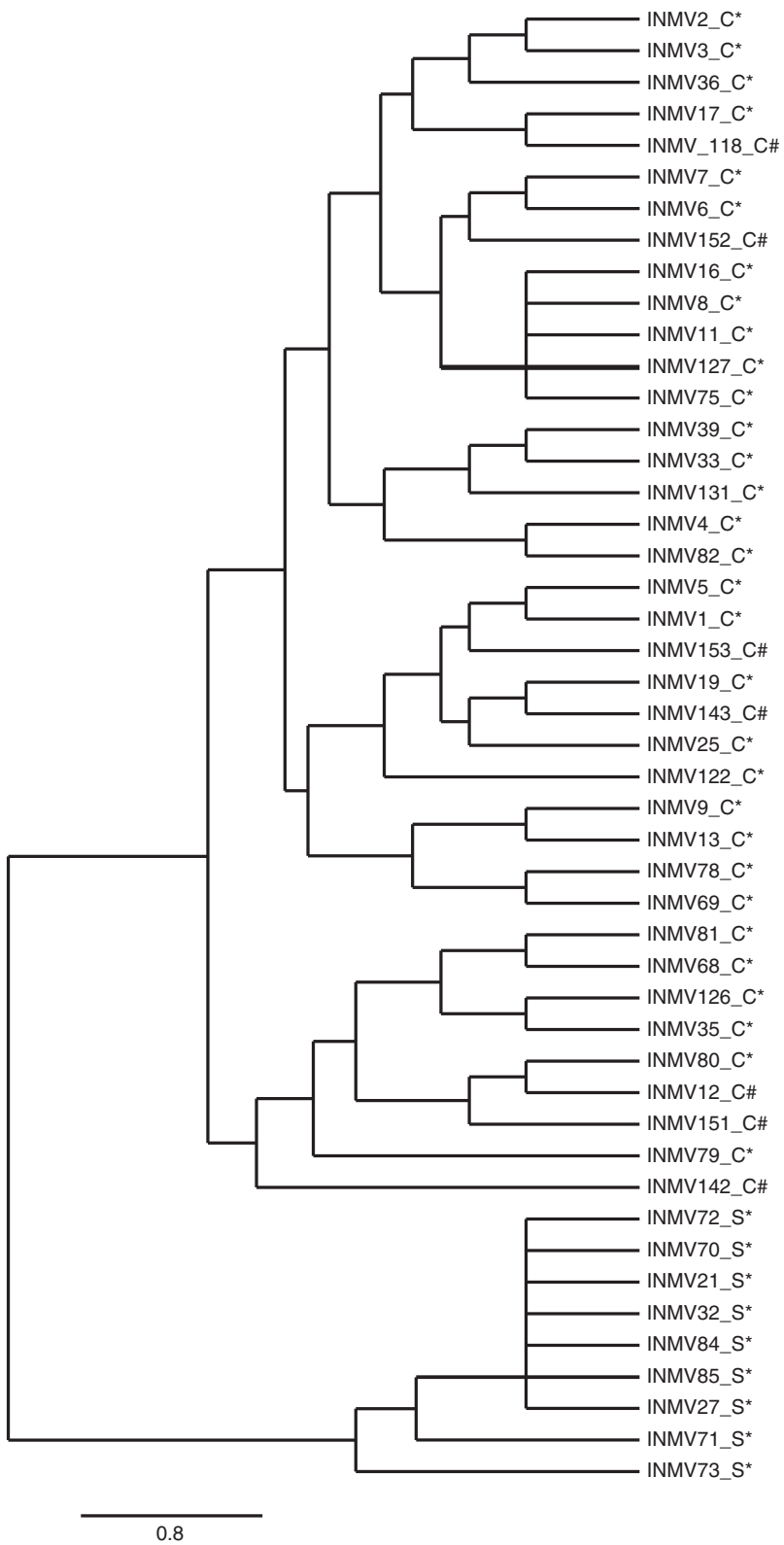

Figure 2 INMV-based phylogenetic tree (UPGMA) for the discrimination into MAP types C and S. INMV types are those uniquely described for MAP type C (#) cultures in this study (Table S1) and other isolates (types C and S (*)) previously described in two typing studies (Biet *et al*. 2012; Bryant *et al*. 2016). The phylogenetic tree was constructed and visualized with the help of two online tools (see Materials and Methods section). The branch lengths are proportional to the number of substitutions per site (bar).

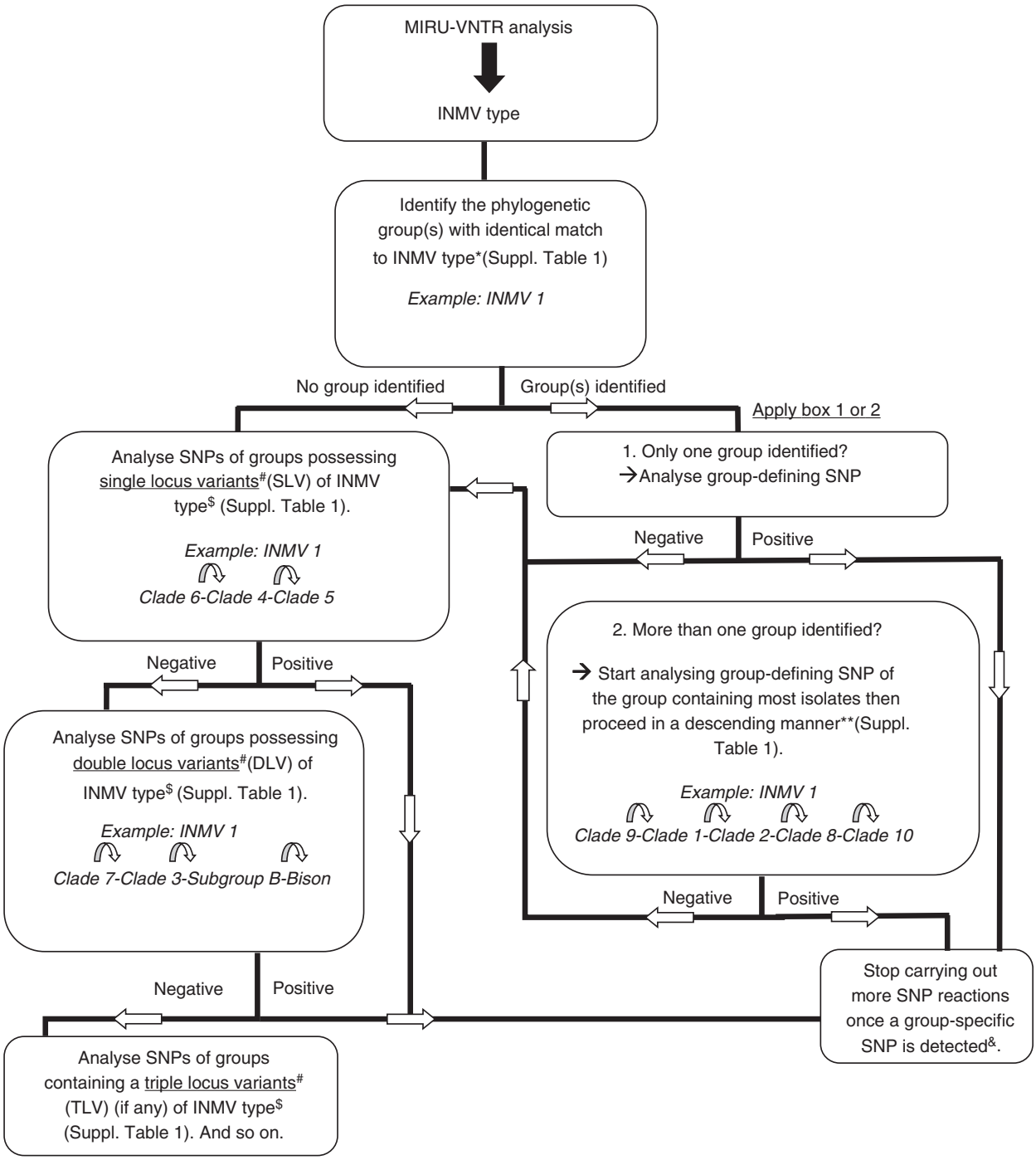

Figure 3 A schematic workflow for a prioritized and stepwise SNP analysis of MAP Type C isolates/cultures based on phylogenetic signals provided by INMV types. *The first applied phylogenetic signal is identity match (see Results section). &SNPs described for phylogenetic groups of MAP type C isolates are mutually exclusive, that is, when an isolate possesses a certain SNP, it lacks all other ones (Leao *et al*. 2016). The only exception would be in case of detecting the subgroup B-specific SNP, where the bison group-specific SNP must also be analysed. **The second applied phylogenetic signal is the relative abundance of each INMV type in the different genetic groups (see Results section). #SLV, DLV or TLV are INMV types that differ from the INMV type under analysis at one, two or three of the eight loci. $The order of genetic groups (SNPs) to be analysed in these cases is not fixed. Therefore, the number of reactions could slightly vary when different persons apply this approach. However, the second phylogenetic signal (**) applies for groups having identical SLVs or DLVs.

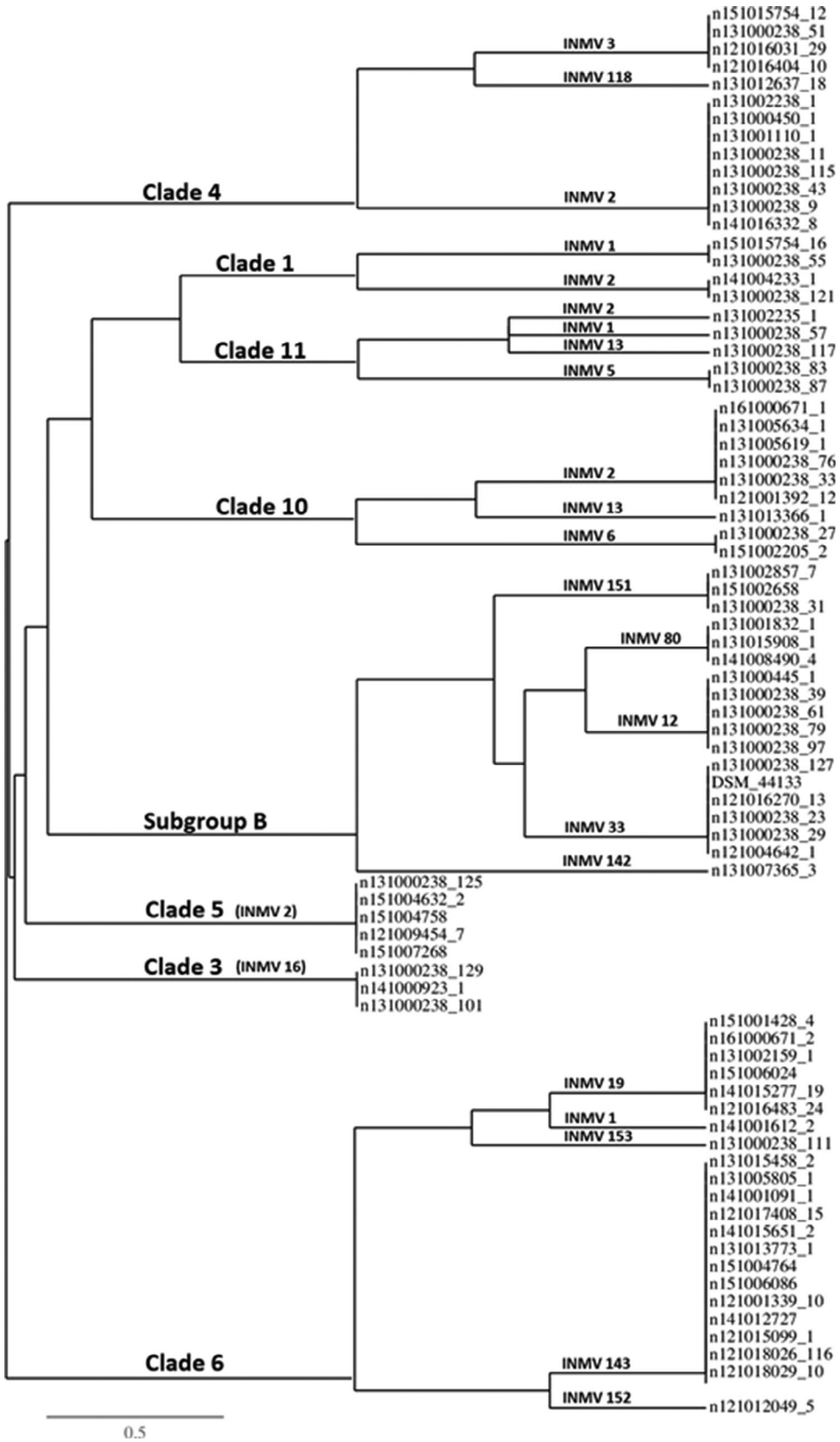

Table 1 INMV genotypes of the epidemiologically unrelated MAP field cultures under study (n = 78)

INMV type	INMV code*	No. of cultures	%
1	42332228	4	5
2	32332228	22	28
3	32332218	4	5
5	42332218	2	3
6	32332128	2	3
12	22522228	5	6
13	22332228	2	3
16	32332528	3	4
19	42332128	6	8
33	32522228	5	6
80	22522226	3	4
118	31332218	1	1
142	32422128	1	1
143	43332128	13	17
151	22522328	3	4
152	33332128	1	1
153	42332118	1	1

*See legend of Table S1.

SNP analysis (phylogenetic groups)

With regard to results of SNP analysis, all MAP cultures under study are of the type C and are differentiated into eight different phylogenetic groups (clades 1, 3, 4, 5, 6, 10 and 11 as well as subgroup B; Table S1). Figure S1 depicts the geographical distribution of the identified phylogenetic groups in the three German federal states included in this study.

Clade 6 represented about 40% of MAP cultures in Hesse (21/52), while only one culture was found to belong to the same phylogenetic group in Thuringia (1/20; 5%). The frequencies of other phylogenetic groups in Hesse were as following: subgroup B (9/52; 17%), clade 4 (8/52; 15%), clade 10 (6/52; 12%), clade 5 (4/52; 8%), clade 1 (2/52; 4%), clade 3 (1/52; 2%) and clade 11 (1/52; 2%). In Thuringia, on the other hand, equal numbers of cultures (n = 5) were found to belong to clade 4 and subgroup B respectively, each representing 25% of the total number of cultures in this federal state. The frequencies of other phylogenetic groups in Thuringia were as following: clade 11 (3/20; 15%), clade 1 (2/20; 10%), clade 3 (2/20; 10%), clade 5 (1/20; 5%) and clade 10 (1/20; 5%). Among the six MAP cultures investigated in Saxony, three cultures (50%) were found to belong to subgroup B; two cultures (33%) belonged to clade 10 and one culture (17%) to clade 11.

Potential application of SNP+INMV typing in epidemiological tracing

In the epidemiological context, the hierarchical typing approach (SNP+INMV) was able to differentiate between MAP cultures within the same phylogenetic lineage even in small geographic areas where Fig. S2 illustrates practical examples. In Hesse, cultures belonging to clade 6 in each of the three provinces were further differentiated into distinct INMV types (Fig. S2a). In Thuringia, cultures belonging to clade 4 and subgroup B were also discriminated into variant INMV types (Fig. S2b). This, in turn, has the potential of identifying the MAP transmission chains between herds of each province at a higher resolution.

Intra-herd strain diversity

The MAP cultures used in this work enabled the investigation of the intra-herd strain diversity in herds, from which MAP cultures were obtained either from environmental (n = 52), multiple individual faecal (n = 4) or a combination of both sample types (n = 14; Table S1).

Multiple strains (different alleles at one or more MIRU-VNTR loci) were detected in 13% of the herds (7/52; six from Thuringia and one from Hesse) that provided only environmental samples (Table S1). To exclude nonspecific PCR products as a cause of the apparent multiple alleles, PCR of MIRU-VNTR loci showing multiple alleles were repeated three times along with positive and negative controls yielding the same results. Moreover, two cultures obtained from two different environmental samples from the same herd (ID 75) showed identical patterns of multiple alleles at the VNTR 10 locus (Table S1). This further excludes PCR artefacts as a possible cause of the multiple alleles in these herds. Cultures obtained from two herds (IDs 76 and 80) displayed double bands at two different MIRU-VNTR loci. Therefore, no INMV types were assigned to these cultures since the real combination of alleles at the respective loci could not be inferred. In one of the seven herds (ID 77), two phylogenetic groups (clades 1 and 10) were identified upon SNP analysis (Table S1). These two groups most likely

Figure 4 Composite tree for the MAP cultures from this study combining two phylogenetic trees based on SNPs (UPGMA) and MIRU-VNTR (INMV) data (UPGMA) within each SNP-clade. The phylogenetic trees were constructed and visualized with the help of two online tools (see Materials and Methods section). Only epidemiologically unrelated MAP cultures (Table S1) were used for tree construction. SNP clusters and INMV types were named according to Leao *et al.* (2016) and Thibault *et al.* (2007) respectively. All SNP clusters other than subgroup B belong to subgroup A described by Leao *et al.* (2016). Black arrows exemplify a case of homoplasy, where cultures have the same INMV type (e.g. INMV 2), but indeed belong to five different SNP-defined phylogenetic groups.

coexist in the respective herd due to purchasing animals from different sources.

Four herds (IDs 2, 11, 58 and 81) were represented only by multiple faecal cultures (2–8 cultures/herd). In only one of the respective herds (ID 81), two MAP strains were identified (Table S1). These two strains displayed different phylogenetic groups (clades 6 and 10) and different INMV types (INMV 2 and 19) respectively. The genotype designation clade 10-INMV 2 was most frequently found (6/8; 75%).

Both environmental and faecal cultures were available for 14 herds (2–19 cultures/herd). Strain diversity was evident in 21% (3/14) of the respective herds (IDs 42, 61 and 68; Table S1). In one herd (ID 42), two strains were identified in different cultures. One strain (culture ID 141016332-8) belonged to the phylogenetic group (clade 4) and INMV 2 type. The other strain (culture ID 141000923-1) belonged to a different phylogenetic group (clade 3) and had a different INMV type (INMV 16). In the second herd (ID 61), two different strains were evident (Table S1). They belonged to two different phylogenetic clades (clade 1 and 4) and displayed different INMV types (INMV 1 and 3), respectively. In the third herd (ID 68), one environmental culture (ID 131015036-01) displayed two single locus variants at locus 292 (Table S1).

Discussion

In this work, we developed a sophisticated hierarchical MAP typing approach with high discriminatory power. This approach starts by defining the phylogenetic groups using SNPs as stable evolutionary markers followed by within-group MIRU-VNTR analysis for a high-resolution typing. Furthermore, the genetic structure of MAP field cultures obtained from dairy herds in three German federal states (Hesse, Thuringia and Saxony) was unravelled.

The hierarchical approach (SNP+INMV) has proven to be advantageous in two main aspects. Regarding phylogenetic accuracy, a baseline definition of phylogenetic groups by SNP analysis helped indeed to overcome homoplasy associated with MIRU-VNTR typing (Fig. 4). Using the hierarchical approach will in turn guarantee an accurate evaluation of relatedness among tested individuals. With respect to discriminatory power, the hierarchical approach displayed a significantly higher DI value compared to both typing panels independently (see Results section). The discrimination level achieved by the proposed approach (DI 0·938) is sufficient for short-term epidemiological applications such as outbreak investigations and regional surveillance. This follows previous recommendations that a DI value of ≥0·9 is important to interpret epidemiological data with confidence (Hunter and Gaston 1988; Mokrousov 2017), which is not achieved by both typing panels alone.

In contrast to findings of Leao *et al.* (2016), MIRU-VNTR (INMV panel) exhibited a slightly higher discriminatory power in this work compared to SNP analysis (see Results section), however, the differences were not statistically significant (*P*-value 0·143). This contrast in the results is probably due to the variation in the genetic background of the typed isolates/cultures in both studies.

Similar hierarchical typing approaches combining SNP and VNTR exist in the literature for a phylogenetically robust and highly discriminating characterization of other monomorphic bacteria like *B. anthracis* and *M. tuberculosis* (Keim *et al.* 2004; Comas *et al.* 2009). However, the approach implemented in the present work employed a novel modification that led to reduction in the number of SNP reactions required and consequently the overall analysis costs. In future studies, analogous approaches for the aforementioned pathogens could benefit from this modification as well.

By carrying out INMV typing first, the phylogenetic signals carried by INMV types have been utilized to stepwisely guide the SNP analysis. First, the SNP was previously employed in the assay described by Leao *et al.* (2016) to discriminate between MAP types C and S, can be skipped. Alternatively, we tentatively propose a phylogenetic analysis based on INMV types (Fig. 2) as a means to discriminate between MAP types C and S. These findings are in accordance with the observations of Biet *et al.* (2012) and Bryant *et al.* (2016). Further verification of these findings is recommended to unequivocally prove this step with more strains' information becoming available. Second, a novel workflow was designed in this work to tailor the SNP analysis of individual MAP type C isolates/cultures (Fig. 3). Although this workflow seems to be sophisticated, it could be easily programmed as a simple computer algorithm that would greatly facilitate its application in further studies. Similar workflows can also be developed for other monomorphic pathogens, for which a hierarchical typing approach was developed (Keim *et al.* 2004; Comas *et al.* 2009).

By applying this workflow in the present work, the majority of cultures required one or two SNPs for identification of the phylogenetic groups (Table S1). On the other hand, 12 SNPs would have been required for the analysis of the majority of cultures by the workflow suggested by Leao *et al.* (2016). In the latter workflow, 12 SNPs are employed for assigning MAP type C isolates to 12 phylogenetic groups/clades in three sequential steps. In the last step, the authors recommend carrying out 10 SNP reactions in parallel to assign each isolate within subgroup A, which constitutes the majority of cultures in the present work, to 1 of 10 different clades. In fact, these

10 clade-specific SNPs are mutually exclusive ((Leao *et al.* 2016); Table S2), that is, when an isolate/culture possesses a certain clade-specific SNP, it lacks all other ones. Therefore, this last step could actually be split into 10 sequential steps in a certain predefined order (e.g. SNP 1–SNP 10) and the analysis is to be discontinued once a clade-defining SNP is identified. If the latter approach would be applied, the analysis of each isolate/culture in subgroup A would begin with the two sequential steps followed by a number of sequential SNP reaction equivalents to the order of its clade. Compared to the workflow described in this study, this would require profoundly more SNP reactions and time. Although INMV analysis (eight PCRs) is required in our approach as a basis for SNP prioritization, it leads to increased discriminatory power when combined to SNP analysis. Therefore, taking both the discrimination power and the number of reactions required to complete the analysis into consideration, our approach outperforms the scenarios of SNP analysis, INMV analysis or the analysis of both of them without applying our prioritization approach.

All MAP cultures in this work were classified as type C. This is in accordance with previous reports from Germany, where only one publication reported the isolation of MAP type S isolates from sheep (Möbius *et al.* 2009). This could be due to the relatively low population size of sheep in Germany compared to other countries with larger populations such as Great Britain and Australia where MAP type S isolates are found more frequently.

In previous studies, DI values for the INMV panel ranged from 0·381 for Canadian (Sohal *et al.* 2014) to 0·88 for Austrian MAP isolates (Gerritsmann *et al.* 2014). An almost identical value was detected in this work for MAP cultures in three German states (DI 0·879). This reflects the considerable diversity of MAP strains in Germany in agreement with previous reports (Möbius *et al.* 2008; Fernández-Silva *et al.* 2012). This might be due to the relatively large population of dairy cattle in Germany. Moreover, Germany's geographical location in central Europe and regular trading across its borders might have played a role in the introduction of MAP from different sources.

In agreement with some epidemiological studies (Thibault *et al.* 2007; Fritsch *et al.* 2012), INMV type 2 was the most common type (28%) in the tested MAP cultures. However, in contrast to other studies (Thibault *et al.* 2007; Stevenson *et al.* 2009; Douarre *et al.* 2011; Fernández-Silva *et al.* 2012), INMV type 1 was much less frequent (5%). In the present work, six new INMV types were identified. However, the genetic profile of one of these types (INMV 143), which represented the second most common genotype among cultures of the present study (17%), was already reported in a previous German study (Fernández-Silva *et al.* 2012). In this study, four MAP isolates possessing the respective genetic profile were reported; one isolate from Hesse and three from Rhineland-Palatinate, a directly adjacent German federal state that was not investigated in the present work. Until the time of writing this work, INMV 143 has not been described from outside Germany, yet. Further MAP epidemiological studies from other geographical regions would support or refute the hypothesis of geographical clustering of this genotype. This could have implications on control measures in Germany such as selecting certain MAP strains for diagnostics and vaccine production.

Concerning SNP analysis, MAP type C cultures under study were differentiated into eight different phylogenetic groups out of a total of 13 yet described groups (Leao *et al.* (2016); Table S1). This confirms the relatively high genetic diversity of MAP strains in Germany (Möbius *et al.* 2008) also for the three federal states under study despite overall genetic homogeneity of MAP worldwide. Four of the phylogenetic groups described in the present work (clades 1, 5, 6 and 10) were also described for German MAP type C isolates in a recent study (Leao *et al.* 2016). However, the other four groups (subgroup B and clades 3, 4 and 11) are described here for the first time. Cultures belonging to clade 11 are of special interest, since they are assigned to this clade based on the absence of the 10 SNPs specific for clades 1–10. However, they could belong to more than one phylogenetic group. Carrying out WGS for these cultures in upcoming studies will extend our knowledge about the genetic diversity of MAP and help defining new SNP(s) for a better characterization. Subgroup B was previously reported only for MAP isolates from Greece and Spain, while clade 6 was identified only in Spain and Germany (Bryant *et al.* 2016; Leao *et al.* 2016). Whether these phylogenetic groups are restricted to Europe only should be investigated in future studies.

Interestingly, with respect to geographical distribution clade 6 was almost exclusively (21/22 cultures) identified in Hesse which probably reflects a common source of infection or the movement of animals between herds via trade in Hesse (Fig. S1). However, the distribution of phylogenetic groups described here is not a true indication of the prevalence since no sampling design was attempted to infer prevalence.

The hierarchical typing approach (SNP+INMV) was useful in the fine discrimination between MAP cultures belonging to a given phylogenetic group even in small geographic areas such as in the three Hessian regional provinces (Kassel, Giessen and Darmstadt) and in Thuringia (Fig. S2). Consequently, a combination of SNP+INMV has the potential to be applied for deciphering MAP transmission chains between the dairy herds in

the respective areas. This could help in evaluation of the local MAP control programmes currently applied.

In the present work, genotyping results have shown that multiple MAP strains existed in 13–25% of the investigated herds (Table S1). However, this figure could be an underestimation of the magnitude of this phenomenon due to some limitations in the typing strategy used in this work. First, the investigated herds were not well represented. Each herd provided mainly one environmental culture and/or few to several faecal cultures. Second, one faecal culture per cow was used. This cannot exclude the possibility of not identifying strains that would exist in faecal parts other than the inoculum used for cultivation or that were not shed in faeces at the sampling time. Third, one medium (HEYM-MJ) was used for MAP cultivation throughout this work and the type of culture media was found to introduce bias in the spectrum of MAP genotypes that could be isolated from a sample (Cernicchiaro *et al.* 2008). Fourth, the genetic material that was subjected to genotyping could underrepresent the original MAP population in the samples. This assumption is based on the diluting effect that could have been introduced by using only some of the growing bacterial biomass for the subculture(s) and DNA extraction. Fifth, the sensitivity of detecting multiple alleles at the employed genetic markers is not known, but it is assumed not to be perfect, especially when variants exist at low concentrations.

As a possible explanation for the detected intra-herd strain diversity, one can assume two different scenarios: 1. multiple introduction occasions possibly by purchasing infected animals from different sources; 2. intra-herd mutation over time. Despite the high-resolution genotyping applied in this work, it is difficult to designate a cut-off point of genetic relatedness that could distinguish between the two possibilities. However, the first scenario rather fits for the herds, where the identified strains belong to different phylogenetic groups. On the other hand, the existence of strains with minor genetic variation at MIRU-VNTR loci could actually be the result of either of the two scenarios. Nevertheless, it is tempting to speculate that the existence of two strains displaying different alleles at only one MIRU-VNTR locus is probably due to intra-herd mutation over time. This assumption is supported by the known relatively rapid mutation rate of tandem repeat loci (Zhou *et al.* 2014). Moreover, Ahlstrom *et al.* (2015) has previously assumed a within-herd MAP evolution in a Canadian dairy herd, in which two closely related MAP strains (based on genome sequence data) displayed different alleles at the VNTR 7 locus.

These findings support previous studies, which even reported higher rates of intra-herd strain multiplicity. In a high-resolution MAP typing study, a combination of short sequence repeat (SSR) and pulsed-field gel electrophoresis analysis revealed that 60% (20/33) of the investigated cattle herds had more than one genotype (Sevilla *et al.* 2008). However, using MIRU-VNTR analysis, van Hulzen *et al.* (2011) detected multiple strains in 55% (6/11) of the investigated dairy herds that provided more than one isolate for analysis. These accumulating pieces of evidence are changing our understanding of the MAP molecular epidemiology and questioning the utility of analysing a single isolate from one animal as a representative of the MAP genetic background in a herd. Therefore, novel approaches are needed that could facilitate analysing the genetic diversity in a sample at a reasonable cost. In this regard, high throughput sequencing has the capacity to be a good candidate especially with its ever-decreasing costs. Recently, Bachmann *et al.* (2015) have successfully used culture-independent genome sequencing to investigate the genetic diversity of *Chlamydia pecorum* at high resolution.

Altogether, we developed a novel and relatively cost effective MAP typing approach employing both SNPs and MIRU-VNTRs in a hierarchical manner. Under field conditions, the high resolution of this approach was capable of differentiating between closely related MAP cultures originating from small geographical scale in Central Germany. Furthermore, we unravelled important epidemiological aspects of the MAP infection in this region. All the tested MAP cultures were of the cattle type (type C) and displayed a relatively high level of heterogeneity. Multiple genotypes were evident in some of the investigated dairy herds.

Acknowledgements

We thank Ulrike Kling for her assistance to isolate MAP strains and the German academic exchange service (DAAD) for their financial support. The Hessian State Laboratory is supported by the Hessian Ministry for the Environment, Climate Change, Agriculture and Consumer Protection. We also thank Dr Karsten Donat from the Thuringian Animal Disease Fund and the veterinarians in the district administration in Giessen for collecting and sending clinical samples. Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

No conflict of interest declared.

References

Ahlstrom, C., Barkema, H.W., Stevenson, K., Zadoks, R.N., Biek, R., Kao, R., Trewby, H., Haupstein, D. *et al.* (2015)

Limitations of variable number of tandem repeat typing identified through whole genome sequencing of *Mycobacterium avium* subsp. *paratuberculosis* on a national and herd level. *BMC Genom* **16**, 161.

Bachmann, N.L., Sullivan, M.J., Jelocnik, M., Myers, G.S., Timms, P. and Polkinghorne, A. (2015) Culture-independent genome sequencing of clinical samples reveals an unexpected heterogeneity of infections by *Chlamydia pecorum*. *J Clin Microbiol* **53**, 1573–1581.

Bannantine, J.P., Li, L.-L., Sreevatsan, S. and Kapur, V. (2013) How does a *Mycobacterium* change its spots? Applying molecular tools to track diverse strains of *Mycobacterium avium* subspecies *paratuberculosis*. *Lett Appl Microbiol* **57**, 165–173.

Biet, F., Sevilla, I.A., Cochard, T., Lefrancois, L.H., Garrido, J.M., Heron, I., Juste, R.A., McLuckie, J. *et al.* (2012) Inter- and intra-subtype genotypic differences that differentiate *Mycobacterium avium* subspecies *paratuberculosis* strains. *BMC Microbiol* **12**, 264.

Bryant, J.M., Thibault, V.C., Smith, D.G., McLuckie, J., Heron, I., Sevilla, I.A., Biet, F., Harris, S.R. *et al.* (2016) Phylogenomic exploration of the relationships between strains of *Mycobacterium avium* subspecies *paratuberculosis*. *BMC Genom* **17**, 79.

Castellanos, E., Romero, B., Rodríguez, S., de Juan, L., Bezos, J., Mateos, A., Domínguez, L. and Aranaz, A. (2010) Molecular characterization of *Mycobacterium avium* subspecies *paratuberculosis* Types II and III isolates by a combination of MIRU-VNTR loci. *Vet Microbiol* **144**, 118–126.

Cernicchiaro, N., Wells, S.J., Janagama, H. and Sreevatsan, S. (2008) Influence of type of culture medium on characterization of *Mycobacterium avium* subsp. *paratuberculosis* subtypes. *J Clin Microbiol* **46**, 145–149.

Chevenet, F., Brun, C., Banuls, A.L., Jacq, B. and Christen, R. (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinform* **7**, 439.

Comas, I., Homolka, S., Niemann, S. and Gagneux, S. (2009) Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS One* **4**, e7815.

Donat, K., Hahn, N., Eisenberg, T., Schlez, K., Köhler, H., Wolter, W., Rohde, M., Pützschel, R. *et al.* (2016) Within-herd prevalence thresholds for the detection of *Mycobacterium avium* subspecies *paratuberculosis*-positive dairy herds using boot swabs and liquid manure samples. *Epidemiol Infect* **144**, 413–424.

Douarre, P.E., Cashman, W., Buckley, J., Coffey, A. and O'Mahony, J. (2011) Molecular characterization of *Mycobacterium avium* subsp. *paratuberculosis* using multi-locus short sequence repeat (MLSSR) and mycobacterial interspersed repetitive units-variable number tandem repeat (MIRU-VNTR) typing methods. *Vet Microbiol* **149**, 482–487.

Eisenberg, T., Wolter, W., Lenz, M., Schlez, K. and Zschock, M. (2013) Boot swabs to collect environmental samples from common locations in dairy herds for *Mycobacterium avium* ssp. *paratuberculosis* (MAP) detection. *J Dairy Res* **80**, 485–489.

Fawzy, A., Zschock, M., Ewers, C. and Eisenberg, T. (2016) New polymorphisms within the variable number tandem repeat (VNTR) 7 locus of *Mycobacterium avium* subsp. *paratuberculosis*. *Mol Cell Probes* **30**, 132–137.

Fernández-Silva, J.A., Abdulmawjood, A., Akineden, Ö., Dräger, K., Klawonn, W. and Bülte, M. (2012) Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis* at a regional scale in Germany. *Res Vet Sci* **93**, 776–782.

Fritsch, I., Luyven, G., Köhler, H., Lutz, W. and Möbius, P. (2012) Suspicion of *Mycobacterium avium* subsp. *paratuberculosis* transmission between cattle and wild-living red deer (*Cervus elaphus*) by multitarget genotyping. *Appl Environ Microbiol* **78**, 1132–1139.

Gerritsmann, H., Stalder, G.L., Spergser, J., Hoelzl, F., Deutz, A., Kuebber-Heiss, A., Walzer, C. and Smith, S. (2014) Multiple strain infections and high genotypic diversity among *Mycobacterium avium* subsp. *paratuberculosis* field isolates from diseased wild and domestic ruminant species in the Eastern Alpine region of Austria. *Infect Genet Evol* **21**, 244–251.

Grissa, I., Bouchon, P., Pourcel, C. and Vergnaud, G. (2008) On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing. *Biochimie* **90**, 660–668.

Grundmann, H., Hori, S. and Tanner, G. (2001) Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms. *J Clin Microbiol* **39**, 4190–4192.

Hahn, N., Failing, K., Eisenberg, T., Schlez, K., Zschöck, P.M., Donat, K., Einax, E. and Köhler, H. (2017) Evaluation of different diagnostic methods for the detection of *Mycobacterium avium* subsp. *paratuberculosis* in boot swabs and liquid manure samples. *BMC Vet Res* **13**, 259.

Hunter, P.R. and Gaston, M.A. (1988) Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* **26**, 2465–2466.

Keim, P., van Ert, M.N., Pearson, T., Vogler, A.J., Huynh, L.Y. and Wagner, D.M. (2004) Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. *Infect Genet Evol* **4**, 205–213.

Leao, C., Goldstone, R.J., Bryant, J., McLuckie, J., Inacio, J., Smith, D.G. and Stevenson, K. (2016) Novel single nucleotide polymorphism-based assay for genotyping *Mycobacterium avium* subsp. *paratuberculosis*. *J Clin Microbiol* **54**, 556–564.

Li, W., Raoult, D. and Fournier, P.E. (2009) Bacterial strain typing in the genomic era. *FEMS Microbiol Rev* **33**, 892–916.

Liverani, E., Scaioli, E., Cardamone, C., Dal Monte, P. and Belluzzi, A. (2014) *Mycobacterium avium* subspecies *paratuberculosis* in the etiology of Crohn's disease, cause or epiphenomenon? *World J Gastroenterol* **20**, 13060–13070.

Losinger, W.C. (2005) Economic impact of reduced milk production associated with Johne's disease on dairy operations in the USA. *J Dairy Res* **72**, 425–432.

Mamedov, T.G., Pienaar, E., Whitney, S.E., TerMaat, J.R., Carvill, G., Goliath, R., Subramanian, A. and Viljoen, H.J. (2008) A fundamental study of the PCR amplification of GC-rich DNA templates. *Comput Biol Chem* **32**, 452–457.

Möbius, P., Fritsch, I., Luyven, G., Hotzel, H. and Köhler, H. (2009) Unique genotypes of *Mycobacterium avium* subsp. *paratuberculosis* strains of Type III. *Vet Microbiol* **139**, 398–404.

Möbius, P., Luyven, G., Hotzel, H. and Köhler, H. (2008) High genetic diversity among *Mycobacterium avium* subsp. *paratuberculosis* strains from German cattle herds shown by combination of IS*900* restriction fragment length polymorphism analysis and mycobacterial interspersed repetitive unit-variable-number tandem-repeat typing. *J Clin Microbiol* **46**, 972–981.

Mokrousov, I. (2017) Revisiting the Hunter Gaston discriminatory index: note of caution and courses of change. *Tuberculosis* **104**, 20–23.

Oakey, J., Gavey, L., Singh, S.V., Platell, J. and Waltisbuhl, D. (2014) Variable-number tandem repeats genotyping used to aid and inform management strategies for a bovine Johne's disease incursion in tropical and subtropical Australia. *J Vet Diagn Invest* **26**, 651–657.

Octavia, S. and Lan, R. (2009) Multiple-locus variable-number tandem-repeat analysis of *Salmonella enterica* serovar Typhi. *J Clin Microbiol* **47**, 2369–2376.

Rónai, Z., Csivincsik, Á., Gyuranecz, M., Kreizinger, Z., Dán, Á. and Jánosi, S. (2015) Molecular analysis and MIRU-VNTR typing of *Mycobacterium avium* subsp. *paratuberculosis* strains from various sources. *J Appl Microbiol* **118**, 275–283.

Severiano, A., Carrico, J.A., Robinson, D.A., Ramirez, M. and Pinto, F.R. (2011) Evaluation of jackknife and bootstrap for defining confidence intervals for pairwise agreement measures. *PLoS One* **6**, e19539.

Sevilla, I., Li, L., Amonsin, A., Garrido, J.M., Geijo, M.V., Kapur, V. and Juste, R.A. (2008) Comparative analysis of *Mycobacterium avium* subsp. *paratuberculosis* isolates from cattle, sheep and goats by short sequence repeat and pulsed-field gel electrophoresis typing. *BMC Microbiol* **8**, 204.

Sohal, J.S., Arsenault, J., Labrecque, O., Fairbrother, J.-H., Roy, J.-P., Fecteau, G. and L'Homme, Y. (2014) Genetic structure of *Mycobacterium avium* subsp. *paratuberculosis* population in cattle herds in Quebec as revealed by using a combination of multilocus genomic analyses. *J Clin Microbiol* **52**, 2764–2775.

Stevenson, K., Alvarez, J., Bakker, D., Biet, F., de Juan, L., Denham, S., Dimareli, Z., Dohmann, K. *et al.* (2009) Occurrence of *Mycobacterium avium* subspecies *paratuberculosis* across host species and European countries with evidence for transmission between wildlife and domestic ruminants. *BMC Microbiol* **9**, 212.

Thibault, V.C., Grayon, M., Boschiroli, M.L., Hubbans, C., Overduin, P., Stevenson, K., Gutierrez, M.C., Supply, P. *et al.* (2007) New variable-number tandem-repeat markers for typing *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* strains: comparison with IS*900* and IS*1245* restriction fragment length polymorphism typing. *J Clin Microbiol* **45**, 2404–2410.

Thibault, V.C., Grayon, M., Boschiroli, M.L., Willery, E., Allix-Béguec, C., Stevenson, K., Biet, F. and Supply, P. (2008) Combined multilocus short-sequence-repeat and mycobacterial interspersed repetitive unit-variable-number tandem-repeat typing of *Mycobacterium avium* subsp. *paratuberculosis* isolates. *J Clin Microbiol* **46**, 4091–4094.

Van Hulzen, K.J.E., Heuven, H.C.M., Nielen, M., Hoeboer, J., Santema, W.J. and Koets, A.P. (2011) Different *Mycobacterium avium* subsp. *paratuberculosis* MIRU-VNTR patterns coexist within cattle herds. *Vet Microbiol* **148**, 419–424.

Zhou, K., Aertsen, A. and Michiels, C.W. (2014) The role of variable DNA tandem repeats in bacterial adaptation. *FEMS Microbiol Rev* **38**, 119–141.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Geographical distribution of the phylogenetic groups identified for MAP cultures in three German federal states, Hesse (A), Thuringia (B) and Saxony (C).

Figure S2. Application of the hierarchical typing approach (SNP + INMV) in detecting relationships between MAP cultures within the same phylogenetic groups in small geographic areas.

Table S1. MAP cultures/strains used in this work and respective genotyping results sorted by federal state.

Table S2. INMV types previously reported by Leao *et al.* (2016) and Bryant *et al.* (2016) for MAP type C isolates belonging to 12 SNP-defined phylogenetic groups.

Table S3. INMV types identified in this work and the order of SNPs investigated for each type.