Aus dem Institut für Insektenbiotechnologie
Professur für Angewandte Entomologie
der Justus-Liebig-Universität Gießen

How the integration of phylogenetics and venomics resolves persistent challenges in evolutionary systematics and toxinology

lessons from the spider kingdom

Dissertation
zur Erlangung des Doktorgrades (Dr. rer. nat.)
Im Fachbereich Agrarwissenschaften, Ökotrophologie und Umweltmanagement der Justus-Liebig-Universität Gießen

vorgelegt von

M.Sc. Tim Lüddecke
aus Salzgitter

Gießen, 2021

Mit Genehmigung des Fachbereichs Agrarwissenschaften, Ökotrophologie und Umweltmanagement der Justus-Liebig-Universität Gießen

Prüfungskommission

Vorsitz
Professor Dr. Marc F. Schetelig
Institut für Insektenbiotechnologie
Fachbereich Agrarwissenschaften, Ökotrophologie und Umweltmanagement
Justus-Liebig-Universität Gießen

1. Gutachter
Professor Dr. Andreas Vilcinskas
Institut für Insektenbiotechnologie
Fachbereich Agrarwissenschaften, Ökotrophologie und Umweltmanagement
Justus-Liebig-Universität Gießen

2. Gutachter
Professor Dr. Thomas Wilke
Institut für Tierökologie und Spezielle Zoologie
Fachbereich Biologie und Chemie
Justus-Liebig-Universität Gießen

3. Gutachter
Professor Dr. Steffen U. Pauls
Institut für Insektenbiotechnologie
Fachbereich Agrarwissenschaften, Ökotrophologie und Umweltmanagement
Justus-Liebig-Universität Gießen

1. Prüfer
Professor Dr. Günter Lochnit
Biochemisches Institut
Fachbereich Medizin
Justus-Liebig-Universität Gießen

2. Prüfer
Professor Dr. Adriaan Dorresteijn
Institut für Allgemeine Zoologie und Entwicklungsbiologie
Fachbereich Biologie und Chemie
Justus-Liebig-Universität Gießen

Datum der Disputation: 10. März 2021

Erklärung gemäß der Promotionsordnung des Fachbereichs 09 vom 07. Juli 2004 § 17 (2), geändert am 29. Mai 2019

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus- Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" niedergelegt sind, eingehalten.

_____ Gießen, 31.08.2020

Tim Lüddecke

Die vorliegende Arbeit wurde in den Jahren 2018 bis 2020 in der Abteilung Bioressourcen des Fraunhofer-Instituts für Molekularbiologie und Angewandte Ökologie (IME-BR) in Gießen unter der Leitung von Prof. Dr. Andreas Vilcinskas angefertigt. Die Promotion wurde in dieser Zeit durch das LOEWE-Zentrum für Translationale Biodiversitätsgenomik (LOEWE TBG) finanziell gefördert und war innerhalb des Zentrums dem Schwerpunkt Naturstoffgenomik zugeordnet.

Acknowledgments

Firstly, I sincerely want to thank my mentor and doctor father Prof. Dr. Andreas Vilcinskas for accepting me as his student and for providing me the means of pursuing a PhD under his guidance. In 2018 he offered me the opportunity to follow my lifelong passion for venomous creatures, despite that we never met before. For me, this act of blind trust was a rather life changing opportunity and ranks among the most generous things I´ve ever experienced. During my time in his group, Andreas constantly had an open ear for me and my concerns. He was always available when I needed him - even during the „special" times, such as: Society shut downs through a pandemic, institute wide IT shutdowns through hacker attacks, important project evaluations being on the horizon or all the random catastrophes linked to the new building phenomenon. I am grateful for the time I had in his group and acknowledge that his mentorship largely shaped me into the scientist I am today.

Furthermore, I need to thank two people who were instrumental for this thesis: Dr. Sarah Lemke successfully managed to supervise me and my project in the everyday business, may it be in the office or in the lab. For reasons that I have not been able to decipher, she remained calm and patient throughout my whole PhD and was never annoyed by my venom-monologues. On several occasions she would take custody for general tasks that I was in charge in order to create more opportunities for me doing science. It was her advice on many smaller and bigger issues along the way that provided the means for me successfully writing this work. Thanks Sarah! In addition to her, Dr. Björn M. Von Reumont was crucial for this work. His vast knowledge on basically all aspects of our field where I have been missing out, was a major contributor to this thesis. Moreover, the things I´ve learned from him about project planing, personal management and beer are essential for my future career (yeah, I know that I still have to focus ;D). Despite we had arguments and often pursue research by different strategies, I acknowledge his experience and character. I am really happy to have had him among my mentors and friends.

I am forever indebted to my previous supervisors Prof. Dr. Miguel Vences and Prof Dr. Sebastian Steinfartz who, during my times as a naive bachelor´s student in Braunschweig, took me under their guidance and kickstarted with me a project on amphibian toxins. These experiences and their teaching provided the baseline of me starting a PhD at all. I am equally grateful to a selected pair of people who I´ve met during my time in Braunschweig and whose friendship and knowledge still remains as a powerful source of joy and inspiritaion to date. Firstly, I want to mention the worlds leading authority for office-neighbourhood and Steinfartz-wrangling: Kathleen Preißler. Thanks a lot for our time in the salamander office, your help on understanding aposematism and for our productive brainstormings. Oh yeah, and thanks for supporting me during my stupid idea of collecting Komodo dragon saliva… Next, I want to thank the girl that saves the salamanders: Vanessa Schulz. Thank you so much for party, spider-hunting-hollidays and for not giving up in making me a bird-enthusiast. At some point I will still convince you on doing a PhD!

Apart from these persons I owe a lot to all the current and former members of the Animal Venomics Group who also supported me whenever help was needed and who contributed a lot to the success of this work. That said, I am grateful to Dr. Andre´ Billion, Dr. Anne Paas, Angela Paul, Dr. Frank Förster, Gesa Schmidt, Dr. Henrike Schmidtberg, Dr. Ivan Koludarov, Dr. Kim N. Kirchhoff, Dr. Lea Talmann, Dr. Miray Tonk, Rabia Özbek, Dr. Stephan Drukewitz and Tanja Berghöfer. From other groups of the LOEWE-TBG and the Justus Liebig University I want to thank Prof. Dr. Günther Lochnit, Dr. Thomas Timm and Dr. Stefan Prost for providing me machines, material and teaching for different experiments. I sincerely thank Martin Reinartz and Wolfgang Dibiasi for providing me with their gorgeous photographies used to assemble spider related figures herein.

Beyond fellow scientists working with me, I was very lucky in having people in my private life who constantly kept me motivated and supported me on my pursuit for knowledge. I am aware that my constant chatting about various creatures, their biology and their toxins can be rather disturbing. However, these folks patiently accepted me for who I am and generously offered their friendship. For whatever reason, some of these are already present in my life for almost 30 years. Thank You Frederik, Christiane, Tobias, Michi, Rene, Marius and Nik! In addition, I´ve met several people during my time in

Gießen who became rather close and helped me to survive the bad times I certainly had at some point. Katharina, Philipp, Silke, Anne, Lea, Kathy Bäm, Miriam and Maren - I am grateful to have met you guys! Most of all I want to thank my family. Having me in your tribe comes with rather unusual life situations. Firstly for large fractions of my existence I messed up our house with multitudes of creatures in all shapes and sizes, many of which were life threatening toxic. Secondly, as a child I proudly announced, that I will become a scientist studying venomous animals and since then never stopped chatting about these. Lastly, I actually followed on this naive announcement and quitted my well payed industry job and instead, on purpose, moved into a financially insecure and health threatening existence as a toxinologist. Despite this difficult to digest life planing of mine and the abundant consequences you had (sorry for the crickets…) none of you challenged my sanity and, instead, you´ve always been supporting me - no matter what happened. Thank you for everything! Finally, I have to mention my personal little chaos - Katharina. During the time of writing and while thinking, working and pondering about my work, it was you who made me smile. It was you who brought my attention back to the real world when I was about to get lost. Whether it was the hacker attack or the Corona pandemic, you were there and brought light into my frequently emerging inner dark. Without you, these lines would probably never have been written. Thank you.

Summary

Spiders represent one of the most successful branches of metazoan life. Throughout their long-lasting evolutionary trajectory, spiders diversified into almost 50,000 species. They conquered all continents except antarctica and established themselves as predators in virtually all ecosystems. The invention of venom systems, that are present in all but one spider lineages, contributed significantly to their evolutionary success. Albeit research on spiders, referred to as Arachnology, is an old field of study, it is hampered by a variety of persistent challenges awaiting scientific resolution. A subset of four such challenges, relating to evolutionary systematics and toxinology, are of pivotal importance. First, the taxonomic status of many spiders, in particular within the mygalomorph infraorder, and their phylogenetics remains largely ambiguous. Secondly, knowledge on spider venoms is so far fully derived from selected taxa and biased towards the few medically significant or exceptionally large species. Third, the sheer diversity of spiders makes it rather difficult to select promising focal taxa for venom bioprospecting studies. Lastly, knowledge upon the evolutionary forces driving spider venom evolution remains in its infancy. Addressing these important issues via phylogenetic and venomic approaches is the scope of this work.

Systematic ambiguity is addressed by using tarantulas (Theraphosidae) as a model group. In two experimental setups, a molecular phylogenetic study utilizing six sequenced genes plus a phylogenomic study on ca. 2,000 genes, the first phylogenetic trees for Theraphosidae are constructed. These recovered monophyly of Theraphosidae as a whole and supported validity of formerly questionable subfamilies Poecilotheriinae, Psalmopoeinae and Stromatopelminae. It clarifies the position of *Brachionopus* and *Harpactirella* and argues for paraphyly of Schismatothelinae. In a trait evolution analysis, this work finds that defensive hairs likely evolved convergently within neotropical tarantulas. To make bioprospecting studies more efficient, this work developed a phylogeny-driven strategy for rational taxon selection in biodiscovery, exemplified on the proposed tarantula phylogeny. Applying this strategy towards the whole spider kingdom recovered the family Araneidae as especially promising focal group. Consequently, the wasp spider *Argiope bruennichi* as a member of this family is subsequently studied. A morphological analysis of its venom apparatus found, that gland and chelicerae mirror structures present in the few other studied spider venom apparatuses. However, the venom duct that connects fang and venom gland was found to be substructured into four distinct units, displaying a previously hidden complexity within spider venom systems. A Venomic analysis revealed, that the wasp spider venom is rather simply composed and that CAP proteins dominates the venom profile. As other spider venoms are mostly composed of small neurotoxic peptides, the venom of *A. bruennichi* is considered as arachno-atypical. This work proposes an evolutionary scenario, in which an economic dilemma between the venom system and the silk system during hunting led to the loss of venom components in the wasp spider. Lastly, a selection of novel biomolecules that mirror insect-neuropeptides are identified within the wasp spider venom, highlighting the underestimated importance of neuropeptides as evolutionary starting points for the birth of toxic components.

This work contributes to the field of Arachnology as it significantly advances the *status quo* within the four selected challenges in evolutionary systematics and toxinology through synthesis of phylogenetics and venomics. It clarifies the taxonomic placement of several spider lineages and proposes the first well supported hypothesis upon tarantula evolution. A novel approach towards a rational taxon selection is developed and explored. As a consequence, the study of an araneid venom expanded the general understanding of spider venoms and the architecture of their venom apparatus beyond the taxonomic bias. The underestimated importance of larger proteins versus small neurotoxic peptides is emphasized and the role of neuropeptides in venom evolution is supported. The role of negative selection in spider venom evolution is discussed in perspective to loss of toxicity in defensive hair-bearing tarantulas and the economic dilemma between both weapon systems in *A. bruennichi*. This work thus contemplates novel insights and concepts towards the four persistent challenges and provides an experimentally supported framework on which future systematic-, evolutionary-, bioprospective- and general venomic works can be informed upon.

How the integration of phylogenetics and venomics resolves persistent challenges in evolutionary systematics and toxinology

lessons from the spider kingdom

By Tim Lüddecke

"There are four million different kinds of animals and plants in the world. That's four million solutions to the problem of staying alive."

-Sir David Attenborough

To Katharina

Table of contents

I. Introduction

Biodiversity and evolutionary relationships of spiders and their kin

Spiders (Araneae) belong to Arachnida, a class that includes the orders scorpions (Scorpiones), camel spiders (Solifugidae), whip scorpions (Uropygii), harvestmen (Opiliones), whip spiders (Amblypygii), pseudoscorpions (Pseudoscorpiones), and mites (Acari), along with hooded tick spiders (Ricinulei) and palpigrades (Palpigradi). Together with horseshoe crabs (Xiphosura) and sea spiders (Pycnogonida), Arachnida forms the subphylum Chelicerata within the phylum Arthropoda (Sharma, 2018).

Extant spiders are divided into the three infraorders: Mesothelae, Mygalomorphae, and Araneomorphae (Fig. 1). The most ancestral infraorder is represented by the monotypic Mesothelae, which harbors only the family Liphistidae: segmented trapdoor spiders comprising 135 species from Asia (World Spider Catalog, 2019). With ca. 3,000 valid species, Mygalomorphae accounts for a much higher percentage of global spider diversity (World Spider Catalog, 2019). Although mygalomorphs occur globally, most species described are from the tropics and subtropics (World Spider Catalog, 2019). Several prominent spider families belong to Mygalomorphae, such as tarantulas (Theraphosidae), trapdoor spiders (Ctenizidae), and funnel-web spiders (Atracidae). Finally, the most derived and diverse spider infraorder is represented by Araneomorphae (World Spider Catalog, 2019). Among others, the group harbors charismatic orb-weaver spiders (Araneidae), wolf spiders (Lycosidae), and jumping spiders (Salticidae). Araneomorphs display an unprecedented diversity of ecological specializations and have undergone a multitude of radiations culminating in an array of highly biodiverse families (World Spider Catalog, 2019).

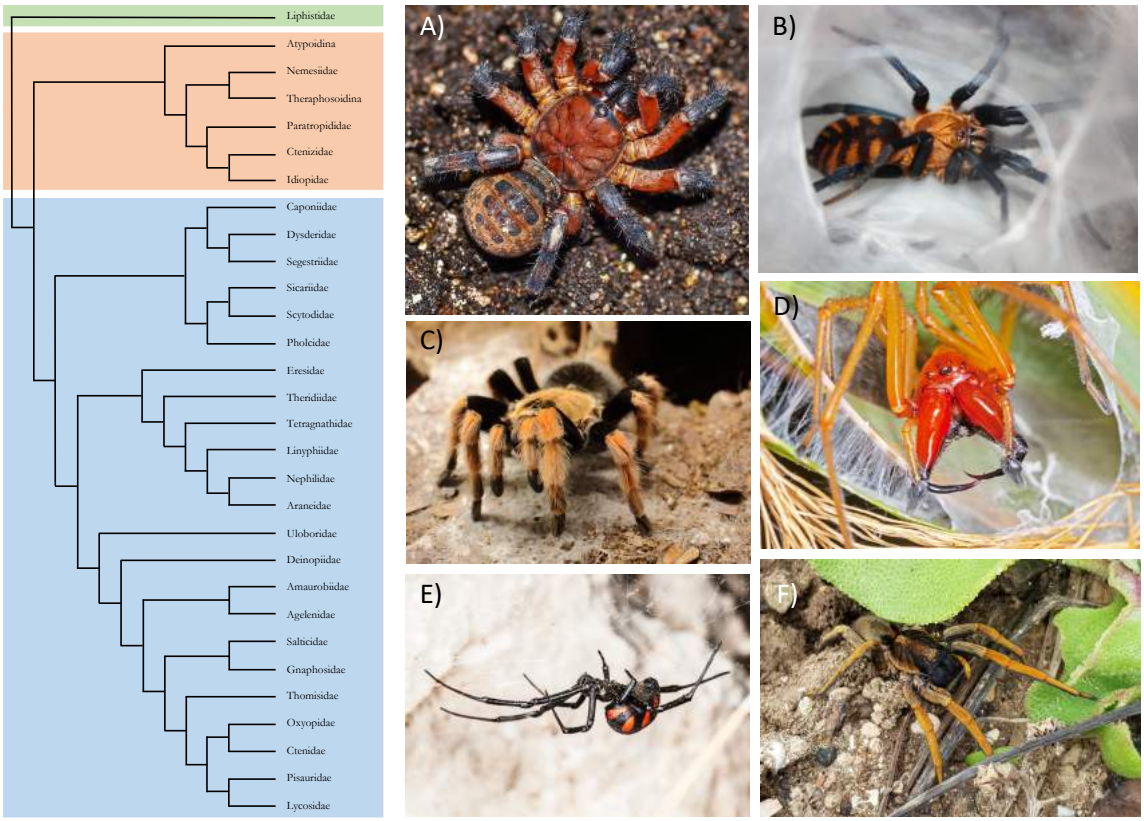


Fig. 1: Phylogenetic relationships between spider infraorders and illustration of spider diversity. The simplified cladogram, based on Garrison et al. (2016), depicts the relationships between Mesothelae (green), Mygalomorphae (red) and Araneomorphae (blue). Given are selected families of each infraorder. The righthand side illustrates representative taxa of some important lineages. Mesothelae: A) *Liphistius yangae* (Liphistidae). Mygalomorphae: B) *Linothele fallax* (Dipluridae), C) *Brachypelma boehmei* (Theraphosidae). Araneomorphae: D) *Cheiracanthium punctorium* (Eutichuridae), E) *Latrodectus tredecimguttatus* (Theridiidae) and F) *Hogna schmitzi* (Lycosidae). Images courtesy of: A), F) M. Reinartz; B), C) T. Lüddecke and D), E) W. Dibiasi.

The fossil record of spiders dates back to the Carboniferous (Dunlop et al., 2015; Selden & Penney, 2010). Since then, the overall spider body plan has remained largely unaltered. Today, spiders inhabit virtually all ecosystems and have successfully conquered all continents except Antarctica (Piel, 2018; World Spider Catalog, 2019). Their vast distribution and occurrence are paralleled in their tremendous diversification. In total, the order comprises 48,424 extant species in 120 families, but it has been estimated that a total of 90,000 species will be discovered. Thus, only a fraction of the global arachnofauna has been described (Pennisi, 2017; World Spider Catalog, 2019). Such biodiversity is almost unprecedented in the animal kingdom, and only the realm of insects is more prolific in terms of diversification (Engel, 2015).

The biology and ecology of spiders

One of the possible explanations for the outstanding evolutionary success of spiders is their biological organization. They display a conserved body plan that is shared between all taxa in all infraorders. It is partitioned into two tagmata, namely the prosoma and opisthosoma. Both are covered by an exoskeleton consisting of chitin. Therefore, spiders perform periodic ecdysis to facilitate growth (Foelix, 1983; Nentwig, 2013). The prosoma carries numerous appendages. First, it has four pairs of legs symmetrically dispersed around it. Second, the pedipalps, an additional pair of leg-like structures is localized to the anterior prosoma in proximity to the oral cavity. These serve a variety of functions, and often play a role during reproduction for the male specimen (Calbacho-Rosa et al., 2013; Cargnelutti et al., 2018; Foelix, 1983; Mahmoudi et al., 2008). Lastly, the prosoma carries a pair of chelicerae covering the oral cavity. In spiders, unlike all other chelicerates, these are modified into fangs and harbor a glandular system that produces a venom, which is released from an opening close to the tip of each. The opisthosoma houses most of a spider's organs, including book lungs for respiration and large parts of the digestive and vascular systems (Foelix, 1983). Posttereoventrally located are the spinnerets, the major components of the silk spinning apparatus. Lastly, the opisthosoma contains the reproductive system of females.

Silk is an omnipresent trait for spiders, and all taxa feature a functional silk apparatus that is used in a multifunctional manner (Eisoldt et al., 2011; Gosline et al., 1986; Vollrath, 1999). Silk is plesiotypically applied for the construction and stabilization of burrows and trapdoors in Mesothelae and Mygalomorphae. However, apomorphic silk functionality evolved in Araneomorphae, and is often implemented for the construction of complex foraging webs (Foelix, 1983; Harmer et al., 2011). Across all spiders, a myriad of silk types with specific functions and properties emerged (Vollrath, 1999). All silk types are composed of repetitive protein elements that are hyphenated and compose a macromolecular protein fiber (Vollrath, 1999).

Principally, spiders are predatory. Most taxa feed on a variety of invertebrates, mostly insects. Therefore, spiders contribute to maintaining the equilibration of insect populations in several ecosystems and occupy an ecological niche of pivotal importance (Foelix, 1983). As most spiders are general predators that prey on a diversity of species, it is noteworthy that several groups evolved high degrees of trophic specialization, such as myrmecophagy (*Zodarion* sp.), oniscophagy (*Dysdera* sp.), lepidopterophagy (e.g. *Mastophora* sp.), or even occasional herbivory (Clark et al., 2000; Forster, 1977; Nyffeler et al., 2016; Pekár, 2004; Pekár & Toft, 2015; Řezáč et al., 2008; Yeargan, 1988). All spiders perform extraintestinal digestion and feed on enzymatically pre-liquefied prey items (Foelix, 1983; Nentwig, 2013).

Independently from a taxonomic assignment, ecological niche, or trophic specialization, spiders evolved a remarkable array of traits that enable their unparalleled diversity and abundance. For instance, rather complex hunting, mating, and defensive behaviors are distributed through the spider tree of life (Clark et al., 2000; Forster, 1977; Riechert & Singer, 1995; Welke & Schneider, 2012). However, the most outstanding development in spiders is the widespread implementation of a versatile toolbox of biomolecules that enables them to prevail throughout the ongoing struggle of survival. This biochemical toolbox is composed of two components, each including a plethora of different molecules and reflecting extraordinary complexity: the above-discussed silk and venom.

The biological role of spider venom

Venoms independently evolved in several animal lineages, and are present in each phylum of the animal kingdom (Casewell et al., 2013). They are defined as secretions produced in specialized glands of an

animal that are injected into another animal through the infliction of a wound, leading to the disruption of viable physiological processes in the victim (Fry et al., 2009). This physiological disruption is facilitated by bioactive molecules that compose the venom and are referred to as toxins. Most venom toxins are proteins and peptides that evolved from normal physiological isoforms via different processes, often following gene duplications and subsequent neo- and/or subfunctionalization ("weaponization") (Fry et al., 2009). Other processes such as gene co-option and *de novo* evolution have also been claimed to have produced toxin proteins and peptides (Casewell et al., 2011; Casewell, 2017; Drukewitz et al., 2019). Originally, it was proposed that venoms serve the three biological functions of predation, defense, and sexual competition (Fry et al., 2009). However, this functional trinity was expanded in the recent past and additional biological functions were assigned to venom, including immune system function and applications in reproduction and digestion (Schendel et al., 2019).

Spiders employ venoms for two predominant functions: predation and defense. During predation, venoms are essential to overpowering prey. Upon capturing its prey, a spider will utilize its chelicerae to bite the victim and to inject its venom. After envenomation, the toxic components rapidly immobilize the prey and allow the spider to feed (Nentwig, 2013). If applied defensively, a spider will deliver a bite to its potential predator. The negative effects caused by the envenomation may either enhance its chance to escape, lead to an abortion of the attack, or at least trigger a learning behavior that leads to the future avoidance of similar-looking prey by the predator. The latter is often enhanced via aposematic warning colorations or distinct defensive behaviors that flag the spider's toxicity towards a predator (Pekár, 2014).

Components and biochemistry of spider venom

All spiders – with the exception of Uloboridae – feature a functional venom system. Spiders are hence recognized as one of the most successful groups of venomous animals (Weng et al., 2006; Saez et al., 2010). Spider venoms are outstandingly complex, and a single species can harbor up to 1,000 different components in its venom, which is unparalleled within the animal kingdom (Herzig, 2019). It has been estimated that a total of 10 million different compounds could be discovered in spider venoms (Saez et al., 2010). Components essentially fall into four categories: larger proteins, cysteine-rich peptides, small organic molecules, and antimicrobial linear peptides.

Small organic molecules from spider venom are mostly acylpolyamines. These are composed of aromatic acyl groups linked to polyamine backbones and, sometimes, are extended to harbor amino acid moieties (Jasys et al., 1990; Langenegger et al., 2019; Tzouros et al., 2013; Wilson et al., 2017). Toxins of this class are insecticidal and may have antimicrobial activities (Langenegger et al., 2019). Likewise, antimicrobial linear peptides from spider venom have antimicrobial activities, but they also interfere with eukaryotic cells by disrupting the integrity of their cell membranes (Corzo et al., 2002; Garcia et al., 2013; Yan & Adams, 1998).

The importance and abundance of large proteins in spider venom has recently been under discussion (Langenegger et al., 2019). In a few spiders, these are key components of their venoms. In particular, black widow spiders (*Latrodectus* sp.) have severe toxicity to humans that is derived from alpha Latrotoxin (LTX), a homotetrameric protein of 130 kDa that forms pores in presynaptic neuronal membranes of vertebrates. This leads to an uncontrolled flux of Ca^{2+} and neurotransmitters resulting in nociception, convulsions, and sometimes death (Grishin, 1998; Henkel & Sankaranarayanan, 1999; Orlova et al., 2000; Ushkaryov et al., 2008). The venom of recluse spiders (*Loxosceles* sp.) and related Sicariidae contains phospholipase D (PLD), a sphingomyelin hydrolyzing enzyme of 50 kDa that is responsible for the rapid cytotoxicity of recluse spider venom (Swanson & Vetter, 2006). In both black widows and recluse spiders, large proteins are essential parts of the venom and have effects during envenomation. Apart from these, a cytotoxic hyaluronidase-like enzyme has recently been isolated that enhanced the uptake of co-administered neurotoxins, thus acting as a spreading factor (Biner et al., 2015).

Except these few examples, the role of larger proteins in spider venoms is ambiguous. Members of the cysteine-rich secretory protein/antigen 5/pathogenesis-related 1 (CAP) and neprilysin metalloproteases protein families have been identified in many spider venoms, but their biological role has not yet been illuminated (Kuhn-Nentwig et al., 2019; Langenegger et al., 2019; Undheim et al., 2013; Zobel-Thropp et al., 2019). Additionally, disulfide isomerases, carboxypeptidases, and serine proteases

were recently discovered in spider venom. It was proposed that these may facilitate the maturation and post-translational modification of toxins (Langenegger et al., 2019). While chymotrypsin-like activity and a role in toxin maturation could be experimentally determined for one of these serine proteases, the biological roles of the other proteins remain questionable (Langenegger et al., 2018).

The last group of spider venom components is cysteine-rich peptides. These are rather small polypeptides, typically with molecular masses below 10 kDa, that are rich in disulfide bonds derived from cysteines. Many of them are thought to interact with ion channels and receptors, thus representing the principal neurotoxic component of spider venoms (Langenegger et al., 2019). This group comprises different protein families such as Kunitz-type serine protease inhibitors, helical arthropod neuropeptide derived peptides (HAND), colipase fold (MIT-1) peptides, disulfide-directed beta-hairpin fold (DDH) peptides, and, most importantly, peptides with an inhibitor cysteine knot (ICK) scaffold (Langenegger et al., 2019). While the first few mostly represent understudied components, ICK peptides are the most diverse, abundant, and well-studied group within spider venom systems (Langenegger et al., 2019). The secondary structure of ICK peptides is a triple-stranded antiparallel beta-sheet, and its tertiary structure is determined by at least six cysteines. After oxidation, these form disulfide-bonds with each other and lead to a characteristic pseudo-knot motif (Buczek et al., 2007; Cardoso & Lewis, 2019; Langenegger et al., 2019; Norton & Pallaghy, 1998). Most ICK peptides feature six cysteine residues and thus constitute three disulfide bridges. However, several derivatives with expanded cysteine scaffolds, as well as members with additional ICK motifs (double ICK, or dICK) were discovered recently (Chassagnon et al., 2017; Escoubas et al., 2003; McCarthy et al., 2015; Pineda et al., 2014). Inhibitor cysteine knot peptides are not restricted to spiders, and are present in other venomous animals as well (Drukewitz et al., 2018; Fry et al., 2009; Özbek et al., 2019; Von Reumont et al., 2014). However, spiders are the most prolific source of ICK peptides, as one species usually harbors dozens of different ICKs, and spiders exceed all other taxa for ICK diversity by far (Langenegger et al., 2019).

Pharmacology and translational potential of spider venom toxins

Inhibitor cysteine knot peptides in spider venom are neurotoxins that interfere with ion channels and receptors. They form stable complexes and perturb the normal biochemical mode of action in these vitally important targets, often by inhibiting their activation, delaying their deactivation, or shifting their potential limits (Langenegger et al., 2019). Key targets affected by spider venom neurotoxins are voltage-gated sodium, voltage-gated potassium, and voltage-gated calcium channels, but targets also include acid-sensing ion channels (ASIC), glutamate receptors, and transient receptor potential channels (TRP) (Langenegger et al., 2019). Given that such targets are of pivotal importance for signal transduction and cellular communication, their functional disruption disturbs the physiological homeostasis of the intoxicated organism.

The binding of a neurotoxic ICK is commonly facilitated indirectly via a toxin's partial penetration of a membrane and subsequent lateral migration towards the target (Deplazes et al., 2016). In some cases, a transient trimeric complex of the membrane, toxin, and the target can be employed (Agwa et al., 2017). The utilization of foreign biomembranes for the binding of neurotoxins to their target is energetically efficient, as large parts of the process's binding energy deviates from membrane adhesion (Lee & MacKinnon, 2004). The proximity of venom proteins to membranes was shown to induce structural changes and support three-dimensional orientation prior to target binding (Mihailescu et al., 2014; Ryu et al., 2017). Along their evolutionary trajectory, ICK peptides acquired outstanding biological performance. Their binding to a respective target is facilitated with unprecedented specificity, often on the level of ion channel subtypes. Similarly, the displayed bioactivity is of significant efficiency, as ICK peptides already exhibit their effects at concentrations that lie one order of magnitude below those afforded from other, less specific components (Langenegger et al., 2019). Lastly, their physicochemical properties relate to powerful pharmacodynamics. As molecular size negatively correlates with distribution time, small ICKs exert their physiological effects quickly post-injection. In addition, the ICK motif provides these peptides with outstanding stability against proteolytic degradation, thus maximizing the toxin's biological half-life (Pineda et al., 2014).

The promising bioactivities, pharmacodynamics, diversity, and physicochemical properties that spider venom toxins display are why research efforts were previously made to study them as potential

bioresources (Pineda et al., 2014). Many such bioprospecting programs looking at spider venom were successful, and a variety of novel biomolecules with valuable bioactivities were discovered and are currently under closer investigation concerning potential translational applications. For instance, several ICK peptides that modulate sodium channels were shown to represent promising lead structures for subsequent development into novel analgetics (Park et al., 2008; Pineda et al., 2014; Saez et al., 2010). A spider toxin has also recently been used to rescue Dravet-syndrome mice from seizures and premature death (Richards et al., 2018). The most striking example, however, was derived from *Hadronyche infensa*, an Australian funnel-web spider. A dICK peptide isolated from its venom was successfully used to protect mice from neuronal damage following ischemic stroke, even when administered hours later (Chassagnon et al., 2017). Apart from therapeutics, spider venom is considered a prolific source for eco-friendly bioinsecticides, with great potential for plant protection in agriculture (Herzig et al., 2014; King, 2019; King & Hardy, 2013; Saez & Herzig, 2019; Windley et al., 2012). Antimicrobial peptides from spider venoms could potentially yield novel tools for the ongoing battle against drug-resistant prokaryotes (Samy et al., 2017). Lastly, the neglected array of proteins displaying enzymatic activities may be harvested as innovative enzymes for the production and degradation of industrial goods (Fig. 2).

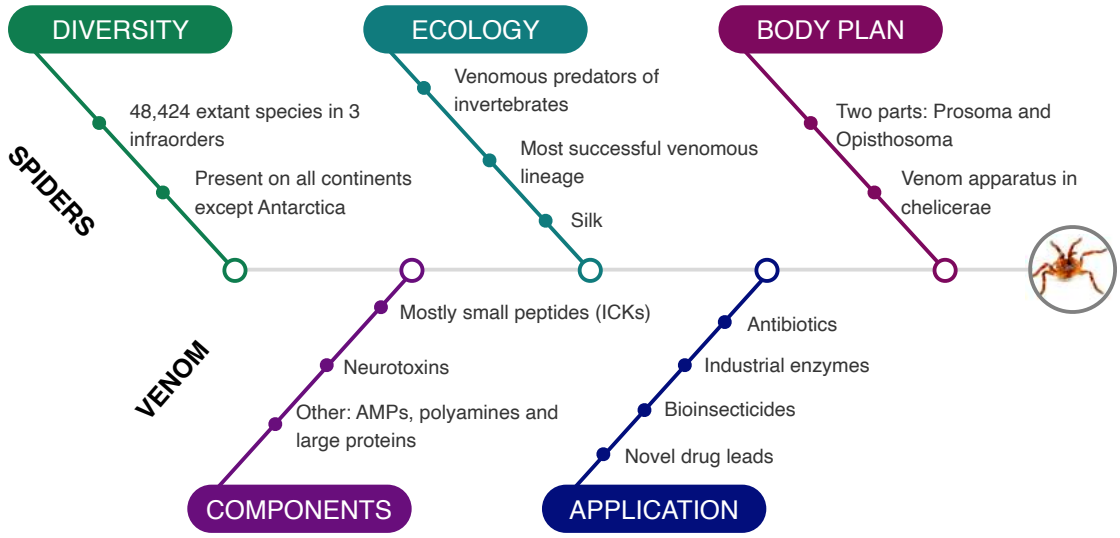


Fig. 2: Spiders in a nutshell. Summary of major points regarding spiders (top row) versus their venom (bottom row). Summarised are spider diversity, ecology, and body plan together with accounts on main venom components and their potential in applied research as further discussed in the text above.

.

Persistent challenges in Arachnology

As described above, arachnological research already has a long and successful history. Several aspects of spider biology have already been scientifically addressed and demystified. However, a large number of questions regarding spider biology is still unanswered and awaiting scientific exploration. In particular, four areas related to spider evolutionary systematics and toxinology are of relevance. These are outlined below and henceforth referred to as (1) systematic ambiguity, (2) the taxonomic bias in spider venom research, (3) the taxon selection dilemma, and (4) the venom evolution conundrum. Along with a discussion of each of these four challenges, this work develops a series of working hypotheses that will be tested and serve as an informative baseline.

Spider systematics is largely ambiguous

During their ca. 300 million years of evolution, spiders evolved unique phenotypic and ethological adaptions that contribute to their survival (Pennisi, 2017; Selden & Penney, 2010). It is a longstanding aim of scientists to elucidate the diversity of global arachnofauna, expand its species inventory, and disentangle the evolutionary processes behind spider radiations and trait acquisitions.

Most studies have emphasized morphological and ethological data to infer systematic relationships. However, in Mygalomorphae and, in particular, Theraphosidae (tarantulas), such morphological characters are often structurally conserved between lineages, thus limiting the amount of information that could be extracted. Moreover, on multiple occasions, it has been demonstrated that in the realm of theraphosids, morphological characters are affected by high degrees of homoplasy (Ortiz et al., 2018). Hence, it is likely that the combination of uninformativeness plus homoplasy in frequently used characters explains the prevailing systematic ambiguity in Theraphosidae. An alternative to such morphology-driven systematic and evolutionary studies is represented by a molecular workflow that sequences genetic material and uses molecular information instead. The implementation of such approaches on theraphosids, however, has been a rare exception and has prior exclusively been performed for studies on the genera *Aphonopelma* and *Brachypelma* (Hamilton et al., 2011; Turner et al., 2018; Wilson et al., 2013). As a likely result of the large neglect of molecular data, the inner systematics of Theraphosidae are still questionable. On one hand, intrafamilial relationships among major lineages have been studied by several authors, but the results of these works were largely contradictory, and no consensus has been established so far (Pérez-Miles et al., 1996; Raven, 1985, 1990; Wilson et al., 2013). On the other hand, morphologically informed alpha- and beta taxonomy of Theraphosidae remains mostly incongruent, and the taxonomic boundaries in and around the family are not well understood (Bertani et al., 2012; Hamilton et al., 2011; Hendrixson et al., 2015; Wilson et al., 2013). As a consequence, recurrent taxonomic and nomenclatural changes are frequent in tarantulas (Pérez-Miles et al., 1996; Fukushima et al., 2005; Guadanucci, 2007; Guadanucci & Gallon, 2008; Guadanucci & Wendt, 2014; Hamilton et al., 2016; Mendoza & Francke, 2017; Nagahama et al., 2009; West et al., 2008).

Another mygalomorph family that illustrates the prevailing problem of systematic ambiguity is the Australian funnel-web spiders of the genera *Atrax* and *Hadronyche*. Commonly acknowledged as the world's most venomous spiders, these have drawn significant scientific interest due to their medical relevance. Based on morphology, they were placed within Hexathelidae in close relationship with genera such as *Macrothele*, *Porrhothele,* and *Illawara* (Hedin et al., 2018). This long-standing hypothesis was recently reconsolidated by phylogenomic approaches, which recovered a paraphyly of Hexathelidae. Consequently, three new families (Atracidae, Macrothelidae, and Porrhothelidae) were erected to yield all non-hexathelid taxa and to reestablish the monophyly of Hexathelidae (Hedin et al., 2018). In this context, *Atrax* and *Hadronyche* were transferred to Atracidae, thus demonstrating that even focal taxa are not reliably assigned. Although such examples of systematic ambiguity are most frequent in Mygalomorphae, they occur in all of the three spider infraorders (Pennisi, 2017).

Systematics represents the fundamental pillar for organismic research. Hence, such ambiguities cause direct repercussions in other branches of biology. For example, it is more difficult to expand the existing arachnological species inventory without a solid taxonomic framework in place, as new taxa need to be allocated within. Moreover, systematic ambiguity indirectly threatens the conservation of species, as the taxonomic system is the groundwork upon which extinction risks and conservation needs are predicted. This is an essential problem for some theraphosid spiders, which are currently threatened by habitat loss and fragmentation as well as by illegal wildlife trade (Fukushima et al., 2019; Hendrixson et al., 2015; Mendoza & Francke, 2017; Turner et al., 2018). The last repercussion is the potential impact that systematic ambiguity can have on the medication of envenomated entities. Recurrent nomenclatural changes of medically relevant taxa cause disturbances in choosing proper therapy for envenomated patients, as the literature accumulates inconsistent names for species. Henceforth, physicians' selection of an optimal treatment may take more time, thus threatening a patient's health.

Taking these negative implications into account, it is clear that the prevailing systematic ambiguity in Arachnology urgently needs to be resolved.

Taxonomic bias in spider venom research

The great diversity of spiders, currently comprising 48,424 species in 4,160 genera assigned to 120 families, is increasing annually. Despite problems imposed by systematic ambiguity, on average, 800 new species are described every year (World Spider Catalog, 2019). Unfortunately, this expansion in knowledge regarding biodiversity is not reflected in venom research.

As of today, spiders whose venom has been studied belong to only 28 different families. The other 90 families have been fully omitted from venom research (Jungo et al., 2010; Pineda et al., 2018). Metrics for venomically assessed spiders decrease even more on lower taxonomic levels. At the time of writing, venoms from 61 species of Mygalomorphae plus 66 Araneomorphae had been investigated, reflecting only ca. 1% of all mygalomorphs and 0.1% of all araneomorph taxa (Herzig et al., 2019). Moreover, no representatives of Mesothelae have been studied, and the discovery that Mesothelae harbor a functional venom system at all is only a recent development (Foelix & Erb, 2010). In total, only a marginal fraction of ca. 0.3% of extant spiders have been studied for their venom. The vast majority of taxa are a toxinological black box (Herzig et al., 2019). The taxon sampling in spider venom research is rather disproportionate, and the majority of studied taxa fall into one of two categories: spiders that are either of clinical relevance or of extraordinary size (Herzig et al., 2019; Kuhn-Nentwig et al., 2011).

Clinically relevant spiders encompass a comprehensible subset of families, specifically Atracidae, Ctenidae, Paratropididae, Sicariidae, and Theridiidae (Hauke & Herzig, 2017). Thus, potentially dangerous species are exceptional, and account for only 0.5% of spider biodiversity (Hauke & Herzig, 2017). However, as these are of great medical concern in some countries, the necessity of studying their venoms is emphasized by public health demands. The tendency to study exceptional large spiders is explained by several factors. Spider size correlates with venom yield and ease of collection (Herzig et al., 2019). Research on spider venom has mostly been driven by pharmacology. In those studies, crude venoms were fractionated via liquid chromatography and subsequently investigated for bioactivity and structure (Herzig et al., 2019). Such workflows require large quantities of sample material, often on the milligram level. Given that this demand usually far exceeds the venom yield of small species, working with them is significantly more difficult. Moreover, the crude venom of spiders is usually obtained via electrical stimulation of the chelicerae. In small species, this approach fails regularly due to the petite venom system (Herzig et al., 2019). As an alternative, venom glands can be dissected for venom collection, thus sacrificing the animals (Herzig et al., 2019). This requires a high number of collected specimens and is therefore of ethical concern. Lastly, large spiders are kept and reproduced by arachno-enthusiasts worldwide (Jäger, 2003). The private collections of hobbyists often encompass hundreds of specimens that are readily available for venom collection, thus circumventing laborious fieldwork. In particular, Theraphosidae, which harbor virtually all of the largest spider taxa known and are the predominantly kept as pets, have been frequently studied for their venoms. As a consequence, they currently account for a third of all spider venoms studied (Herzig et al., 2019).

Previous research on spider venom suffers from a taxonomic bias consisting of two components. A first, anthropocentric component derives from medical significance and pet trade availability. The second, and methodological, component is fed by instrumental constraints. The result of this taxonomic bias towards larger and dangerous species is that today, available knowledge on spider venom is only inferred from a small, non-representative fraction of its total taxonomic breadth. The majority of families remain fully neglected in venom research, although these harbor most of the ecologically specialized and hyperdiverse lineages (Herzig et al., 2019). Given this fractionation of knowledge on venoms throughout the spider tree of life, it is questionable to what extent current assumptions about spider venom biology reflect the biological truth. Understanding the biology of spider venom systems in their totality relies upon a holistic understanding. It is therefore an important task to include as many neglected taxa as possible in detailed venom surveys. Only that will allow us to resolve the taxonomic bias in spider venom research, to make meaningful general inferences upon spider venom biology, and to understand spider venom systems beyond taxonomically imposed constraints.

The taxon selection dilemma

At present, the pharmaceutical industry is facing unparalleled economic challenges, and the imminent collapse of its current system has been predicted (Lindgardt et al., 2008). The main contributor to this situation is the subsiding number of truly innovative therapeutics that are developed and approved (Paul et al., 2010). Spider venom represents a naturally occurring chemical library that accommodates diversified yet unexplored pharmacopeias (Herzig, 2019; Pineda et al., 2014; Saez et al., 2010). Consequently, harvesting this bioresource will likely reveal novel leads for the development of in-demand

therapeutics. However, the diversity that spiders and spider venoms deploy represents a dilemma for bioprospecting programs.

Taxon selection is a fundamental step in bioprospecting, as it largely determines the potential of a project. Only molecules of the respective organisms will be identified, and this identification is time and cost-intensive. Ergo, insufficient taxon selection can drastically reduce the potency and success rate of a bioprospecting program. Despite its importance, the means by which a taxon selection for spider venom bioprospecting should be performed is rather questionable. The current approach obeying to taxonomic bias (sensu Herzig et al., 2019) is not applicable to making strategic decisions regarding taxon selection, as it will neglect large swaths of overall spider diversity. The dilemma in spider venom bioprospecting has, therefore, two components: first, the great diversity deployed by spiders and their venom components highlights them as prime sources for novel bioressources. Second, the same factors drastically complicate a rational taxon selection for bioprospecting programs.

Prevention of the anticipated crisis within the pharmaceutical industry will largely depend on the swift discovery of novel biomolecules (Paul et al., 2010). Patients suffering from diseases that could potentially be treated by such also depend on their near-term discovery. Therefore, rapid bioprospecting – of which reliable taxon selection is a major component – is demanded by producers and consumers in the pharmaceutical sector. Reliable taxon selection is further critical in the face of the continuing global biodiversity loss related to the prevailing sixth mass extinction (Barnosky et al., 2011). The conservation status of most spiders has not been assessed, but several representatives are already considered threatened (Fukushima et al., 2019; Turner et al., 2018). As largely insectivorous predators, their prosperity is interdependent with abundances of sympatric insects. As insects are declining worldwide (Hallmann et al., 2017; Sánchez-Bayo & Wyckhuys, 2019), it is of considerable concern that their decease may trigger the extinction of spiders. If this happens, for venom-wise unstudied species, their valuable venom components will never be studied and a whole library of bioresources will be lost forever. From a bioprospecting perspective, it is thus of utmost importance to apply reliable taxon selection to enhance spider venom biodiscovery.

The importance of swift drug discovery for industrial applications as well as for patients, plus the threat of losing valuable bioresources through extinction, highlights the importance of addressing the taxon selection dilemma. Innovative strategies for a rational and optimized selection of taxa for spider venom research must be developed.

The venom evolution conundrum

Venom systems are key innovations and promote integral biological functions for each species that features them (Schendel et al., 2019). Their prevalence throughout the animal kingdom has led to an array of eco-evolutionary studies (Fry et al., 2006; Fry et al., 2008; Hargreaves et al., 2014; Sanggaard et al., 2014; Warren et al., 2008). Questions asked in this context are mostly related to the origin of venom systems as well as the ecological factors driving their evolution.

Although a variety of different venomous taxa have been included to these studies, the largest proportion of knowledge on venom evolution has been derived from snakes. Here, the origin of the venom system as a transition from salivary glands to venom glands and their interconnection with apotypic dentition was discovered (Fry et al., 2006; Fry, Scheib, van der Weerd, Young, McNaughtan, Ryan Ramjan, et al., 2008; Hargreaves et al., 2014; Vonk et al., 2008). Further, the close connection of venom phenotypes and trophic niches was of interest. In this framework, venom has been discovered as a trait with high degrees of intra-specific plasticity, and dietary transitions between life-history stages and between sexes were frequently observed (Alape-Girón et al., 2008; Amazonas et al., 2018; Casewell et al., 2014; Chippaux et al., 1991). Moreover, trophic specialization in snakes was shown to reduce or defunctionalize venom systems (Daltry et al., 1996; Li et al., 2005). Beyond species-level studies, the underlying machinery behind venom genes concerning evolutionary processes on the genomic level attracted much attention. The illuminated proteins involved in vital metabolic and regulatory processes are the substrate giving rise to toxins, based on different mechanisms such as duplication, co-option, or domain losses (Casewell, 2017; Casewell et al., 2011; Fry et al., 2008; Kordiš & Gubenšek, 2000; Vonk et al., 2013; Wong & Belov, 2012). Based on the convergent character of venoms, it is possible that some insights derived from snakes can be applied to other venom systems as well. However, the expansion of

studies beyond the scope of snakes has indicated that venom evolution and ecology are much more complex than previously thought (Aminetzach et al., 2009; Drukewitz et al., 2019; Harris & Arbuckle, 2016; Jenner et al., 2019; Koludarov et al., 2017; Sanggaard et al., 2014). Each venom system is shaped by unique eco-evolutionary dynamics, and thus generalizations about the fundamental processes at work are difficult to propose.

In contrast to snakes, many aspects of spider venom evolution are poorly understood, although great effort has also been made to study spiders' venom. For instance, knowledge on the morphology and functionality of venom delivery systems is minuscule for most spider lineages, and only for a small fraction are profound examinations available (Dos Santos et al., 2000; Garb, 2014; Silva et al., 2008; Yiğit et al., 2004). The morphology of venom systems across animal groups has increasingly been investigated, and adjacent studies have revealed evolutionary constraints and functional implications for venoms caused by venom system morphology (Schendel et al., 2019). Because of data unavailability, insights on the connectivity between morphology, functionality, and evolutionary constraints are currently hampered in spiders. It is thus important to examine more spider venom systems to derive insights about this interplay.

Moreover, the processes underlying venom evolution in spiders are largely unknown. A comprehensive understanding of venom molecules throughout the spider kingdom is also missing, and it is challenging to infer the preferred evolutionary substrates that may give rise to toxins in spiders. That said, from available studies, it was deduced that lateral gene transfer is among the mechanisms involved in spider toxin evolution. For instance, PLDs have recently been traced back to their evolutionary origin (Cordes & Binford, 2018). This work demonstrated that PLDs are present in different organismic classes, but only acquired a venom function in sicariid spiders (Cordes & Binford, 2018). Further, PLD are found to derive from a single proteobacterial ancestor and, from there, seemingly radiated into other organismic groups, at least partially facilitated via horizontal gene transfer (Cordes & Binford, 2018). Apart from PLD, many studies have been performed on LTX in black widows (*Latrodectus* sp.) and related, yet medically insignificant Theridiiade (*Steatoda* sp. and *Parasteatoda* sp.). An evolutionary trajectory largely shaped by purifying selection was revealed for LTX (Garb & Hayashi, 2013). However, LTX acquired dissimilar properties and sequence patterns in different species: members of *Steatoda* and *Parasteatoda*, which are not toxic to vertebrates, differ drastically sequence-wise from their counterparts in *Latrodectus* that are of severe vertebrate toxicity (Garb & Hayashi, 2013). Thus, it was proposed that the functional shift towards vertebrate toxicity coincided with the evolutionary split from black widows off other Theridiidae (Garb & Hayashi, 2013). As with PLD, an origin via horizontal gene transfer has been proposed for LTX, mostly governed by the recently published genome of *Parasteatoda tepidariorum* (Gendreau et al., 2017; Schwager et al., 2017). The genome illuminated the role of gene duplications in the evolution of LTX as those substantially duplicated in *Latrodectus* compared to relatives (Gendreau et al., 2017; Schwager et al., 2017). The consecutive increase of available arachnid genomes (Garb et al., 2018), in particular those of *Acanthoscurria geniculata* and *Stegodyphus mimosarum*, further strengthened the importance of gene duplication events for spider venom evolution (Sanggaard et al., 2014). Beyond the scope of PLD and LTX, knowledge of evolutionary processes on spider toxins is limited. Lynx spider (*Oxyopes takobius*) ICKs were found to evolve under purifying selection and to be likely descendants of spiderine toxins, although the exact process of their evolution remains questionable (Sachkova et al., 2014).

The venom system is a critical innovation in spiders and a key contributor for their evolutionary success and diversification (Nentwig, 2013). Ergo, it is of critical importance to derive a holistic understanding of its morphological and molecular evolution to gain detailed insights into the processes that shape the diversity and abundance of extant spiders. Therefore, the study of venom systems across the spider tree of life needs to be expanded urgently.

The persistent problems are interconnected

Scientists working on subfields linked to these four challenges have barely scratched the tip of the iceberg regarding the full spectrum of discoveries to be made (Fig. 3). Each of the above-outlined factors represents an individual problem that hampers Arachnology. They have a high degree of interconnectivity but pose multi-pronged hurdles across the field.

Systematic ambiguity and taxonomic bias interfere with the venom evolution conundrum. The first reason is the absence of a well-supported phylogenetic system for most spiders, as such is the prerequisite for making meaningful evolutionary inferences about deployed venom systems or other traits. The neglect of most venoms has resulted in a multitude of blind spots regarding venom components throughout the spider tree of life. It for this reason that no general patterns of evolution within spider venoms can be deduced so far. Likewise, the taxon selection dilemma is partially a result of taxonomic bias, but is, again, linked to systematic ambiguity. Without a reliable systematic framework, it is difficult to make rational decisions for venom bioprospecting. A major component that highlights the importance of this problem is the ongoing biodiversity loss. The need for enhanced venom bioprospecting that allows the exploitation of venom before species extinction is among the key components within this problem. Since international species conservation relies on systematic data as a foundation to assess conservation status, systematic ambiguity threatens conservation assessments and protective means in spiders. As a consequence, rapid venom bioprospecting in spiders becomes even more critical.

Fig. 3: The four prevailing challenges in Arachnology summarised. Systematic ambiguity, the taxonomic bias in spider venom research, the taxon selection dilemma, and the venom evolution conundrum represent individual- but also interconnected challenges to the field of arachnology. As a result, for many subdisciplines that are related to these challenges scientists just scratched the tip of the iceberg of awaiting discoveries.

The problems of systematic ambiguity, the taxonomic bias in spider venom research, the taxon selection dilemma, and the venom evolution conundrum each represent considerable impediments. Separately, as well as synergistically, they pose major challenges to fundamental and applied spider research. Resolving this Gordian knot would significantly contribute to the science of Arachnology.

Aim of this work and working hypotheses

This work is devoted to the four outlined challenges in arachnology and aims to advance their underlying status quo. Accordingly, it intends to answer a selection of questions related to evolutionary systematics and toxinology throughout the spider tree of life. It is designated to unravel evolutionary relationships between unresolved taxa and to support taxonomic stability within problematic groups. Moreover, it aims to advance bioprospecting from spider venoms and to optimize its processes by identifying a rationale on which taxon selections can be based. Lastly, this work contemplates expanding the existing knowledge on spider venom systems beyond the scope of taxa previously studied as a result of taxonomic bias. For this purpose, a series of six working hypotheses are proposed below, which then are tested experimentally throughout the chapters of this thesis.

Hypothesis 1: Molecular phylogenetics is in advantage over morphological analyses for systematic inference in spiders. In spiders, and in particular within Mygalomorphae, morphological homoplasy, the limited characteristics that are available, and their minuscule informativeness all affect the ability to make systematic inferences. Molecular phylogenetic approaches, on the other hand, circumvent these factors, as these technologies generate orders of magnitude more datapoints and can be bioinformatically normalized against a variety of effects that cause a bias. Thus, they may have the means to reconstruct the first reliable phylogenies for problematic spider lineages. Therefore, these methodologies represent a promising and powerful alternative to the traditionally used approaches that may at least augment or, very likely, fully replace morphology as the principal source of systematic data in spiders.

Hypothesis 2: Reconstruction of trait evolution and systematic placements are supported by molecular phylogenetics. As morphology-based studies previously failed to determine the systematic placement of several spiders, a direct consequence is that the placement of many taxa is ambiguous. An opportunity of the first molecular phylogenetic trees will be, that the current systematics of spiders can be tested against this evolutionary framework. It is to be expected that molecular phylogenetically informed systematics will resolve large amounts of the taxonomic ambiguity in spiders and determine the systematic placement of several questionable groups. Moreover, an available phylogenetic system could serve as a baseline to study the evolution of important biological traits of relevant taxa.

Hypothesis 3: Bioprospecting from spider venom can be optimized via rational taxon selection. The focus on dangerous and large spiders for venom bioprospecting enabled the discovery of powerful drug candidates but has been a slow and cost intensive process. Finding a rational criterion upon which taxon selection for bioprospecting in spiders can be based, may allow for streamlined, more efficient biodiscovery.

Hypothesis 4: Studies beyond the taxonomic bias will challenge current assumptions on spider venom systems. The study of venom systems from large and dangerous spiders led to the current assumption that spider venoms are highly complex and dominated by small ICK peptides, with larger proteins being only minor components with limited biological importance. As these studies are based on a non-representative minority of the total spider diversity, this current picture on spider venoms should be challenged. It appears reasonable that the subsequent study of spider venom systems beyond the previously selected taxa may reveal different venom profiles based on other evolutionary innovations than ICK peptides and may eventually favor larger proteins within the venom system.

Hypothesis 5: Venom system morphology is functionally underestimated in spiders. For most spiders, the cellular architecture of their centralized venom system remains unexplored. As in most animals, the venom system of spiders is thought of as a simple injector connected to a venom producing gland via a duct. However, in centipedes, cone snails, and snakes, it has been shown that venom systems can impose major evolutionary constraints on the venom, and that these systems can be physiologically rather complex when investigated in more detail. Hence, a similar underestimated complexity in function and organization is equally likely for spiders.

Hypothesis 6: The study of more venom systems will illuminate novel evolutionary substrates and mechanisms for venom components. Previous work on spider venom systems addressed the evolutionary origins of ICKs, PLD, and LTX, but the evolutionary substages for other venom components remains unexplored. Moreover, the evolutionary processes in spider venom systems are only poorly understood in comparison to other venomous animals. The expansion of venomic data beyond the few previously studied venom systems will likely lead to the exploration of novel toxic proteins with independent evolutionary trajectory and thus advance our understanding of which genetic blueprints are weaponized in spider venoms and which mechanisms are involved in this process.

II. Thesis overview

Chapters I-II. The taxonomic integrity and evolutionary relationships of many spiders are poorly understood. In particular, Mygalomorphae are rendered as systematically ambiguous. Theraphosidae, commonly referred to as tarantulas, comprises the most prominent and species-rich family within the infraorder. Despite their widespread recognition and long research history, their intrafamilial relationships have not been disentangled, as previous studies have delivered contradicting results and no consensus has been reached for most included taxa. Unsurprisingly, many such taxa were consistently questioned for their taxonomic validity and their monophyly. As is commonplace in arachnological research, theraphosid systematics relied almost exclusively on morphological data, and the utilization of molecular informed analysis is a rare exception. However, recent studies revealed that, in particular, the systematic ambiguity of theraphosids is reasoned by high degrees of morphological homoplasy.

Chapter I infers the first molecular phylogeny of Theraphosidae based on 3,500 base pairs of genetic information derived from six genes. It shows that most of the traditionally recognized subfamilies are indeed monophyletic groups, and the previously questionable subfamilies Poecilotheriinae, Psalmopoeinae, and Stromatopelminae represent valid taxonomic entities. Moreover, this chapter recovered the paraphyly of Avicullariinae and Schismatothelinae and clarified the placement of *Brachionopus* and *Harpactirella*, two taxa formerly assigned to Barychelidae, as members of Theraphosidae. Lastly, it finds that statistical support was commonly absent from deep supra-generic clades, whereas shallow clades consistently received high support. Research conducted in this chapter contributes to the battle against systematic ambiguity in mygalomorph spiders, as it proposes the first molecular-based hypothesis on tarantula evolutionary systematics. It circumvents the traditional morphological workflow and delivers, for the first time, a reliable answer to some of the most critical questions within the field. However, as it is unable to resolve deep relationships of Theraphosidae, this research likewise explores the limitations of molecular phylogenetics based only on a few genes. It indicates a demand for more data-intensive approaches as a means to unravel such relationships with certainty. Addressing this demand and taking another major leap against mygalomorph systematic ambiguity are the major issues of Chapter II. Here, a phylogenomics approach is applied to illuminate the evolution and systematic relationships within Theraphosidae.

In Chapter II, a core ortholog approach based on 2,460 genes from transcriptomes of 33 taxa is used to infer a first phylogenomically informed hypothesis on tarantula evolutionary systematics. It recovers high statistical support for shallow as well as deeper clades. Thus, Chapter II recovers the first reliable backbone phylogeny for theraphosids. In agreement with findings in Chapter I, Chapter II recovers the validity of Poecilotheriinae, Psalmopoeinae, and Stromatopelminae as well as the paraphyly of Avicullariinae and Schismatothelinae. The study furthermore shows that the paraphyletic subfamily Ischnocolinae is composed of at least two genetic lineages. Chapter II, moreover, introduces a previously unrecognized young clade of neotropical subfamilies that comprises ca. 60% of all described theraphosid species, and thus proposes that Neotropic tarantulas were subject to a relatively recent rapid diversification. A defensive system composed of varying types of urticating setae was found to be exclusive to members of this clade. The likelihood of different evolutionary scenarios behind these setae is compared, and their evolution is determined as most likely driven by convergence. Lastly, the study suggests that the evolution of urticating setae may have direct repercussions with defensive components in tarantula venoms, and that urticating setae represent a key innovation for neotropical tarantulas and facilitated their rapid diversification.

Chapter II highlights the importance of phylogenomics to infer deep relationships within Theraphosidae and illustrates the first well-supported evolutionary hypothesis for the family. Moreover, this chapter underlines the power of this technology to deduce systematic validity and to disentangle paraphyletic groups that were previously difficult to study. Lastly, Chapter II explores how such phylogenies can be used for subsequent studies in spider trait evolution and thus contribute to the holistic understanding of spiders in an eco-evolutionary context. Like Chapter I, the research in Chapter II represents a major advance in the battle against systematic ambiguity in spiders, as it answers a plethora of open questions on tarantulas and tarantula trait evolution.

By using Theraphosidae as a model, Chapter I and Chapter II underline how molecular phylogenetic approaches can be deployed as a powerful tool for studies in spider evolutionary systematics. They provide a framework upon which subsequent studies in the field can be informed.

Chapter III. The extent of potentially identifiable biomolecules and the diversity of species, represents a prevailing challenge for bioprospecting in spiders. It depicts the taxon selection dilemma of spider venoms as being (a) on one hand a promising and rich source for novel therapeutics and (b) on the other hand displaying a species hyperdiversity that makes it tremendously difficult to exploit this resource efficiently. So far, taxon selection in spider venom bioprospecting has been driven by taxonomic bias, and thereby has only discovered less than one percent of all expected venom biomolecules. Accordingly, this approach is deemed insufficient to deduce a representative view on spider venoms. Finding a rationale on which such bioprospecting can instead be based upon would represent a critical advance, eventually solving the taxon selection dilemma. Developing innovative strategies to overcome this dilemma is the subject of this chapter.

Chapter III, again using Theraphosidae as a model group, combines data mining from bioinformatics resources with available phylogenetic data from Chapter I and Chapter II. It found that, despite Theraphosidae being the most studied spider group in terms of venom, available venomic data is restricted to a small subset of lineages. Therefore, Chapter III recovers a new level of taxonomic bias within a single spider family. Based on the phylogenetic placement and considering toxin data coverage throughout the phylogeny, this chapter deduced priority groups for future venom bioprospecting in tarantulas. Chapter III proposes that a taxon selection based on phylogenetics and big data mining may represent a promising approach to tackling the taxon selection dilemma.

Chapter III shows that by reducing taxonomic ambiguity in spiders (Chapter I and Chapter II), preexisting insights can be accompanied by data mining approaches to illuminate the status quo in bioprospecting, which can be used to translate these findings into novel strategies for biodiscovery. The strategy that it proposes, however, needs to be evaluated on a larger scale.

Chapters IV-V. Knowledge of spider venoms, their components, and their evolution is rather limited beyond species investigated in the context of taxonomic bias. By applying the bioprospecting strategy developed in Chapter III, Chapter IV and Chapter V aim to verify its usefulness and to expand knowledge on spider venom systems in understudied groups. When the phylogeny-guided approach for taxon selection is applied to spiders in their totality, a series of ecologically relevant, phylogenetic distant, and venom-wise understudied groups are revealed. Among these, the family Araneidae holds an outstanding position, as they comprise one of the most diverse spider lineages whose members are frequent model systems for the evolution of silk genes. This multi-pronged significance renders Araneidae a promising group for venom bioprospecting, and consequentially their venom systems are explored in Chapter IV and Chapter V. The wasp spider *Argiope bruennichi* was selected as a representative species as it is an established model organism for a variety of ecological questions, and thus elucidating its venom system will directly contribute to answering other research questions. For *A. bruennichi*, no venom components have been previously described, and the morphology of its venom apparatus has not been explored in much detail.

Therefore, as a foundation, Chapter IV illuminates the morphological organization behind the venom apparatus in *A. bruennichi* via microscopy. It verified that in its outer morphology, the venom apparatus conformed to the basic structure, and used the jackknife motion that is typical for araneomorph spiders. Cheliceral teeth are present, in which the fangs rest when not in use, and which may support prey handling by the wasp spider. The investigation further revealed that *A. bruennichi* harbors relatively large venom glands that reach from the chelicerae into the prosoma. Lastly, this chapter depicts the cellular organization throughout a wasp spider's venom apparatus. The most interesting observation of Chapter III is the underestimated complexity of the venom duct, which can be divided into three subregions that differ in cellular complexity.

Chapter V investigates the venom composition of *A. bruennichi* via a proteo-transcriptomics-based venomics workflow based on two mass spectrometry technologies for proteomics. Enabled by insights from Chapter IV, venom gland dissection was conducted as a means to collect venom from *A.*

bruennichi, as electrostimulation mostly fails in such small spiders. This chapter shows that proteo-transcriptomic workflows are powerful tools for exploring the venoms of small neglected invertebrates. Furthermore, with the araneid *A. bruennichi* as a model, Chapter V contributes to the reduction of taxonomic bias in spider venom research. It shows that the venom of wasp spiders is comprised of several components that are typical for spider venoms, such as ICK peptides, MIT atracotoxins, neprilysin metalloproteases, and HAND peptides. More importantly, Chapter V reveals the presence of several novel venom components belonging to different protein families, some of which have never before been described from spider venoms. By analyzing these novel compounds for their sequences and relationships with known proteins, the study found that many of these represent apotypic neuropeptides or hormones recruited to the venom gland of *A. bruennichi*. This chapter also finds that the venom of wasp spiders is of an arachno-atypical composition, as it is of astonishing simplicity and dominated by CAP proteins instead of small neurotoxic peptides. Therefore, Chapter V discusses the biological role of CAPs and, based on sequence homology with functionally assessed CAPs, predicts a proteolytic activity within the venom of *A. bruennichi*. Lastly, it proposes an evolutionary hypothesis that may explain the observed arachno-atypical venom. Here, the venom system metabolically competes with the unique silk apparatus-based hunting behavior of wasp spiders, which leads to an energetic dilemma. Chapter V argues that in *A. bruennichi,* this dilemma was solved in favor of the silk apparatus, and most venom components fell victim to purifying selection along the species evolutionary trajectory.

Chapter IV and Chapter V explore for the first time the venom system of wasp spiders. Chapter IV depicts the morphological organization of the venom apparatus in *A. bruennichi* and thus delivers critical insights into the architecture and functionality of these spiders. Chapter V illuminates the biochemical diversity of wasp spider venom. It illustrates the composition, identifies novel components, and reveals the arachno-atypical nature of wasp spider venom. Moreover, it proposes hypotheses that may explain the chapter's findings. In tandem, the results of Chapter IV and Chapter V support a phylogeny-based taxon selection as proposed in chapter III and argue that it may indeed represent a powerful approach to tackling the taxon selection dilemma. Moreover, by adding data from the venom of an understudied lineage, research in these chapters contributes to the resolution of taxonomic bias in spider venom research. Both chapters feature a variety of empirical and theoretical advancements on the understanding of araneid venom systems, in particular in an eco-evolutionary context. Thus, the work in Chapter IV and Chapter V helps to disentangle the venom evolution conundrum

III Discussion

Molecular approaches are a powerful tool in spider systematics

The mygalomorph family Theraphosidae represents an excellent model group for questions regarding systematic ambiguity, as this phenomenon is especially prevalent in this infraorder. Among Mygalomorphae, they represent the most diverse family, and members are present on all continents except Antarctica (Piel, 2018; World Spider Catalog, 2019). Their size and charismatic appearance have led to their widespread recognition and sparked a long-lasting pursuit to disentangle their evolutionary history, as well as to decipher their species inventory. This pursuit was pioneered by Robert Raven, who postulated the first hypothesis about intrafamilial relationships within Theraphosidae (Raven, 1985). Since Raven's initial work, a series of follow-up studies from a variety of authors have been conducted. All of these studies relied fully on morphological characters but ranged in scope, and addressed intrafamilial, intrasubfamilial, or intrageneric relationships (Pérez-Miles et al., 1996; Guadanucci, 2014; Hamilton et al., 2011; Samm & Schmidt, 2010; West et al., 2008). Regardless of scope, the resulting phylogenies consistently contained polytomies and nodes of insignificant statistical support. Accordingly, previous studies were unable to resolve relationships for older or for younger taxonomic levels, and the inferred evolutionary hypotheses on Theraphosidae were fully contradictory (Guadanucci & Wendt, 2014; Samm & Schmidt, 2010; Schmidt, 1995; Valencia-Cuéllar et al., 2019; West et al., 2008). Robert Raven later summarized the prevailing systematic ambiguity in Theraphosidae and the difficulties in disentangling their relationships by referring to the family as a taxonomic nightmare (Raven, 1990).

Hypothesis 1 proposed that molecular phylogenetics is an advantageous means to overcome taxonomic ambiguity in mygalomorph spiders, of which the few characteristics available, the uninformativeness of these characteristics, and homoplastic effects are leading causes. By sequencing genetic material, large quantities of informative datapoints can usually be generated. Through selecting orthologous genes, effects such as homoplasy can be excluded from molecular approaches, and a variety of other sources of bias can be controlled for bioinformatically. The results of Chapter I and II agree with the proposed hypothesis, highlighted by the fact that their utilization of molecular phylogenetics was able to derive the first robust backbone phylogeny for Theraphosidae. In particular, the phylogenomics approach of Chapter II is a promising tool because it recovered even the deepest theraphosid branches with high support values. The tremendous potential of such techniques is also reflected in the increasing number of taxonomic studies on Theraphosidae that have utilized molecular approaches since (Hamilton et al., 2016; Hüsser, 2018; Ngamniyom et al., 2014; Ortiz et al., 2018; Ortiz & Francke, 2016, 2017). Some of these works are directly based on the research conducted in this work (see comment within Ortiz et al., 2018) and expand its scope towards shallower taxonomic levels. A result of these studies is that, for the first time in theraphosid systematics, obtained phylogenies are in agreement to each other even though they are derived from independent researchers and independent samples.

Facing this recent rise of molecular approaches and the linked decrease of taxonomically ambiguous taxa in Theraphosidae, it can strongly be presumed that the "taxonomic nightmare" (*sensu* Raven, 1990) will be resolved soon. However, future research on theraphosid evolutionary systematics still needs to overcome some hurdles. For instance, some particularly import taxa of the Ischnocolinae and Selenogyrinae subfamilies could not be collected for the herein presented work. These critical groups are notoriously difficult to access as they live in some extremely remote and dangerous geographical areas. Therefore, collecting such samples will be a rather time-consuming and laborious task for the future. Interestingly, the rise of molecular phylogenetics and phylogenomics as a means to tackle systematic ambiguity in Theraphosidae have shifted the problem from methodological issues towards taxon sampling. It was previously problematic to resolve theraphosid relationships, but this task is now easier to perform. In contrast, obtaining sufficient material for molecular analyses from such hard-to-access groups is now the challenge that the battle against systematic ambiguity in Theraphosidae is facing.

Resolving the systematic status of questionable theraphosids

Linked to the previous difficulties in the inference of an evolutionary framework within Theraphosidae is the uncertain status of many included taxa. Boundaries within and around Theraphosidae have not

been successfully determined and several taxa are regularly transferred between genera and subfamilies. Hypothesis 2 stated that advancements in the field of phylogenetic research via molecular approaches will provide an informative baseline upon which systematic assessments can be informed. It predicts that this framework will enable the clarification of previously questionable taxa. The research in Chapters I and II tested this hypothesis using problematic theraphosids.

Of particular interest were the subfamilies Psalmopoeinae, Poecilotheriinae, and Stromatopelminae. All of these groups exclusively harbor species that follow an arboreal lifestyle. However, they frequent different geographical areas. While Psalmopoeinae occur in the neotropics, Poecilotheriinae are endemic to the Indian subcontinent and Stromatopelminae are native to mainland Africa (Schmidt, 2003; Teyssie, 2015; World Spider Catalog, 2019). In the past, the validity of all three subfamilies was questioned, and, more often than not, they were placed in other subfamilies. For instance, Psalmopoeinae and Poecilotheriinae were sometimes considered as members of Selenocosmiinae (Marshall et al., 1999; Raven, 1985). Others placed Psalmopoeinae together with Stromatopelminae in Aviculariinae – another group of arboreal neotropical theraphosids (Fukushima & Bertani, 2017). The results of Chapters I and II refute these previous assignments and support the independence of Psalmopoeinae, Poecilotheriinae, and Stromatopelminae, thus claiming validity for these taxa. Stromatopelminae were recovered as sister to Harpactirinae, forming a clade composed of African lineages. The proposed relationship of Poecilotheriinae with Selenocosmiinae is not supported by the results in both chapters. Instead, a placement as a sister group to the Asian Ornithoctoninae is recovered with high confidence. Consistently throughout all analyses, Psalmopoeinae are recovered as the sister group to Schismatothelinae, and these, together, as sisters to Avicularinae. An association of Psalmopoeinae with Poecilotheriinae, Selenocosmiinae, or Stromatopelminae is not supported. Recently, Hüsser (2018) was able to reproduce the findings concerning Psalmopoeinae. Hüsser officially re-erected the subfamily based on his findings and the research presented therein. Comprehensive studies to follow the recommended reestablishment of Poecilotheriinae, likewise to Psalmopoeinae, are currently underway (Meyer, pers. Comm).

Besides these, the Ischnocolinae subfamily has been a critical group in the past. Originally erected to comprise all Theraphosidae that could not be assigned to other subfamilies, Ischnocolinae represents a paraphyletic taxonomic "trash bin." Until recently, they were defined upon the absence of informative morphological characters (Schmidt, 2003). Considered the most problematic group within Theraphosidae, they were frequent subjects to systematic studies (Guadanucci, 2007; Guadanucci & Gallon, 2008; Guadanucci, 2014; Guadanucci, 2005; Guadanucci & Wendt, 2014). However, these works were only moderately successful: While some taxa were transferred to an independent subfamily (Schismatothelinae), other ischnocolines have not yet been successfully resolved. The research presented in Chapter I and Chapter II constitutes the most reliable hypothesis proposed for their systematic status. These studies support that Ischnocolinae is composed of at least two independent lineages. Of these, the first is a sister to Eumenophorinae, as one of the most ancestral theraphosid groups, while the other is related to the neotropical subfamilies and is of relatively recent origin. A potential third lineage was recovered that places *Nesiergus insulanus* in close relationship with Selenocosmiinae. However, this finding is exclusively based on a single gene obtained from an exuvia, and the relationship of *N. insulanus* to Selenocosmiinae is only marginally supported. As *N. insulanus* is endemic to the Seychelles archipelago and thus sample material of this genus is scarce, this data was evaluated as too valuable to omit and hence included to Chapter I (Canning et al., 2014). That said, this placement and thus the existence of a potential third lischnocoline lineage needs to be tested in the future.

Another important yet systematically unclarified group of Theraphosidae is a selection of miniaturized taxa from Africa. These, comprising the genera *Brachionopus* and *Harpactirella,* were critically discussed for their placement within Theraphosidae, and some authors placed them instead in Barychelidae (brushed-foot trapdoor spiders) (Raven, 1985; Schmidt, 2002). However, the results of Chapter I reject that these genera are related to barychelids. The molecular analysis clearly recovered both as members of the Harpactirinae subfamily, placed sister to Stromatopelminae in a clade fully composed of taxa from Africa.

Facing the above outlined implications for tarantula systematics made by this work, findings throughout this thesis largely conform to hypothesis 2. Research on tarantula systematics has been unable

to determine the independent characters of Psalmopoeinae, Poecilotheriinae, and Stromatopelminae in the past. Likewise, the status of *Brachionopus* and *Harpactirella* as well as the cryptic genetic diversity of Ischnocolinae have never been reliably recovered before. Only through the molecular phylogenetic framework erected in Chapter I and Chapter II has it been possible to address these important questions and to considerably advance the field of tarantula systematics.

Evolution of urticating setae in Theraphosidae and implications for venom evolution

Investigating the evolutionary processes and patterns in spiders, in particular under consideration of the venom evolution conundrum, was another substantial goal of this research. In this realm, hypothesis 2 proposed that the phylogenetic framework that enabled the clarification of the systematic status of questionable groups could subsequently guide the study of relevant traits in each particular group. For example, resolving the phylogeny of squamates and combining it with morphological data on glandular systems and proteomics illuminated the evolution of reptilian venom systems (Casewell et al., 2012; Fry et al., 2012). Likewise, established phylogenies enabled the detailed study of varanid venom systems and the interconnected evolution of spitting, hooding, and aposematism with cytotoxicity in cobras (Koludarov et al., 2017; Panagides et al., 2017). Unfortunately, the spider kingdom lacks reliable phylogenetic insights for most lineages and thus comparable studies cannot be performed with certainty.

At least for Theraphosidae, this has significantly changed. Within this study, the validity of hypothesis 2 was tested on this spider family and on one of its most characteristic traits relating to anti-predatory defense: urticating setae. These are present in neotropical subfamilies (Aviculariinae, Psalmopoeinae, and Theraphosinae) and are mostly dispersed across the opisthosoma (Bertani & Guadanucci, 2013; Kaderka et al., 2019; Pérez-Miles & Perafán, 2015). Several functional mechanisms of urticating setae are described, with the most prevalent known as bombardment. Here, an agitated tarantula projects its setae into the air and towards a potential threat. If setae get in contact with vulnerable body parts (skin, eyes, respiratory system) of predators, they induce painful micro-injuries owed to their harpoon-like ultrastructure. A rather intriguing defensive trait, urticating setae were unsurprisingly often studied for their evolution (Bertani & Guadanucci, 2013; Kaderka et al., 2019; Pérez-Miles & Perafán, 2015). However, as a result of systematic ambiguity, no hypothesis about their evolution has been proposed with a reliable foundation, thus leading to mostly controversial findings.

The results of Chapter II suggest that all subfamilies in which taxa can carry urticating setae form a rather young monophyletic clade. This neotropical "bombardier clade" consists of Theraphosinae, Aviculariinae, Schismatothelinae, and Psalmopoeinae. Interestingly, urticating setae are not found in every species within the bombardier clade, and considering their distribution throughout the clade, their evolution could be explained by three different scenarios. A first scenario would conform to a multiple-gain process, in which urticating setae evolved independently in Aviculariinae, *Ephebopus,* and Theraphosinae. Second, a multiple-loss scenario would predict that urticating setae evolved once in the last common ancestor (LCA) of the bombardier clade but were then lost twice in Schismatothelinae and Psalmopoeinae after the split off of *Ephebopus*. Lastly, a gain-loss-gain scenario would propose the emergence of urticating setae in the LCA and a secondary loss when Schismatothelinae diverged from Aviculariinae. Then, urticating setae were regained in *Ephebopus*, but not in Psalmopoeinae. The bioinformatic analysis within Chapter II favored the multiple-gain scenario and therefore indicated that convergent evolution of urticating setae likely explains the distribution of this peculiar trait throughout the bombardier clade. Interestingly, the presence of urticating setae relates to anecdotal reports on theraphosid venoms. Theraphosidae carry functional venom systems and are, with very few exceptions, not considered to be medically significant (Hauke & Herzig, 2017). However, anecdotal reports consistently highlighted that bites from old-world theraphosids are causing more severe envenomations in humans than bites from taxa from the bombardier clade (Escoubas & Rash, 2004). Therefore, the theraphosid tree of life appears to be bifurcated in terms of toxicity towards primates: while most plesiotypic lineages seem to cause relatively painful envenomations in humans, the reduction of toxicity towards humans coincides with the apotypic bombardier clade. Defense against predators is among the principal biological functions of venom (Fry et al., 2009). However, venoms are also metabolically expensive (Morgenstern & King, 2013; Wigger et al., 2002). Considering that, the loss of toxicity within the bombardier clade could be explained by the evolution of urticating setae. In a defensive situation, the

projection of urticating setae and the injection of venom could be used by tarantulas as a means to deter predators. While both mechanisms serve the same goal, they differ drastically in their metabolic expensiveness. Venom consists of proteins and other biomolecules that need to be re-synthesized and thus requires an energetic investment in replenishing the venom gland after it has been emptied (Morgenstern & King, 2013; Wigger et al., 2002). Urticating setae, on the other hand, are structures that consist of chitin and are localized on the exoskeleton (Kaderka et al., 2019). As ecdysozoans, tarantulas perform regular ecdysis throughout their life history, in which their old exoskeleton is replaced by a larger one enabling growth. During ecdysis, the arsenal of urticating setae is replaced, probably with significantly lower energetic costs when compared to venom components. Therefore, Chapter II proposed that the evolution of urticating setae within the bombardier clade may have caused the subsequent loss of defensively used venom components. Intriguingly, urticating setae represent key innovations for tarantulas of the bombardier clade that contribute to their evolutionary success, as lineages that harbor such setae exceed other tarantula groups in diversity. The bombardier clade subfamilies comprise ca. 60% of the total biodiversity of Theraphosidae, and genera with multiple types of setae usually have more congenerics than those with a less diverse setae arsenal.

Findings throughout this body of research verify the proposed hypothesis 2, as the erection of a reliable phylogenetic framework indeed enabled the study of trait evolution in tarantulas. However, the conclusions made about the evolutionary interactions of venom versus urticating setae need to be tested in the future. Facing the evolutionary success displayed by the bombardier clade and the fact that the presence of urticating setae is the only obvious ecological difference within Theraphosidae, it is likely that future studies may reveal an evolutionary mechanism that relates to these hypotheses.

Fighting off the taxon selection dilemma: Towards novel strategies for venom bioprospecting in spiders

Previous attempts at mining spider venom for translational applications were largely driven by taxonomic bias. Thus, the current view on the subject is derived from a non-representative minority of taxa (Herzig et al., 2019). To derive a more comprehensive picture and to increase the efficiency of bioprospecting programs, parts of this research were devoted to the optimization of related workflows. Among the most important factors in this realm is the taxon selection dilemma. Related to this specific problem is hypothesis 3, which states that a rational selection of target species may enable accelerated biodiscovery from venom systems. This hypothesis has been tested in Chapters III and V.

Chapter III developed an alternative path towards rational taxon selection in spider venom research that is based on phylogeny instead of size or medical significance. It has been proposed that phylogenetic distance is among the main drivers of venom diversification (Fry et al., 2015). Consequently, the discovery of novel bioresources in venoms is predicted to be more likely if phylogenetically distant groups are studied (Fry et al., 2015). However, this concept has not yet been applied to spider venom bioprospecting, partly because taxonomic ambiguity often prevented the inference of phylogenetic distances. Previous chapters of this thesis provided the needed phylogenetic framework for one spider family, and thus relieved a major constraint for such an approach, at least in this group. Using these spiders of the Theraphosidae family as a model group, the research of Chapter III took advantage of the available phylogeny and combined this with metadata for theraphosid venoms (Pineda et al., 2018) to illuminate priority groups. This work suggests that subfamilies such as Poecilotheriinae or Eumenophorinae likely reflect the most promising groups for venom bioprospecting. However, the proposed phylogeny-based approach needs to be further validated before its usefulness can be established. As a first step in this direction, the strategy has been applied towards the totality of spider families, thereby selecting the wasp spider *Argiope bruennichi* as a promising taxon (Fig. 4). The venomic analysis of *A. bruennichi* in Chapter V, recovered a plethora of biomolecules, including ICK peptides, MIT atracotoxins, and HAND family peptides. All of them, but in particular the ICKs, represent attractive leads for translational applications (Saez et al., 2010). Besides these, the presence of several peptides that seem to derive from neuropeptides has been discovered. Most of these neuropeptide-based venom components have never before been isolated from spider venoms.

Although these studies that sought to battle the taxon selection dilemma are of exploratory nature, the findings support the validity of hypothesis 3. By applying the strategy developed in Chapter

III, a swift and rational taxon selection for venomic analysis was enabled. This yielded, in its first trial, a variety of promising biomolecules worth further study. All of these are molecules that are new to science and belong to different classes of proteins and peptides. For several of these classes, this work was able to report their presence in spider venoms for the first time. It can therefore be deduced that the phylogeny-based strategy represents a promising approach for rational taxon selection. The subsequent study explored the venom of a single taxon, and thus more trials are needed to validate it. That said, if such subsequent trials are successful, the strategy's scope could even be expanded to other venomous lineages. For example, pseudoscorpions reflect a similarly diverse group of small venomous invertebrates (von Reumont et al., 2014). Therefore, bioprospecting from their venoms may be challenged by similar hurdles as in spiders. Overcoming the taxon selection dilemma in pseudoscorpions (or any other hyperdiverse group of venomous metazoans) may thus be enabled by this research as well.

Fig. 4: The wasp spider and their relatives. The wasp spider *A. bruennichi* (A)was selected as a promising target species for venom bioprospecting and subsequently studied in the chapters IV and V. The species belongs to the understudied Araneidae family and displays a charactersistic black and yellow coloration. This phenotype is shared in several congenerics, including *Argiope aurantia* (B) and *Argiope lobata* (C). Images courtesy of W. Dibiasi.

Challenging the current picture on spider venom composition

Among the major questions behind the venom evolution conundrum in spiders is the evolutionary origin from which toxins arose, how venom systems are organized, and how they are composed. One important goal was to gain new knowledge on spider venom evolution beyond the current status quo. Addressing this important task, hypothesis 4 states that the study of venom systems beyond taxonomic bias will recover venom profiles that challenge the current picture on spider venoms, as this is derived from only a marginal fraction of the exorbitantly diverse spider kingdom.

The in-depth venomic analysis of the wasp spider venom by proteo-transcriptomics conformed to this hypothesis, as it recovered an unusual venom profile with an unprecedented degree of simplicity. This finding drastically contradicts the paradigm that spider venoms are outstandingly complex entities. Although the venoms of *Cupiennius salei* and *Physocyclus mexicanus* were recently found to be relatively simple (Kuhn-Nentwig et al., 2019; Zobel-Thropp et al., 2019), their chemical complexity still exceeds that featured in *A. bruennichi* by far. Another commonly accepted concept is that small neurotoxic components, and in particular those with an ICK scaffold, dominate spider venoms. The venom of wasp spiders, again, defies this rule: it is foremost composed of large molecules in the CAP superfamily. Contrary to other spider venoms, that of *A. bruennichi* yields only a few small neurotoxic peptides, and those with an ICK scaffold do not reflect a significant proportion of the venom profile. In parallel to the venom of *A. bruennichi*, the venom of a related araneid species (*Araneus ventricosus*) was re-analyzed because the original work did not provide a functional classification of identified proteins. This approach revealed that the venom of *A. ventricosus* is also predominantly composed of large molecules, mostly thryoglobulins and CAPs, and, similar to wasp spiders, ICKs are underrepresented.

In this perspective, the work that was conducted in Chapter V addresses the venom evolution conundrum as it adds to the increasing body of evidence that large proteins are indeed pivotal components in spider venom systems. It also emphasizes that the commonly made assumptions of spider venoms are taxonomically biased, and that the study of neglected venom systems in spider will reveal rather different venom profiles.

Changing perspectives on the importance of large proteins in light of the dual prey inactivation strategy

For most of toxinological history, large proteins from spider venoms were neglected, and it was questioned whether they constituted important or abundant components at all. Most authors recognized large proteins, except LTX and PLD, as minor components that were unlikely to mediate key functions in spider venoms (Langenegger et al., 2019). This perspective has recently begun to change, as an increasing number of studies have reported on large components (Kuhn-Nentwig et al., 2019; Undheim et al., 2013; Zobel-Thropp et al., 2019). Some of these belong to protein classes that are known as toxins from other animals, and thus represent potential toxins for spiders. Others were functionally linked to cellular homeostasis within the venom system, as they are involved in the maturation of toxins. For most of the large proteins, biological functions remain unstudied (Langenegger et al., 2019). As indicated by the venoms of *A. bruennichi* and *A. ventricosus*, some spider venoms can be largely composed of such proteins and therefore defy the widespread rule of small peptide dominance.

Despite the enigmatic nature of most large proteins, a recent study proposed the dual prey inactivation strategy for spider venom (Kuhn-Nentwig et al., 2019). It postulated that spiders inject a venom cocktail of neurotoxic small peptides together with an array of larger proteins into their prey. Some of the larger proteins are thought to interfere with metabolic pathways that mediate cellular homeostasis. Others are proposed to affect the integrity of cell membranes and therefore enhance the dispersion of neurotoxins, thus acting as spreading factors. Neurotoxins mediate their activity on ion channels or receptors and subsequently lead to prey immobilization (Kuhn-Nentwig et al., 2019). The dual prey inactivation strategy therefore proposes that spider venoms largely rely on the synergistic machinery of large and small components that employ their biochemical properties in two waves. While some large proteins rapidly interfere with the victim's metabolism, thus mounting a first physiological attack, others increase the spread of neurotoxins. Then, in a second wave, these neurotoxic compounds overpower the prey item (Kuhn-Nentwig et al., 2019). In total, the strategy implies that large proteins, although functionally understudied, reflect important parts of spider venom systems, and thus rejects their previous perception as minor compounds.

In light of this novel hypothesis, research in Chapter V recovered several venom components from *A. bruennichi* that may relate to the dual prey inactivation strategy. It identified S10-Peptidases, astacin-like metalloproteases, serine-proteases, and lectins within the venom of *A. bruennichi*. Members of these molecule classes are widespread throughout animal venoms and are known to have cytotoxic activities. Thus, they may represent spreading factors in *A. bruennichi* following the dual prey inactivation strategy (Kuhn-Nentwig et al., 2019). Another spreading factor could be represented by chitinases, as these degrade chitin, a major structural element of insect prey items. For proteins that interfere with cellular homeostasis and thus represent the first wave of chemical attack within the preys metabolism, nucleotidases, Kunitz peptides, and cystatin were recovered as potential candidates. That said, the presence of large proteins with thyroglobulin- and leucine-rich repeat domains inside the venom of *A. bruennichi* have been detected. For these, no data on their biological activities is available from any venom system, and their function needs to be established by future studies to interpret these molecules in perspective to the dual prey inactivation strategy. Lastly, the presence of neurotoxic ICKs, HAND peptides, and MIT atracotoxins, plus a subset of putative novel neurotoxins, likewise conforms to the dual prey inactivation strategy assumptions (Kuhn-Nentwig et al., 2019).

The major wasp spider venom components that this work revealed belong to the class of CAP proteins. These are frequently recruited venom proteins and are featured in virtually all known venom systems (Fry et al., 2009). Throughout the venomous tree of life, they acquired a wide array of biological functions. While they are proteolytic enzymes in cone snails, they reflect neurotoxins in snake venom and hemotoxins in ticks, some insects, and lampreys (Fry et al., 2009). In spider venoms, CAPs are also frequently reported, however, their biological functions and modes of action have not been established (Fry et al., 2009; Kuhn-Nentwig et al., 2019; Undheim et al., 2013). In the context of the dual prey activation strategy, it is assumed that spider venom CAPs are parts of the first wave in envenomation that interferes with the prey's metabolic homeostasis (Kuhn-Nentwig et al., 2019). That said, proteolytic activity, as well as neurotoxic activities, are known in CAPs outside of spider venoms. It is therefore also possible that they reflect spreading factors or neurotoxins involved in the second wave of envenomation.

Consequently, the functional assignment of CAPs within the dual prey inactivation strategy remains to be validated. Chapter V compared the primary structure of CAPs from *A. bruennichi* venom with other, functionally resolved CAPs. It confirmed that the C terminal domain that facilitates the neurotoxic activity of snake venom CAPs is absent from those in *A. bruennichi,* and concluded that a neurotoxic function is unlikely (Fry et al., 2009). Instead, it found that all CAPs recovered from wasp spider venom displayed a high sequence similarity with a proteolytic CAP from the cone snail *Conus textile* (Tex31) (Milne et al., 2003). Tex31 facilitates its proteolytic activity based upon its active site that contains histidine and glutamic acid (Milne et al., 2003). In order to derive insights about the potential biological function of CAPs in *A. bruennichi*, exploratory structural-sequence alignments and homology modelling against Tex31 plus structurally elucidated CAPs from other organisms were conducted (Fig 5). The alignment revealed that the active site of Tex31, which contains glutamic acid and histidine moieties, is conserved in *A. bruennichi* CAPs (Fig. 5 a). However, these feature an insertion of 27 amino acids between both moieties, thus disrupting their presupposed proximity. Interestingly, the insertion within *A. bruennichi* CAPs carries an additional set of glutamic acid and histidine, and therefore comprises an additional putative active site that mirrors Tex31. The homology modelling revealed that these moieties indeed are in proximity and are localized within a potential binding pocket (Fig. 5 b). Moreover, if the protein's predicted folding is taken into consideration, this active site may be expanded by the original histidine that is present in Tex31. Through three-dimensional orientation this distant moiety is brought in proximity to the inserted putative active site of *A. bruennichi* CAPs, forming a potential triad of glutamic acid and two histidines (Fig. 5 b and c). Therefore, the active site that is responsible for proteolytic activity in Tex31 is present and potentially even expanded in the CAPs of wasp spiders, suggesting that these also have a proteolytic function.

Based on these insights, the results of this work reject that CAPs from wasp spider venom are involved in the metabolic perturbance during the first wave as it is proposed by the dual prey inactivation strategy (Kuhn-Nentwig et al., 2019). Instead, it supports the theory that the CAPs of wasp spiders represent spreading factors, although another role within extraintestinal digestion cannot be excluded. This analysis of CAP proteins from wasp spider venom are in agreement with hypothesis 4, as this analysis beyond the taxonomic bias reshaped the perception of biological functions of CAP proteins in spider venom systems.

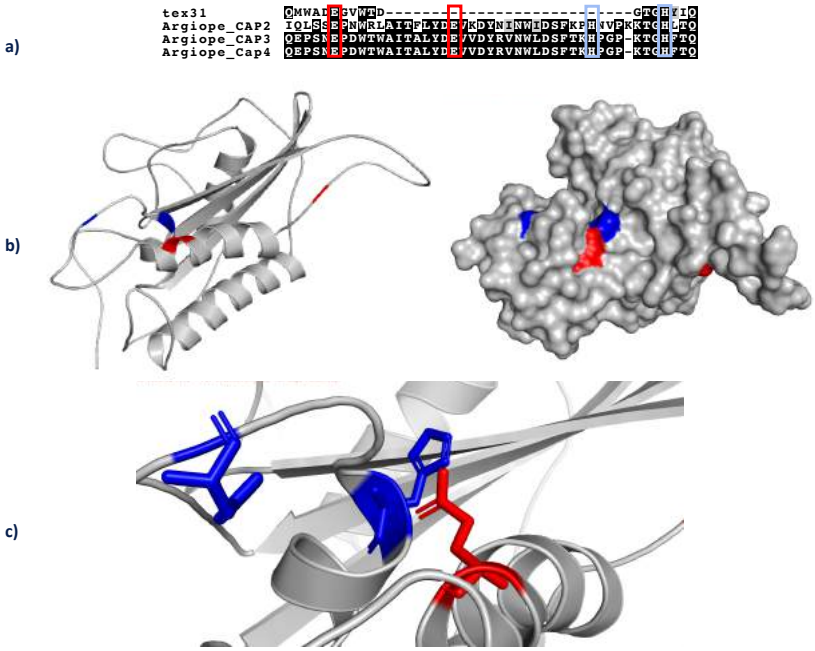


Fig. 5. Structural insights on the putative function of wasp spider CAPs. (a) Structural-sequence alignment of selected *A. bruennichi* CAPs against the active site of Tex31. The active site of Tex31 consists of glutamic acid (E, red) and histidine (H, blue) and is conserved between *A. bruennichi* CAPs. The latter carry a 27 amino acid insertion with an additional set of glutamic acid and histidine, forming a potential active site. (b) homology modelling shows (left protein structure, right protein surface) that the inserted active site and the histidine of the Tex31 active site are nested within a putative binding pocket. (c) Enlarged view on the three putative active site moieties in *A. bruennichi* CAPs reveal that their side chains are in proximity, thus forming a potential triad. The structural insights support a proteolytic activity of *A. bruennichi* CAPs.

The architecture of spider venom systems and hidden complexity of the venom duct

The venoms of most spiders have not yet been studied, and even less is known about the morphological organization of their venom systems. Given the wide variety of spider species that employ their venom systems in different ecological contexts, hypothesis 5 predicts that the morphology of these systems is functionally underestimated.

In all spiders, the outer part of the venom system is composed of chelicerae consisting of a large basal segment and a fang. On the inside, the system comprises a venom gland that produces and stores the venom, an orifice at the fangs end from which venom is released, and a duct that connects the gland with the opening (Foelix, 1983). The morphological analysis of the wasp spider venom system made in Chapter IV illuminated that the outer morphology of the chelicerae conforms to the basic structure that is typically reported in Araneomorphae. At the basal segment, *A. bruennichi* harbors cheliceral teeth into which the fangs are folded in the resting position. Such teeth have been proposed to increase friction and enable the spider to use their chelicerae during the handling of prey (Foelix, 1983). The presence of cheliceral teeth in *A. bruennichi* suggests that this spider utilizes its venom system in more scenarios than just envenoming prey, for example as a means of adjusting the prey item during feeding within its orb web.

A histological analysis of the wasp spider venom system found that the venom gland is composed of a complex system of secretory cells that release vesicles, presumably filled with venom components towards the glands lumen. The gland itself is covered in a dense layer of muscle bundles and is attached to nervous fibers. Innervation and muscles are likely the means of regulating the release of venom. As spiders are known to be capable of controlling the amount of venom used depending on the prey type, such an architecture can be expected, as a high degree of regulation is needed to facilitate the fine tuning of the venom system. Moreover, the histology of the wasp spider venom system largely mirrors the venom systems of the few other spiders studied so far. Despite coming from different infraorders and occupying different ecological niches, spiders seem to rely on the same cellular architecture for their venom systems.

The most important insights gained from Chapter IV concern the often-overlooked venom duct. This structure was found to be composed of four different subsections that differ histologically. While venom ducts in most species are considered to be a simple connection between the injector and venom gland, the study of venom ducts in cone snails and *Bothrops* pit vipers revealed that venom ducts can be metabolically active and synthesize parts of the venom cocktail (Hu et al., 2011; Valente et al., 2018). These findings mean that venom ducts in other venomous species, including spiders, may be of higher importance than previously thought. To what extent these insights apply to the wasp spider is a subject for future investigations. However, a biological function of the venom duct seems rather likely considering the given substructures.

The insights gained by morphologically examining the venom system of wasp spiders partly agree with and partly refute hypothesis 5. On one hand, the expected functional underestimation of spider venom systems was not found for most parts of the chelicerae and the venom gland, as these mirror those from other spiders. On the other hand, the presence of cheliceral teeth and the new insights regarding the potential role of the venom duct as part of the toxin synthesizing machinery are in agreement with hypothesis 5. It is important for future research to test these potential roles, in particular that of the venom duct, via appropriate technologies such as MALDI-imaging approaches.

Neuropeptides as frequently recruited spider venom components

By exploring the wasp spider venom from a chemical perspective, the presence of a variety of venom components that share similarities with arthropod neuropeptides has been revealed. These findings relate to hypothesis 6, which predicts that deep venomic studies in neglected spider lineages will recover venom components that are derived from understudied protein groups. Thus, the identification of the evolutionary substrate giving rise to toxic proteins beyond ICKs, PLD, and LTX will be enabled.

In general, neuropeptides are promising evolutionary substrates for the birth of toxins, as they are involved in the control and regulation of vital processes. Dysregulation of these is likely to cause perturbations of neuronal chemistry and thus neurotoxicity. Recruitment into the venom system may already be sufficient to weaponize neuropeptides. Following initial recruitment, the toxicity of these

neuropeptides can be enhanced via subsequent mutations that alter the surface chemistry. The simplicity by which neuropeptides can be recruited to venom systems is flagged by the diversity of venomous taxa that incorporate such molecules into their venoms, for example, the emerald jewel wasp (*Ampulex compressa*) utilizes neuropeptides in its venom to induce hypokinesia in cockroaches (Arvidson et al., 2019). Helodermatid lizards, centipedes, ticks, and wasps equally feature weaponized neuropeptides in their venoms, and even the skin poison of frogs harbors an arsenal of neuropeptides (Roelants et al., 2013; Undheim et al., 2015; Yap & Misuan, 2019). The recruitment of neuropeptides into spider venoms has been described only on a few occasions. The most prominent examples are HAND peptides that were shown to be recruited into the venom system of spiders and centipedes via weaponization of ion transport peptide/ crustacean hyperglycemic hormone neuropeptides (ITP/CHH) (Undheim et al., 2015). The same family of neuropeptides has been recruited to the venom of Theridiidae, where they gave rise to Latrodectins (McCowan & Garb, 2014). Beyond the modification of ITP/CHH peptides, little is known about the importance of neuropeptides as the evolutionary substrate for spider toxins.

In Chapter V, the presence of HAND peptides was revealed in wasp spider venom. As these evolved from ITP/CHH neuropeptides, this study supports that the recruitment of neuropeptides into the venom system also happened in *A. bruennichi*. This is further emphasized by other results from Chapter V, which illuminated the presence of an array of peptides that represent weaponized neuropeptides. Among these is a set of putative toxins with homology to diuretic hormones. Venom components that are derived from such hormones have not been described from spider venom before. Next, a group of toxins that mirrors insulin-like growth factor binding proteins has been identified. These are functionally unstudied but have previously been found in venoms of several arachnids, including a spider (*Cupiennius salei*) (Kuhn-Nentwig et al., 2019). The last group of putative neuropeptide-derived venom components features high similarity with ITG-like peptides. Representatives within *A. bruennichi* venom mirrors those from the black cutworm moth (*Agrotis ipsilon*) and the Florida carpenter ant (*Camponotus floridanus*). As diuretic hormones, members of ITG-like peptides have not been described from spider venom before.

The findings of Chapter V conform to the predictions made by hypothesis 6 and indicate that neuropeptides and hormones represent a prolific evolutionary substrate for novel tools in spider venom systems. Also, the chapter expands the knowledge on the diversity of neuropeptide-based toxins, as it discovers two new families that have never before been found in spider venoms. These results represent important advances regarding the venom evolution conundrum, as for many spider toxins, it is unknown from which ancestral proteins and peptides they arose. However, it must be established whether the discovered neuropeptide-derived molecules display neurotoxic activities or if they instead participate in other, yet undiscovered processes within the venom of *A. bruennichi*.

A reductive approach to venom systems? The role of purifying selection in spider venom evolution

It has been a long-standing paradigm that venoms evolve rapidly under strong positive Darwinian selection (Gibbs & Rossiter, 2008; Juárez et al., 2008; Lynch, 2007; Sunagar et al., 2013). In agreement with the Red Queen hypothesis of Van Valen, it has been proposed that venoms in predators and venom resistance in prey exert reciprocal positive selection on each other, thus leading to the diverse and powerful pharmacopeia observed in extant venoms (Fry et al., 2015; Holford et al., 2018; Van Valen, 1974). However, this assumption has been derived from studies exclusively using evolutionary young organism classes, in particular reptiles (Sunagar & Moran, 2015). By studying the selection pressures on venom genes throughout the whole animal kingdom, recent studies discovered that selective pressures differ between younger and older clades (Sunagar & Moran, 2015). While venom genes in young clades indeed tend to evolve under strong positive selection, older lineages such as spiders show signatures of purifying selection. This led to the proposal of the "two-speed mode" of venom genes, which predicts that the diversifying nature of positive selection mostly acts on the earlier stages of ecological specialization within a given species (Sunagar & Moran, 2015). This period of diversification is then followed by an extended stage of purifying selection that imposes heavy constraints on the venom system. Although this process has been shown to act on spider venom genes on the nucleotide level, it is

questionable how this type of selection – which conserves advantageous and purges deleterious alleles – shapes the venom systems.

By investigating the evolutionary history of Theraphosidae, Chapter II focused on the bombardier clade. Its members are known to exert less significant envenomations in humans compared to other theraphosid lineages (Escoubas & Rash, 2004). It has been discussed above that the evolution of urticating setae within the bombardier clade caused the loss of defensively used toxins. Toxins are energetically expensive, as their biosynthesis affords metabolic investments (Morgenstern & King, 2013; Wigger et al., 2002). Consequently, any evolutionary innovation that relieves these toxins from their biological function would essentially turn them into a deleterious trait, as their biosynthesis and maintenance would be energetically wasteful. Urticating setae, on the other hand, are composed of chitin and are recovered during each ecdysis, thus representing a metabolically less expensive alternative to defensive toxins. The negative selection that has been proposed for spider venoms could therefore lead to a subsequent purge of toxin alleles. The energy that is saved by such a process could instead be invested in other expensive processes, such as reproduction or growth. The fact that urticating setae-carrying tarantulas are the most diverse lineages within the family and that they also comprise the largest members could suggest that these concepts may apply to Theraphosidae (Mammola et al., 2017).

Chapter V investigated the venom of the wasp spider *A. bruennichi* and proposed that its simple venom results from an energetic dilemma between the venom system and the silk apparatus. In contrast to most spiders, *A. bruennichi* employs a unique hunting behavior. While others tend to overpower prey items via injection of venom, *A. bruennichi* utilizes its silk apparatus instead (Eisner & Dean, 1976). Prey items that enter the foraging web usually have silk spun around them, and only when immobilized by the silk are they injected with venom (Eisner & Dean, 1976). Therefore, the biological function of prey immobilization is relieved from the venom system and instead imposed on the silk apparatus. Given that toxins within the venom, as well as proteins within the silk, are metabolically expensive, the wasp spider faces an economic dilemma in which both systems compete for resources. In the case of *A. bruennichi*, many venom components are therefore energetically wasteful and likely to fall victim to purifying selection, thus relieving the spider from a metabolic burden. However, for some lepidopteran prey items, *A. bruennichi* switches its strategy and applies a venomous bite first (Nyffeler & Benz, 1981), a behavior that could explain the presence of the few neurotoxic components detected.

Little is known about the evolutionary processes behind spider venoms, although their toxins have already been under investigation for a long time. This venom evolution conundrum is still largely unresolved, and many details and nuances will be discovered in future studies, hence the proposal of hypothesis 6. However, the herein presented findings on theraphosid defenses and araneid hunting seem to conform with the proposed importance of purifying selection and with the prediction of hypothesis 6 that the mechanisms at play in spider venom evolution can be experimentally disentangled (Sunagar & Moran, 2015). This work suggests that in these two spider groups, evolutionary forces are driving the venom systems towards simplicity. Components that are made redundant by morphological or ethological alternatives are seemingly removed from venom systems. Therefore, the spiders investigated in this work seem to pursue a reductive approach to venom evolution in which venom systems tend to be simplified instead of diversified.

Conclusions and future perspectives

This work embodies the results of subsequent phylogenetic and venomic investigations of the four prevailing challenges for the science of arachnology guided by six proposed hypothesis. The major conclusions of each chapter are summarized in Fig. 6 and contribute to a potential resolution of these challenges.

By resolving large swathes of the theraphosid phylogeny, this work advanced the battle against systematic ambiguity. It clarified the systematics of several important lineages and made a variety of taxonomic changes. However, due to limited taxon sampling, a few important groups are still missing. In particular, the subfamily Selenogyrinae, as well as the paraphyletic groups Ischnocolinae and Schismatothelinae, demand a broader taxon sampling. Although the first evolutionary hypothesis upon major theraphosid lineages was inferred, large parts of the lower taxonomic levels are still unresolved. A myriad of follow-up studies by experts who devote their scientific interest to these rather specialized

groups will be required. Moreover, a broad sampling of tarantula venoms will be needed in order to test the new hypotheses about the interactions of venom systems and urticating setae. The long and cumbersome process of collecting venoms and specimen from missing taxa has been initiated in the wake of this thesis.

This work aimed to improve bioprospecting from spider venoms. The strategy that has been developed needs to be validated in the future, as the first trial relied only on a single taxon. More studies aiming to expand the strategy towards other spider venom systems and to establish it as a rationale for taxon selection in venom bioprospecting are already under way.

Improving the taxonomic bias in spider venom research was another focus of this research, and applying the taxon selection strategy contributed to its solution. In total, ca. 50 novel components were identified and added to the toxin arsenal of the understudied Araneidae family, reflecting 25% of all known venom components from this group. This major advance regarding the understanding of spider venom compositions beyond the taxonomic bias is accompanied by findings regarding the arachno-atypical nature of wasp spider venom. These findings highlight that the picture of spider venoms, in general, is indeed not representative, and that studies on neglected spider lineages may drastically alter the understanding of spider venoms overall.

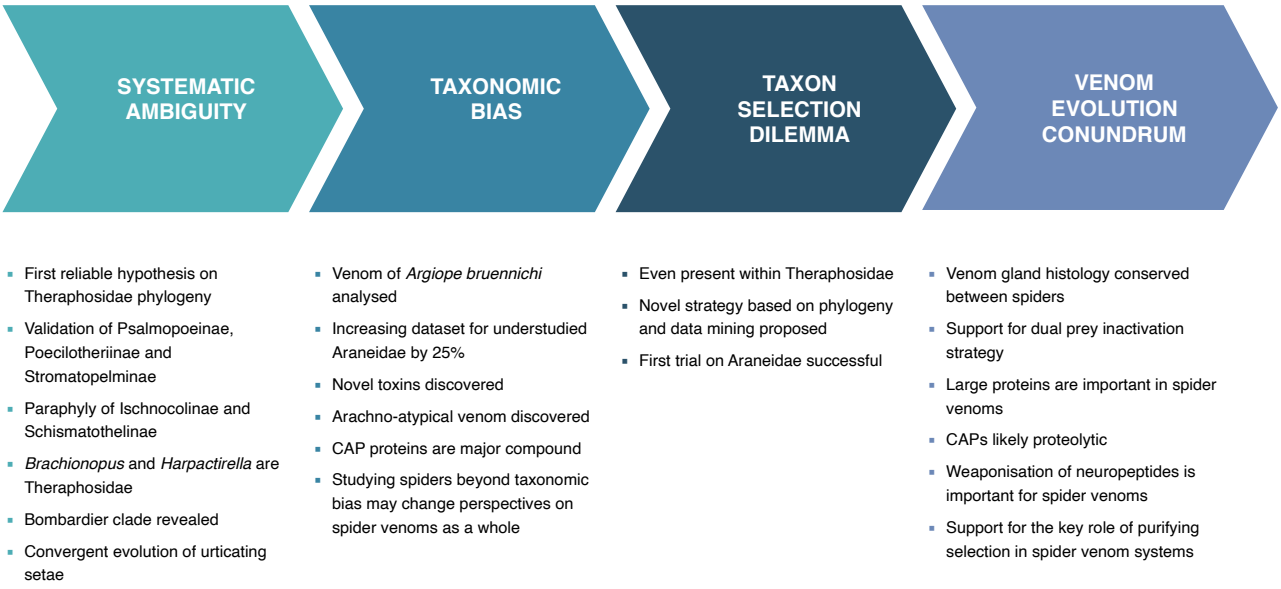


Fig. 6. Summarised results of my research conducted throughout this thesis. Major findings are summarised as bullet-points below each of the addressed prevailing challenges.

Several insights gained by this work expanded the knowledge on spider venom evolution processes. In this regard, Chapter IV demonstrated that spider venom systems employ an astonishing degree of conversation for their cellular organization, but that the complexity of the venom duct has been underestimated. Moreover, it has been established that neuropeptides are important yet understudied evolutionary starting points for the birth of toxic proteins in spiders. Results throughout this thesis support the recently proposed dual prey inactivation strategy of spider venom and indicate that large proteins are more important for spider venoms than previously thought. That said, for the CAP proteins in spider venom, a role in the first wave of envenomation as proposed by the dual prey inactivation strategy could be rejected. Instead, a biological role as a spreading factor or as an agent for extraintestinal digestion is supported. Lastly, this work found evidence for the importance of purifying selection in venom systems of Theraphosidae and Araneidae and postulates mechanistic hypotheses to be tested in the future.

A large fraction of hypotheses derived from insights on the analyzed venom systems are fully theoretical and demand experimental approaches. This is, in particular, true for the novel identified venom components derived from neuropeptides and for establishing the biological role of large proteins in the face of the dual prey inactivation strategy. However, due to the small size of most spiders, it is

currently not possible to collect sufficient amounts of venom to allow the testing of isolated compounds. It is therefore an important task for future studies to develop methods that allow the isolation of trace compounds from venom samples of a limited amount. As an alternative, biotechnology may represent an answer to this fundamental problem. By engineering prokaryote or eukaryote cells for the production of venom components discovered by venomics, it should be possible to produce sufficient amounts of material for accessing their enigmatic bioactivities – even from taxa that have previously been impossible to study in this way due to their size and venom yield. Studies in this direction are currently underway.

IV References

Agwa, A. J., Henriques, S. T., & Schroeder, C. I. (2017). Gating modifier toxin interactions with ion channels and lipid bilayers: Is the trimolecular complex real? In *Neuropharmacology*. https://doi.org/10.1016/j.neuropharm.2017.04.004

Alape-Girón, A., Sanz, L., Escolano, J., Flores-Díaz, M., Madrigal, M., Sasa, M., & Calvete, J. J. (2008). Snake venomics of the lancehead pitviper bothrops asper. Geographic, individual, and ontogenetic variations. *Journal of Proteome Research*. https://doi.org/10.1021/pr800332p

Amazonas, D. R., Portes-Junior, J. A., Nishiyama-Jr, M. Y., Nicolau, C. A., Chalkidis, H. M., Mourão, R. H. V., Grazziotin, F. G., Rokyta, D. R., Gibbs, H. L., Valente, R. H., Junqueira-de-Azevedo, I. L. M., & Moura-da-Silva, A. M. (2018). Molecular mechanisms underlying intraspecific variation in snake venom. *Journal of Proteomics*. https://doi.org/10.1016/j.jprot.2018.03.032

Aminetzach, Y. T., Srouji, J. R., Kong, C. Y., & Hoekstra, H. E. (2009). Convergent Evolution of Novel Protein Function in Shrew and Lizard Venom. *Current Biology*. https://doi.org/10.1016/j.cub.2009.09.022

Arvidson, R., Kaiser, M., Lee, S. S., Urenda, J. P., Dail, C., Mohammed, H., Nolan, C., Pan, S., Stajich, J. E., Libersat, F., & Adams, M. E. (2019). Parasitoid jewel wasp mounts multipronged neurochemical attack to hijack a host brain. *Molecular and Cellular Proteomics*. https://doi.org/10.1074/mcp.RA118.000908

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B., & Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? In *Nature*. https://doi.org/10.1038/nature09678

Bertani, R., & Guadanucci, J. P. L. (2013). Morphology, evolution and usage of urticating setae by tarantulas (Araneae: Theraphosidae). *Zoologia*. https://doi.org/10.1590/S1984-46702013000400006

Bertani, R., Nagahama, R. H., & Fukushima, C. S. (2012). Vitalius nondescriptus comb. nov. (araneae: Theraphosidae: Theraphosinae): An example of theraphosid taxonomic chaos. *Zoologia*. https://doi.org/10.1590/S1984-46702012000500011

Biner, O., Trachsel, C., Moser, A., Kopp, L., Langenegger, N., Kämpfer, U., Von Ballmoos, C., Nentwig, W., Schürch, S., Schaller, J., & Kuhn-Nentwig, L. (2015). Isolation, N-glycosylations and function of a hyaluronidase-like enzyme from the venom of the spider Cupiennius salei. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0143963

Buczek, O., Wei, D., Babon, J. J., Yang, X., Fiedler, B., Chen, P., Yoshikami, D., Olivera, B. M., Bulaj, G., & Norton, R. S. (2007). Structure and sodium channel activity of an excitatory I 1-superfamily conotoxin. *Biochemistry*. https://doi.org/10.1021/bi700797f

Calbacho-Rosa, L., Galicia-Mendoza, I., Dutto, M. S., Córdoba-Aguilar, A., & Peretti, A. V. (2013). Copulatory behavior in a pholcid spider: Males use specialized genitalic movements for sperm removal and copulatory courtship. *Naturwissenschaften*. https://doi.org/10.1007/s00114-013-1038-1

Canning, G., Reilly, B. K., & Dippenaar-Schoeman, A. S. (2014). The Distribution and Population Status of Nesiergus insulanus Araneae: Theraphosidae: Ischnocolinae) on Frégate Island, Seychelles . *Arachnology*. https://doi.org/10.13156/arac.2014.16.4.116

Cardoso, F. C., & Lewis, R. J. (2019). Structure-function and therapeutic potential of spider venom-derived cysteine knot peptides targeting sodium channels. *Frontiers in Pharmacology*. https://doi.org/10.3389/fphar.2019.00366

Cargnelutti, F., Calbacho-Rosa, L., Córdoba-Aguilar, A., & Peretti, A. V. (2018). Patterns of Sperm Transfer Behavior in a Pholcid Spider with Two Distinct Copulatory Phases. *Journal of Insect Behavior*. https://doi.org/10.1007/s10905-018-9702-0

Casewell, N. R. (2017). Evolution: Gene Co-option Underpins Venom Protein Evolution. In *Current Biology*. https://doi.org/10.1016/j.cub.2017.05.091

Casewell, N. R., Huttley, G. A., & Wüster, W. (2012). Dynamic evolution of venom proteins in squamate reptiles. *Nature Communications*. https://doi.org/10.1038/ncomms2065

Casewell, N. R., Wagstaff, S. C., Harrison, R. A., Renjifo, C., & Wüster, W. (2011). Domain loss facilitates accelerated evolution and neofunctionalization of duplicate snake venom metalloproteinase toxin genes. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msr091

Casewell, N. R., Wagstaff, S. C., Wuster, W., Cook, D. A. N., Bolton, F. M. S., King, S. I., Pla, D., Sanz,

L., Calvete, J. J., & Harrison, R. A. (2014). Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1405484111

Casewell, N. R., Wüster, W., Vonk, F. J., Harrison, R. A., & Fry, B. G. (2013). Complex cocktails: The evolutionary novelty of venoms. In *Trends in Ecology and Evolution*. https://doi.org/10.1016/j.tree.2012.10.020

Chassagnon, I. R., McCarthy, C. A., Chin, Y. K. Y., Pineda, S. S., Keramidas, A., Mobli, M., Pham, V., De Silva, T. M., Lynch, J. W., Widdop, R. E., Rash, L. D., & King, G. F. (2017). Potent neuroprotection after stroke afforded by a double-knot spider-venom peptide that inhibits acid-sensing ion channel 1a. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1614728114

Chippaux, J. P., Williams, V., & White, J. (1991). Snake venom variability: methods of study, results and interpretation. In *Toxicon*. https://doi.org/10.1016/0041-0101(91)90116-9

Clark, R. J., Harland, D. P., & Jackson, R. R. (2000). Speculative hunting by an araneophagic salticid spider. *Behaviour*. https://doi.org/10.1163/156853900502736

Cordes, M. H. J., & Binford, G. J. (2018). Evolutionary dynamics of origin and loss in the deep history of phospholipase D toxin genes. *BMC Evolutionary Biology*. https://doi.org/10.1186/s12862-018-1302-2

Corzo, G., Villegas, E., Gómez-Lagunas, F., Possani, L. D., Belokoneva, O. S., & Nakajima, T. (2002). Oxyopinins, large amphipathic peptides isolated from the venom of the wolf spider Oxyopes kitabensis with cytolytic properties and positive insecticidal cooperativity with spider neurotoxins. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M200511200

Daltry, J. C., Wüster, W., & Thorpe, R. S. (1996). Diet and snake venom evolution. *Nature*. https://doi.org/10.1038/379537a0

Deplazes, E., Henriques, S. T., Smith, J. J., King, G. F., Craik, D. J., Mark, A. E., & Schroeder, C. I. (2016). Membrane-binding properties of gating modifier and pore-blocking toxins: Membrane interaction is not a prerequisite for modification of channel gating. *Biochimica et Biophysica Acta - Biomembranes*. https://doi.org/10.1016/j.bbamem.2016.02.002

Dos Santos, V. L. P., Franco, C. R. C., Viggiano, R. L. L., Da Silveira, R. B., Cantão, M. P., Mangili, O. C., Veiga, S. S., & Gremski, W. (2000). Structural and ultrastructural description of the venom gland of Loxosceles intermedia (brown spider). *Toxicon*. https://doi.org/10.1016/S0041-0101(99)00155-5

Drukewitz, S. H., Bokelmann, L., Undheim, E. A. B., & von Reumont, B. M. (2019). Toxins from scratch? Diverse, multimodal gene origins in the predatory robber fly Dasypogon diadema indicate a dynamic venom evolution in dipteran insects. *GigaScience*. https://doi.org/10.1093/gigascience/giz081

Drukewitz, S. H., Fuhrmann, N., Undheim, E. A. B., Blanke, A., Giribaldi, J., Mary, R., Laconde, G., Dutertre, S., & von Reumont, B. M. (2018). A dipteran's novel sucker punch: Evolution of arthropod atypical venom with a neurotoxic component in robber flies (asilidae, diptera). *Toxins*. https://doi.org/10.3390/toxins10010029

Dunlop, J. A., Penney, D., & Jekel, D. (2015). A summary list of fossil spiders and their relatives. *World Spider Catalog*. https://doi.org/10.1111/j.1469-185X.2009.00099.x

Eisner, T., & Dean, J. (1976). Ploy and counterploy in predator - prey interactions: orb weaving spiders versus bombardier beetles. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.73.4.1365

Eisoldt, L., Smith, A., & Scheibel, T. (2011). Decoding the secrets of spider silk. In *Materials Today*. https://doi.org/10.1016/S1369-7021(11)70057-8

Engel, M. S. (2015). Insect evolution. In *Current Biology*. https://doi.org/10.1016/j.cub.2015.07.059

Escoubas, P., Bernard, C., Lambeau, G., Lazdunski, M., & Darbon, H. (2003). Recombinant production and solution structure of PcTx1, the specific peptide inhibitor of ASIC1a proton-gated cation channels. *Protein Science*. https://doi.org/10.1110/ps.0307003

Escoubas, P., & Rash, L. (2004). Tarantulas: Eight-legged pharmacists and combinatorial chemists. *Toxicon*. https://doi.org/10.1016/j.toxicon.2004.02.007

F. Pérez-Miles, S. M. Lucas, P. D. S. Y. R. B. (1996). Systematic revision and cladistic analalysis of

Theraphosinae (Araneae: Theraphosidae). In *Mygalomorph*.

Foelix, R., & Erb, B. (2010). Mesothelae have venom glands. *Journal of Arachnology*. https://doi.org/10.1636/b10-30.1

Foelix, R. F. (1983). Biology of Spiders. *Insect Systematics and Evolution*. https://doi.org/10.1163/187631283X00371

Forster, L. M. (1977). A qualitative analysis of hunting behaviour in jumping spiders (Araneae: Salticidae). *New Zealand Journal of Zoology*. https://doi.org/10.1080/03014223.1977.9517936

Fry, B. G., Casewell, N. R., Wüster, W., Vidal, N., Young, B., & Jackson, T. N. W. (2012). The structural and functional diversification of the Toxicofera reptile venom system. *Toxicon*. https://doi.org/10.1016/j.toxicon.2012.02.013

Fry, B. G., Koludarov, I., Jackson, T. N. W., Holford, M., Terrat, Y., Casewell, N. R., Undheim, E. A. B., Vetterb, I., Alia, S. A., Low, D. H. W., & Sunagar, K. (2015). Seeing the woods for the trees: Understanding venom evolution as a guide for biodiscovery. *RSC Drug Discovery Series*. https://doi.org/10.1039/9781849737876-00001

Fry, B. G., Roelants, K., Champagne, D. E., Scheib, H., Tyndall, J. D. A., King, G. F., Nevalainen, T. J., Norman, J. A., Lewis, R. J., Norton, R. S., Renjifo, C., & de la Vega, R. C. R. (2009). The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms. *Annual Review of Genomics and Human Genetics*. https://doi.org/10.1146/annurev.genom.9.081307.164356

Fry, B. G., Scheib, H., van der Weerd, L., Young, B., McNaughtan, J., Ramjan, S. F. R., Vidal, N., Poelmann, R. E., & Norman, J. A. (2008). Evolution of an Arsenal. *Molecular & Cellular Proteomics*. https://doi.org/10.1074/mcp.m700094-mcp200

Fry, B. G., Scheib, H., van der Weerd, L., Young, B., McNaughtan, J., Ryan Ramjan, S. F., Vidal, N., Poelmann, R. E., & Norman, J. A. (2008). Evolution of an arsenal: Structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Molecular and Cellular Proteomics*. https://doi.org/10.1074/mcp.M700094-MCP200

Fry, B. G., Vidal, N., Norman, J. A., Vonk, F. J., Scheib, H., Ramjan, S. F. R., Kuruppu, S., Fung, K., Hedges, S. B., Richardson, M. K., Hodgson, W. C., Ignjatovic, V., Summerhayes, R., & Kochva, E. (2006). Early evolution of the venom system in lizards and snakes. *Nature*. https://doi.org/10.1038/nature04328

Fukushima, C., Mendoza, J., West, R., Longhorn, S., Rivera, E., Cooper, E., Hénaut, Y., Henriques, S., & Cardoso, P. (2019). Species conservation profiles of tarantula spiders (Araneae, Theraphosidae) listed on CITES. *Biodiversity Data Journal*. https://doi.org/10.3897/BDJ.7.e39342.suppl9

Fukushima, C. S., & Bertani, R. (2017). Taxonomic revision and cladistic analysis of Avicularia lamarck, 1818 (Araneae, Theraphosidae, Aviculariinae) with description of three new aviculariine genera. *ZooKeys*. https://doi.org/10.3897/zookeys.659.10717

Fukushima, C. S., Bertani, R., & Da Silva, P. I. (2005). Revision of Cyriocosmus Simon, 1903, with notes on the genus Hapalopus Ausserer, 1875 (Araneae: Theraphosidae). *Zootaxa*. https://doi.org/10.11646/zootaxa.846.1.1

Garb, J. E. (2014). Extraction of venom and venom gland microdissections from spiders for proteomic and transcriptomic analyses. *Journal of Visualized Experiments*. https://doi.org/10.3791/51618

Garb, J. E., & Hayashi, C. Y. (2013). Molecular evolution of α-latrotoxin, the exceptionally potent vertebrate neurotoxin in black widow spider venom. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/mst011

Garcia, F., Villegas, E., Espino-Solis, G. P., Rodriguez, A., Paniagua-Solis, J. F., Sandoval-Lopez, G., Possani, L. D., & Corzo, G. (2013). Antimicrobial peptides from arachnid venoms and their microbicidal activity in the presence of commercial antibiotics. *Journal of Antibiotics*. https://doi.org/10.1038/ja.2012.87

Gendreau, K. L., Haney, R. A., Schwager, E. E., Wierschin, T., Stanke, M., Richards, S., & Garb, J. E. (2017). House spider genome uncovers evolutionary shifts in the diversity and expression of black widow venom proteins associated with extreme toxicity. *BMC Genomics*. https://doi.org/10.1186/s12864-017-3551-7

Gibbs, H. L., & Rossiter, W. (2008). Rapid evolution by positive selection and gene gain and loss: PLA 2 venom genes in closely related Sistrurus rattlesnakes with divergent diets. *Journal of Molecular*

Evolution. https://doi.org/10.1007/s00239-008-9067-7

Gosline, J. M., DeMont, M. E., & Denny, M. W. (1986). The structure and properties of spider silk. *Endeavour*. https://doi.org/10.1016/0160-9327(86)90049-9

Grishin, E. V. (1998). Black widow spider toxins: The present and the future. *Toxicon*. https://doi.org/10.1016/S0041-0101(98)00162-7

GUADANUCCI, J. P. L. (2007). A revision of the Neotropical spider genus Oligoxystre Vellard 1924 (Theraphosidae, Ischnocolinae). *Zootaxa*. https://doi.org/10.11646/zootaxa.1555.1.1

Guadanucci, José Paulo L., & Gallon, R. C. (2008). A revision of the spider genera Chaetopelma Ausserer 1871 and Nesiergus Simon 1903 (Araneae, Theraphosidae, Ischnocolinae). *Zootaxa*. https://doi.org/10.11646/zootaxa.1753.1.2

Guadanucci, José Paulo L. (2014). Theraphosidae phylogeny: Relationships of the "Ischnocolinae" genera (Araneae, Mygalomorphae). *Zoologica Scripta*. https://doi.org/10.1111/zsc.12065

Guadanucci, José Paulo Leite. (2005). Tarsal Scopula Significance In Ischnocolinae Phylogenetics (Araneae, Mygalomorphae, Theraphosidae). *Journal of Arachnology*. https://doi.org/10.1636/04-75.1

Guadanucci, José Paulo Leite, & Wendt, I. (2014). Revision of the spider genus Ischnocolus Ausserer, 1871 (Mygalomorphae: Theraphosidae: Ischnocolinae). *Journal of Natural History*. https://doi.org/10.1080/00222933.2013.809492

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., & De Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0185809

Hamilton, C. A., Formanowicz, D. R., & Bond, J. E. (2011). Species delimitation and phylogeography of Aphonopelma hentzi (araneae, mygalomorphae, theraphosidae): Cryptic diversity in north american tarantulas. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0026207

Hamilton, C. A., Hendrixson, B. E., & Bond, J. E. (2016). Taxonomic revision of the tarantula genus Aphonopelma Pocock, 1901 (Araneae, Mygalomorphae, Theraphosidae) within the United States. *ZooKeys*. https://doi.org/10.3897/zookeys.560.6264

Hargreaves, A. D., Swain, M. T., Hegarty, M. J., Logan, D. W., & Mulley, J. F. (2014). Restriction and recruitment-gene duplication and the origin and evolution of snake venom toxins. *Genome Biology and Evolution*. https://doi.org/10.1093/gbe/evu166

Harmer, A. M. T., Blackledge, T. A., Madin, J. S., & Herberstein, M. E. (2011). High-performance spider webs: Integrating biomechanics, ecology and behaviour. In *Journal of the Royal Society Interface*. https://doi.org/10.1098/rsif.2010.0454

Harris, R. J., & Arbuckle, K. (2016). Tempo and mode of the evolution of venom and poison in tetrapods. *Toxins*. https://doi.org/10.3390/toxins8070193

Hauke, T. J., & Herzig, V. (2017). Dangerous arachnids—Fake news or reality? In *Toxicon*. https://doi.org/10.1016/j.toxicon.2017.08.024

Hedin, M., Derkarabetian, S., Ramírez, M. J., Vink, C., & Bond, J. E. (2018). Phylogenomic reclassification of the world's most venomous spiders (Mygalomorphae, Atracinae), with implications for venom evolution. *Scientific Reports*. https://doi.org/10.1038/s41598-018-19946-2

Hendrixson, B. E., Guice, A. V., & Bond, J. E. (2015). Integrative species delimitation and conservation of tarantulas (Araneae, Mygalomorphae, Theraphosidae) from a North American biodiversity hotspot. *Insect Conservation and Diversity*. https://doi.org/10.1111/icad.12089

Henkel, A. W., & Sankaranarayanan, S. (1999). Mechanisms of α-latrotoxin action. In *Cell and Tissue Research*. https://doi.org/10.1007/s004410051284

Herzig, V. (2019). Arthropod assassins: Crawling biochemists with diverse toxin pharmacopeias. In *Toxicon*. https://doi.org/10.1016/j.toxicon.2018.11.312

Herzig, V., Bende, N. S., Alam, M. S., Tedford, H. W., Kennedy, R. M., & King, G. F. (2014). Methods for Deployment of Spider Venom Peptides as Bioinsecticides. In *Advances in Insect Physiology*. https://doi.org/10.1016/B978-0-12-800197-4.00008-7

Herzig, V., King, G. F., & Undheim, E. A. B. (2019). Can we resolve the taxonomic bias in spider venom research? *Toxicon: X*. https://doi.org/10.1016/j.toxcx.2018.100005

Holford, M., Daly, M., King, G. F., & Norton, R. S. (2018). Venoms to the rescue. *Science*.

https://doi.org/10.1126/science.aau7761

Hu, H., Bandyopadhyay, P. K., Olivera, B. M., & Yandell, M. (2011). Characterization of the Conus bullatus genome and its venom-duct transcriptome. *BMC Genomics*. https://doi.org/10.1186/1471-2164-12-60

Hüsser, M. (2018). A first phylogenetic analysis reveals a new arboreal tarantula genus from South America with description of a new species and two new species of Tapinauchenius Ausserer, 1871 (Araneae, mygalomorphae, theraphosidae). *ZooKeys*. https://doi.org/10.3897/zookeys.784.26521

Jenner, R. A., Von Reumont, B. M., Campbell, L. I., & Undheim, E. A. B. (2019). Parallel Evolution of Complex Centipede Venoms Revealed by Comparative Proteotranscriptomic Analyses. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msz181

John Jasys, V., Kelbaugh, P. R., Nason, D. M., Phillips, D., Rosnack, K. J., Saccomano, N. A., Stroh, J. G., & Volkmann, R. A. (1990). Isolation, Structure Elucidation, and Synthesis of Novel Hydroxylamine-Containing Polyamines from the Venom of the Agelenopsis aperta Spider. *Journal of the American Chemical Society*. https://doi.org/10.1021/ja00174a037

Juárez, P., Comas, I., González-Candelas, F., & Calvete, J. J. (2008). Evolution of snake venom disintegrins by positive Darwinian selection. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msn179

Jungo, F., Estreicher, A., Bairoch, A., Bougueleret, L., & Xenarios, I. (2010). Animal toxins: How is complexity represented in databases? In *Toxins*. https://doi.org/10.3390/toxins2020261

Kaderka, R., Bulantová, J., Heneberg, P., & Řezáč, M. (2019). Urticating setae of tarantulas (Araneae: Theraphosidae): Morphology, revision of typology and terminology and implications for taxonomy. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0224384

King, G. F. (2019). Tying pest insects in knots: the deployment of spider-venom-derived knottins as bioinsecticides. In *Pest Management Science*. https://doi.org/10.1002/ps.5452

King, G. F., & Hardy, M. C. (2013). Spider-Venom Peptides: Structure, Pharmacology, and Potential for Control of Insect Pests. *Annual Review of Entomology*. https://doi.org/10.1146/annurev-ento-120811-153650

Koludarov, I., Jackson, T. N. W., op den Brouw, B., Dobson, J., Dashevsky, D., Arbuckle, K., Clemente, C. J., Stockdale, E. J., Cochran, C., Debono, J., Stephens, C., Panagides, N., Li, B., Manchadi, M. L. R., Violette, A., Fourmy, R., Hendrikx, I., Nouwens, A., Clements, J., … Fry, B. G. (2017). Enter the dragon: The dynamic and multifunctional evolution of anguimorpha lizard venoms. *Toxins*. https://doi.org/10.3390/toxins9080242

Kordiš, D., & Gubenšek, F. (2000). Adaptive evolution of animal toxin multigene families. *Gene*. https://doi.org/10.1016/S0378-1119(00)00490-X

Kuhn-Nentwig, L., Langenegger, N., Heller, M., Koua, D., & Nentwig, W. (2019). The dual prey-inactivation strategy of spiders—in-depth venomic analysis of cupiennius salei. *Toxins*. https://doi.org/10.3390/toxins11030167

Kuhn-Nentwig, L., Stöcklin, R., & Nentwig, W. (2011). Venom composition and strategies in spiders. is everything possible? In *Advances in Insect Physiology*. https://doi.org/10.1016/B978-0-12-387668-3.00001-5

Langenegger, N., Koua, D., Schürch, S., Heller, M., Nentwig, W., & Kuhn-Nentwig, L. (2018). Identification of a precursor processing protease from the spider Cupiennius salei essential for venom neurotoxin maturation. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M117.810911

Langenegger, N., Nentwig, W., & Kuhn-Nentwig, L. (2019). Spider venom: Components, modes of action, and novel strategies in transcriptomic and proteomic analyses. In *Toxins*. https://doi.org/10.3390/toxins11100611

Lee, S. Y., & MacKinnon, R. (2004). A membrane-access mechanism of ion channel inhibition by voltage sensor toxins from spider venom. *Nature*. https://doi.org/10.1038/nature02632

Li, M., Fry, B. G., & Kini, R. M. (2005). Eggs-only diet: Its implications for the toxin profile changes and ecology of the marbled sea snake (Aipysurus eydouxii). *Journal of Molecular Evolution*. https://doi.org/10.1007/s00239-004-0138-0

Lindgardt, Z., Reeves, M., & Wallenstein, J. (2008). Waking the Giant: Business Model Innovation in the

Drug Industry. *In Vivo*.
Lynch, V. J. (2007). Inventing an arsenal: Adaptive evolution and neofunctionalization of snake venom phospholipase A2 genes. *BMC Evolutionary Biology*. https://doi.org/10.1186/1471-2148-7-2
Mahmoudi, N., Modanu, M., Brandt, Y., & Andrade, M. C. B. (2008). Subtle pedipalp dimorphism: a reliable method for sexing juvenile spiders. *Journal of Arachnology*. https://doi.org/10.1636/sh07-81.1
Mammola, S., Michalik, P., Hebets, E. A., & Isaia, M. (2017). Record breaking achievements by spiders and the scientists who study them. *PeerJ*. https://doi.org/10.7717/peerj.3972
Marshall, S., Raven, R., & Hoeh, W. (1999). A test of alternative phylogenetic relation- ship of theraphosid subfamilies using cladistic analysis of morphological traits. *Am Arachnol*, *60*(5).
McCarthy, C. A., Rash, L. D., Chassagnon, I. R., King, G. F., & Widdop, R. E. (2015). PcTx1 affords neuroprotection in a conscious model of stroke in hypertensive rats via selective inhibition of ASIC1a. *Neuropharmacology*. https://doi.org/10.1016/j.neuropharm.2015.08.040
McCowan, C., & Garb, J. E. (2014). Recruitment and diversification of an ecdysozoan family of neuropeptide hormones for black widow spider venom expression. *Gene*. https://doi.org/10.1016/j.gene.2013.11.054
Mendoza, J., & Francke, O. (2017). Systematic revision of Brachypelma red-kneed tarantulas (Araneae:Theraphosidae), and the use of DNA barcodes to assist in the identification and conservation of CITES-listed species. *Invertebrate Systematics*. https://doi.org/10.1071/IS16023
Mihailescu, M., Krepkiy, D., Milescu, M., Gawrisch, K., Swartz, K. J., & White, S. (2014). Structural interactions of a voltage sensor toxin with lipid membranes. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1415324111
Milne, T. J., Abbenante, G., Tyndall, J. D. A., Halliday, J., & Lewis, R. J. (2003). Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M304843200
Morgenstern, D., & King, G. F. (2013). The venom optimization hypothesis revisited. In *Toxicon*. https://doi.org/10.1016/j.toxicon.2012.11.022
Nagahama, R. H., Fukushima, C. S., & Bertani, R. (2009). Nhandu tripepii is a senior synonym of Nhandu vulpinus (Araneae: Theraphosidae). *Zoologia*. https://doi.org/10.1590/S1984-46702009000300025
Nentwig, W. (2013). Spider ecophysiology. In *Spider Ecophysiology*. https://doi.org/10.1007/978-3-642-33989-9
Ngamniyom, A., Manaboon, M., Panyarachun, B., & Showpittapornchai, U. (2014). Phylogenetic relationships of two earth tiger tarantulas, Haplopelma lividum and H. longipes (Araneae, Theraphosidae), within the infraorder Mygalomorph using 28S ribosomal DNA sequences. *International Journal of Zoological Research*. https://doi.org/10.3923/ijzr.2014.15.19
Norton, R. S., & Pallaghy, P. K. (1998). The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon*. https://doi.org/10.1016/S0041-0101(98)00149-4
Nyffeler, M., & Benz, G. (1981). Freilanduntersuchungen zur Nahrungsökologie der Spinnen: Beobachtungen aus der Region Zürich. *Anzeiger Für Schädlingskunde Pflanzenschutz Umweltschutz*. https://doi.org/10.1007/BF01905916
Nyffeler, Martin, Olson, E. J., & Symondson, W. O. C. (2016). Plant-eating by spiders. *Journal of Arachnology*. https://doi.org/10.1636/p15-45.1
Orlova, E. V., Atiqur Rahman, M., Gowen, B., Volynski, K. E., Ashton, A. C., Manser, C., Van Heel, M., & Ushkaryov, Y. A. (2000). Structure of α-latrotoxin oligomers reveals that divalent cation-dependent tetramers form membrane pores. *Nature Structural Biology*. https://doi.org/10.1038/71247
Ortiz, D., & Francke, O. F. (2016). Two DNA barcodes and morphology for multi-method species delimitation in Bonnetina tarantulas (Araneae: Theraphosidae). *Molecular Phylogenetics and Evolution*. https://doi.org/10.1016/j.ympev.2016.05.003
Ortiz, D., & Francke, O. F. (2017). Reconciling morphological and molecular systematics in tarantulas (Araneae: Theraphosidae): Revision of the Mexican endemic genus Bonnetina. *Zoological Journal of the Linnean Society*. https://doi.org/10.1093/zoolinnean/zlw013
Ortiz, D., Francke, O. F., & Bond, J. E. (2018). A tangle of forms and phylogeny: Extensive morphological homoplasy and molecular clock heterogeneity in Bonnetina and related tarantulas.

Molecular Phylogenetics and Evolution. https://doi.org/10.1016/j.ympev.2018.05.013

Özbek, R., Wielsch, N., Vogel, H., Lochnit, G., Foerster, F., Vilcinskas, A., & Von Reumont, B. M. (2019). Proteo-transcriptomic characterization of the venom from the endoparasitoid wasp pimpla turionellae with aspects on its biology and evolution. *Toxins*. https://doi.org/10.3390/toxins11120721

Panagides, N., Jackson, T. N. W., Ikonomopoulou, M. P., Arbuckle, K., Pretzler, R., Yang, D. C., Ali, S. A., Koludarov, I., Dobson, J., Sanker, B., Asselin, A., Santana, R. C., Hendrikx, I., van der Ploeg, H., Tai-A-Pin, J., van den Bergh, R., Kerkkamp, H. M. I., Vonk, F. J., Naude, A., … Fry, B. G. (2017). How the cobra got its flesh-eating venom: Cytotoxicity as a defensive innovation and its co-evolution with hooding, aposematic marking, and spitting. *Toxins*. https://doi.org/10.3390/toxins9030103

Park, S. P., Kim, B. M., Koo, J. Y., Cho, H., Lee, C. H., Kim, M., Na, H. S., & Oh, U. (2008). A tarantula spider toxin, GsMTx4, reduces mechanical and neuropathic pain. *Pain*. https://doi.org/10.1016/j.pain.2008.02.013

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., & Schacht, A. L. (2010). How to improve RD productivity: The pharmaceutical industry's grand challenge. In *Nature Reviews Drug Discovery*. https://doi.org/10.1038/nrd3078

Pekár, S. (2004). Predatory behaviour of two european ant-eating spiders (Araneae, Zodariidae). *Journal of Arachnology*. https://doi.org/10.1636/s02-15

Pekár, S. (2014). Comparative analysis of passive defences in spiders (Araneae). *Journal of Animal Ecology*. https://doi.org/10.1111/1365-2656.12177

Pekár, S., & Toft, S. (2015). Trophic specialisation in a predatory group: The case of prey-specialised spiders (Araneae). *Biological Reviews*. https://doi.org/10.1111/brv.12133

Pennisi, E. (2017). Untangling spider biology. *Science*. https://doi.org/10.1126/science.358.6361.288

Pérez-Miles, F., & Perafán, C. (2015). Geographic patterns of abdominal urticating setae types in Neotropical tarantulas (Araneae, Theraphosidae). *Bol.Soc.Zool. Uruguay*.

Perumal Samy, R., Stiles, B. G., Franco, O. L., Sethi, G., & Lim, L. H. K. (2017). Animal venoms as antimicrobial agents. In *Biochemical Pharmacology*. https://doi.org/10.1016/j.bcp.2017.03.005

Piel, W. H. (2018). The global latitudinal diversity gradient pattern in spiders. *Journal of Biogeography*. https://doi.org/10.1111/jbi.13387

Pineda, S. S., Chaumeil, P. A., Kunert, A., Kaas, Q., Thang, M. W. C., Le, L., Nuhn, M., Herzig, V., Saez, N. J., Cristofori-Armstrong, B., Anangi, R., Senff, S., Gorse, D., & King, G. F. (2018). ArachnoServer 3.0: An online resource for automated discovery, analysis and annotation of spider toxins. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btx661

Pineda, S. S., Undheim, E. A. B., Rupasinghe, D. B., Ikonomopoulou, M. P., & King, G. F. (2014). Spider venomics: Implications for drug discovery. In *Future Medicinal Chemistry*. https://doi.org/10.4155/FMC.14.103

Řezáč, M., Pekár, S., & Lubin, Y. (2008). How oniscophagous spiders overcome woodlouse armour. *Journal of Zoology*. https://doi.org/10.1111/j.1469-7998.2007.00408.x

Richards, K. L., Milligan, C. J., Richardson, R. J., Jancovski, N., Grunnet, M., Jacobson, L. H., Undheim, E. A. B., Mobli, M., Chow, C. Y., Herzig, V., Csoti, A., Panyi, G., Reid, C. A., King, G. F., & Petrou, S. (2018). Selective NaV1.1 activation rescues Dravet syndrome mice from seizures and premature death. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1804764115

Riechert, S. E., & Singer, F. D. (1995). Investigation of potential male mate choice in a monogamous spider. *Animal Behaviour*. https://doi.org/10.1016/0003-3472(95)80204-5

Robert, R. (1985). The spider infraorder Mygalomorphae (Araneae): cladistics and system- atics. *Bull. Am. Mus. Nat. Hist*, *182*, 1–180.

Robert, R. (1990). Comments on the proposed precedence of Aphonopelma Pocock 1901 (Arachnida, Araneae) over Rechostica Simon, 1892. *Bull. Zool. Nom.*, *42*(126).

Roelants, K., Fry, B. G., Ye, L., Stijlemans, B., Brys, L., Kok, P., Clynen, E., Schoofs, L., Cornelis, P., & Bossuyt, F. (2013). Origin and Functional Diversification of an Amphibian Defense Peptide Arsenal. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1003662

Ryu, J. H., Jung, H. J., Konishi, S., Kim, H. H., Park, Z. Y., & Kim, J. Il. (2017). Structure-activity relationships of ω-Agatoxin IVA in lipid membranes. *Biochemical and Biophysical Research Communications*. https://doi.org/10.1016/j.bbrc.2016.11.025

Sachkova, M. Y., Slavokhotova, A. A., Grishin, E. V., & Vassilevski, A. A. (2014). Genes and evolution of two-domain toxins from lynx spider venom. *FEBS Letters*. https://doi.org/10.1016/j.febslet.2014.01.018

Saez, N. J., & Herzig, V. (2019). Versatile spider venom peptides and their medical and agricultural applications. *Toxicon*. https://doi.org/10.1016/j.toxicon.2018.11.298

Saez, N. J., Senff, S., Jensen, J. E., Er, S. Y., Herzig, V., Rash, L. D., & King, G. F. (2010). Spider-venom peptides as therapeutics. In *Toxins*. https://doi.org/10.3390/toxins2122851

Samm, R., & Schmidt, G. (2010). Psalmopoeinae subfamilia nov. - eine neue Unterfamilie der Theraphosidae (Araneae). *Tarantulas World*, *142*, 35–41.

Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. In *Biological Conservation*. https://doi.org/10.1016/j.biocon.2019.01.020

Sanggaard, K. W., Bechsgaard, J. S., Fang, X., Duan, J., Dyrlund, T. F., Gupta, V., Jiang, X., Cheng, L., Fan, D., Feng, Y., Han, L., Huang, Z., Wu, Z., Liao, L., Settepani, V., Thøgersen, I. B., Vanthournout, B., Wang, T., Zhu, Y., … Wang, J. (2014). Spider genomes provide insight into composition and evolution of venom and silk. *Nature Communications*. https://doi.org/10.1038/ncomms4765

Schendel, V., Rash, L. D., Jenner, R. A., & Undheim, E. A. B. (2019). The diversity of venom: The importance of behavior and venom system morphology in understanding its ecology and evolution. In *Toxins*. https://doi.org/10.3390/toxins11110666

Schmidt, G. (1995). Die Stellung der Gattung Poecilotheria im System. *Arachnida*, *10*, 1–2.

Schmidt, G. (2002). Gehören Brachionopus Pocock, 1897 und Harpactirella Purcell, 1902 zu den Theraphosiden? *Arthropoda*, *10*, 12–17.

Schmidt, G. (2003). *Die Vogelspinnen* (1st ed.). Die neue Brehm-Bücherei.

Schwager, E. E., Sharma, P. P., Clarke, T., Leite, D. J., Wierschin, T., Pechmann, M., Akiyama-Oda, Y., Esposito, L., Bechsgaard, J., Bilde, T., Buffry, A. D., Chao, H., Dinh, H., Doddapaneni, H. V., Dugan, S., Eibner, C., Extavour, C. G., Funch, P., Garb, J., … McGregor, A. P. (2017). The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biology*. https://doi.org/10.1186/s12915-017-0399-x

Selden, P. A., & Penney, D. (2010). Fossil spiders. *Biological Reviews*. https://doi.org/10.1111/j.1469-185X.2009.00099.x

Sharma, P. P. (2018). Chelicerates. In *Current Biology*. https://doi.org/10.1016/j.cub.2018.05.036

Silva, L. M., Carvalho Botelho, A. C., Nacif-Pimenta, R., Martins, G. F., Alves, L. C., Brayner, F. A., Fortes-Dias, C. L., & Paolucci Pimenta, P. F. (2008). Structural analysis of the venom glands of the armed spider Phoneutria nigriventer (Keyserling, 1891): Microanatomy, fine structure and confocal observations. *Toxicon*. https://doi.org/10.1016/j.toxicon.2007.12.009

Sunagar, K., Jackson, T. N. W., Undheim, E. A. B., Ali, S. A., Antunes, A., & Fry, B. G. (2013). Three-fingered RAVERs: Rapid Accumulation of Variations in Exposed Residues of snake venom toxins. *Toxins*. https://doi.org/10.3390/toxins5112172

Sunagar, K., & Moran, Y. (2015). The Rise and Fall of an Evolutionary Innovation: Contrasting Strategies of Venom Evolution in Ancient and Young Animals. *PLoS Genetics*. https://doi.org/10.1371/journal.pgen.1005596

Swanson, D. L., & Vetter, R. S. (2006). Loxoscelism. *Clinics in Dermatology*. https://doi.org/10.1016/j.clindermatol.2005.11.006

Teyssie, F. (2015). *Tarantulas of the World*. NAP Editions.

Turner, S. P., Longhorn, S. J., Hamilton, C. A., Gabriel, R., PÉrez-Miles, F., & Vogler, A. P. (2018). Re-evaluating conservation priorities of New World tarantulas (Araneae: Theraphosidae) in a molecular framework indicates non-monophyly of the genera, Aphonopelma and Brachypelma. *Systematics and Biodiversity*. https://doi.org/10.1080/14772000.2017.1346719

Tzouros, M., Chesnov, S., Bigler, L., & Bienz, S. (2013). A template approach for the characterization of linear polyamines and derivatives in spider venom. *European Journal of Mass Spectrometry*.

https://doi.org/10.1255/ejms.1213

Undheim, E. A. B., Grimm, L. L., Low, C. F., Morgenstern, D., Herzig, V., Zobel-Thropp, P., Pineda, S. S., Habib, R., Dziemborowicz, S., Fry, B. G., Nicholson, G. M., Binford, G. J., Mobli, M., & King, G. F. (2015). Weaponization of a Hormone: Convergent Recruitment of Hyperglycemic Hormone into the Venom of Arthropod Predators. *Structure*. https://doi.org/10.1016/j.str.2015.05.003

Undheim, E. A. B., Sunagar, K., Herzig, V., Kely, L., Low, D. H. W., Jackson, T. N. W., Jones, A., Kurniawan, N., King, G. F., Ali, S. A., Antunes, A., Ruder, T., & Fry, B. G. (2013). A Proteomics and Transcriptomics investigation of the venom from the Barychelid spider Trittame loki (brush-foot trapdoor). *Toxins*. https://doi.org/10.3390/toxins5122488

Ushkaryov, Y. A., Rohou, A., & Sugita, S. (2008). *a-Latrotoxin and Its Receptors*. https://doi.org/10.1007/978-3-540-74805-2_7

Valencia-Cuéllar, D., Perafán, C., Guerrero, R. J., & Leite Guadanucci, J. P. (2019). Schismatothelinae spiders (Araneae, Mygalomorphae, Theraphosidae) from Colombia: Four new species and an approach to their diversity. *Zootaxa*. https://doi.org/10.11646/zootaxa.4545.4.6

Valente, R. H., Sakai, F., Portes-Junior, J. A., Viana, L. G., Carneiro, S. M., Perales, J., & Yamanouye, N. (2018). The primary duct of bothrops jararaca glandular apparatus secretes toxins. *Toxins*. https://doi.org/10.3390/toxins10030121

Van Valen, L. (1974). Molecular evolution as predicted by natural selection. *Journal of Molecular Evolution*. https://doi.org/10.1007/BF01796554

Vollrath, F. (1999). Biology of spider silk. *International Journal of Biological Macromolecules*. https://doi.org/10.1016/S0141-8130(98)00076-2

von Reumont, Bjoern Marcus, Campbell, L. I., & Jenner, R. A. (2014). Quo Vadis venomics? A roadmap to neglected venomous invertebrates. In *Toxins*. https://doi.org/10.3390/toxins6123488

Von Reumont, Björn M., Blanke, A., Richter, S., Alvarez, F., Bleidorn, C., & Jenner, R. A. (2014). The first venomous crustacean revealed by transcriptomics and functional morphology: Remipede venom glands express a unique toxin cocktail dominated by enzymes and a neurotoxin. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/mst199

Vonk, F. J., Admiraal, J. F., Jackson, K., Reshef, R., De Bakker, M. A. G., Vanderschoot, K., Van Den Berge, I., Van Atten, M., Burgerhout, E., Beck, A., Mirtschin, P. J., Kochva, E., Witte, F., Fry, B. G., Woods, A. E., & Richardson, M. K. (2008). Evolutionary origin and development of snake fangs. *Nature*. https://doi.org/10.1038/nature07178

Vonk, F. J., Casewell, N. R., Henkel, C. V., Heimberg, A. M., Jansen, H. J., McCleary, R. J. R., Kerkkamp, H. M. E., Vos, R. A., Guerreiro, I., Calvete, J. J., Wüster, W., Woods, A. E., Logan, J. M., Harrison, R. A., Castoe, T. A., De Koning, J., Pollock, D. D., Yandell, M., Calderon, D., … Richardson, M. K. (2013). The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1314702110

Warren, W. C., Hillier, L. D. W., Marshall Graves, J. A., Birney, E., Ponting, C. P., Grützner, F., Belov, K., Miller, W., Clarke, L., Chinwalla, A. T., Yang, S. P., Heger, A., Locke, D. P., Miethke, P., Waters, P. D., Veyrunes, F., Fulton, L., Fulton, B., Graves, T., … Wilson, R. K. (2008). Genome analysis of the platypus reveals unique signatures of evolution. *Nature*. https://doi.org/10.1038/nature06936

Welke, K. W., & Schneider, J. M. (2012). Sexual cannibalism benefits offspring survival. *Animal Behaviour*. https://doi.org/10.1016/j.anbehav.2011.10.027

Weng, J. L., Barrantes, G., & Eberhard, W. G. (2006). Feeding by Philoponella vicina (Araneae, Uloboridae) and how uloborid spiders lost their venom glands. *Canadian Journal of Zoology*. https://doi.org/10.1139/Z06-149

West, R. C., Marshall, S. D., Fukushima, C. S., & Bertani, R. (2008). Review and cladistic analysis of the Neotropical tarantula genus Ephebopus Simon 1892 (Araneae: Theraphosidae) with notes on the Avicullariinae. *Zootaxa*. https://doi.org/10.11646/zootaxa.1849.1.3

Wigger, E., Kuhn-Nentwig, L., & Nentwig, W. (2002). The venom optimisation hypothesis: A spider injects large venom quantities only into difficult prey types. *Toxicon*. https://doi.org/10.1016/S0041-0101(01)00277-X

Wilson, D., Boyle, G. M., McIntyre, L., Nolan, M. J., Parsons, P. G., Smith, J. J., Tribolet, L., Loukas, A.,

Liddell, M. J., Rash, L. D., & Daly, N. L. (2017). The aromatic head group of spider toxin polyamines influences toxicity to cancer cells. *Toxins*. https://doi.org/10.3390/toxins9110346

Wilson, J. S., Gunnell, C. F., Wahl, D. B., & Pitts, J. P. (2013). Testing the species limits of the tarantulas (Araneae: Theraphosidae) endemic to California's Southern Coast Ranges, USA. *Insect Conservation and Diversity*. https://doi.org/10.1111/icad.12000

Windley, M. J., Herzig, V., Dziemborowicz, S. A., Hardy, M. C., King, G. F., & Nicholson, G. M. (2012). Spider-venom peptides as bioinsecticides. In *Toxins*. https://doi.org/10.3390/toxins4030191

Wong, E. S. W., & Belov, K. (2012). Venom evolution through gene duplications. In *Gene*. https://doi.org/10.1016/j.gene.2012.01.009

World Spider Catalog. (2019). World Spider Catalog Version 20.5. *Natural History Museum Bern*. https://doi.org/10.24436/2

Yan, L., & Adams, M. E. (1998). Lycotoxins, antimicrobial peptides from venom of the wolf spider Lycosa carolinensis. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.273.4.2059

Yap, M. K. K., & Misuan, N. (2019). Exendin-4 from Heloderma suspectum venom: From discovery to its latest application as type II diabetes combatant. In *Basic and Clinical Pharmacology and Toxicology*. https://doi.org/10.1111/bcpt.13169

Yeargan, K. V. (1988). Ecology of a bolas spider, Mastophora hutchinsoni: phenology, hunting tactics, and evidence for aggressive chemical mimicry. *Oecologia*. https://doi.org/10.1007/BF00380049

Yiğit, N., Güven, T., Bayram, A., & Çavuşoğlu, K. (2004). A morphological study on the Venom apparatus of the spider Agelena labyrinthica (Araneae, Agelenidae). *Turkish Journal of Zoology*.

Zobel-Thropp, P. A., Mullins, J., Kristensen, C., Kronmiller, B. A., David, C. L., Breci, L. A., & Binford, G. J. (2019). Not so Dangerous After All? Venom Composition and Potency of the Pholcid (Daddy Long-Leg) Spider Physocyclus mexicanus. *Frontiers in Ecology and Evolution*. https://doi.org/10.3389/fevo.2019.00256

V Published works

Chapter I

Discovering the silk road: Nuclear and mitochondrial sequence data resolve the phylogenetic relationships among theraphosid spider subfamilies

Tim Lüddecke, Henrik Krehenwinkel, Gregory Canning, Frank Glaw, Stuart J. Longhorn, René Tänzler, Ingo Wendt & Miguel Vences
Molecular Phylogenetics and Evolution (2018) 119: 63-70.

Molecular Phylogenetics and Evolution 119 (2018) 63–70

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Discovering the silk road: Nuclear and mitochondrial sequence data resolve the phylogenetic relationships among theraphosid spider subfamilies

Tim Lüddecke[a,*], Henrik Krehenwinkel[b], Gregory Canning[c], Frank Glaw[d], Stuart J. Longhorn[e], René Tänzler[d], Ingo Wendt[f], Miguel Vences[a]

[a] Zoological Institute, Technische Universität Braunschweig, Mendelssohnstr. 4, 38106 Braunschweig, Germany
[b] Department of Environmental Science, Policy and Management, University of California Berkeley, 130 Mulford Hall #3114, Berkeley, CA 94720-3114, USA
[c] Department of Nature Conservation, Tshwane University of Technology, P. Bag X680, Pretoria 0001, South Africa
[d] Zoologische Staatssammlung München (ZSM-SNSB), Münchhausenstr. 21, 81247 München, Germany
[e] Hope Entomological Collections, Oxford University Museum of Natural History (OUMNH), Parks Road, Oxford, England OX1 3PW, United Kingdom
[f] Staatliches Museum für Naturkunde Stuttgart, Rosenstein 1, 70191 Stuttgart, Germany

ARTICLE INFO

Keywords:
Arachnida
Mygalomorphae
Theraphosidae
Molecular phylogeny

ABSTRACT

The mygalomorph spiders in the family Theraphosidae, also known as "tarantulas", are one of the most popular and diverse groups of arachnids, but their evolutionary history remains poorly understood because morphological analyses have only provided mostly controversial results, and a broad molecular perspective has been lacking until now. In this study we provide a preliminary molecular phylogenetic hypothesis of relationships among theraphosid subfamilies, based on 3.5 kbp of three nuclear and three mitochondrial markers, for 52 taxa representing 10 of the 11 commonly accepted subfamilies. Our analysis confirms the monophyly of the Theraphosidae and of most recognized theraphosid subfamilies, supports the validity of the Stromatopelminae and Poecilotheriinae, and indicates paraphyly of the Schismatothelinae. The placement of representatives of Schismatothelinae also indicates possible non-monophyly of Avicullariinae and supports the distinction of the previously contentious subfamily Psalmopoeinae. Major clades typically corresponded to taxa occurring in the same biogeographic region, with two of them each occurring in Africa, South America and Asia. Because relationships among these major clades were poorly supported, more extensive molecular data sets are required to test the hypothesis of independent colonization and multiple dispersal events among these continents.

1. Introduction

The family Theraphosidae, commonly referred to as "tarantulas", are a group of mygalomorph spiders which includes some of the largest arachnids on earth. Tarantulas are among the most prominent spiders, widely known to the general public as horror movie performers and also commonly kept as pets (e.g. Klaas, 1989). Given this distinguished position among spiders, it is striking that their deep phylogeny and higher classification remain unstudied from a molecular perspective. The Theraphosidae contain 962 nominal species (World Spider Catalogue, 2017), and most authors accept 11 theraphosid subfamilies (Kambas, 2017) occurring on all continents except Antarctica (West et al., 2008) (Table 1). This diversity is also reflected in different ecological adaptations and natural histories. The group includes arboreal, terrestrial and cave dwelling taxa. Different theraphosid species display a variety of defensive mechanisms such as stridulation, venomous bite or urticating setae (Marshall et al., 1995; Bertani and Guadanucci, 2013). Furthermore, considering their wide range of distribution and their presumably poor dispersal capacity, theraphosids represent good models for biogeographic research.

As is commonplace in arachnological taxonomy (Coddington, 2005; Ramirez, 2014), the systematics of theraphosid spiders is largely based on analysis of morphological characters. Using those, Raven (1985) formulated a first evolutionary hypothesis on intrafamilial relationships, but later recognized the limits of this morphology-based approach by referring to Theraphosidae as a "taxonomic nightmare" (Raven, 1990). Subsequent revisions of theraphosid subfamilies and genera continued using mostly or only morphological characters (e.g. Smith, 1990; Pérez-Miles et al., 1996; Bertani, 2000; Gallon, 2003, 2005a; West et al., 2008, 2012; West and Nunn, 2010; Guadanucci, 2011a, 2011b, 2014; Fukushima and Bertani, 2017), but other studies have shown that many traditional morphological characters of spiders can be affected by a high degree of homoplasy (Platnick and Gertsch, 1976; Goloboff, 1993; Bond and Opell, 2002; Hedin and Bond, 2006;

* Corresponding author.
E-mail address: t.lueddecke@tu-braunschweig.de (T. Lüddecke).

https://doi.org/10.1016/j.ympev.2017.10.015
Received 24 May 2017; Received in revised form 14 October 2017; Accepted 22 October 2017
Available online 02 November 2017

Table 1
Overview of subfamily classification within the Theraphosidae, including geographic range, species numbers, and representatives included in this study. Classification at subfamily and genus level according to the World Spider Catalogue (2017) with modifications as explained in the text.

Subfamily	Distribution range	Number of species	Genera included in this study
Avicularinae	Mexico, Caribbean, Central- South America	86	*Avicularia, Caribena, Psalmopoeus, Tapinauchenius, Ybyrapora*
Eumenophorinae**	Western-Eastern Africa, Madagascar, Yemen, Indian Ocean Archipelagos, India**	64	*Monocentropus, Pelinobius, Hysterocrates*
Harpactirinae	Central-Southern Africa	61	*Augacephalus, Brachionopus**, *Harpactira, Harpactirella**, *Pterinochilus*
Ischnocolinae*	Africa, Asia, Europe, Caribbean, Central-South America	86	*Nesiergus*
Ornithoctoninae	S.E. Asia	27	*Cyriopagopus, Omothymus, Lampropelma*
Schismatothelinae	South America	13	*Euthycaelus, Neoholothele*
Selenocosmiinae	Indian subcontinent, S.E. Asia, Australasia	123	*Chilobrachys, Haplocosmia, Lyrognathus, Orphnaecus, Phlogiellus, Poecilotheria, Selenocosmia*
Selenogyrinae	Western Africa	7	Not included
Stromatopelminae	Western-Central Africa	10	*Heteroscodra, Stromatopelma*
Theraphosinae	North America (USA/ Mexico), Caribbean, Central-South America	489	*Aphonopelma, Brachypelma, Crassicrus, Grammostola, Kochiana, Lasiodora, Lasiodorides, Megaphobema, Nhandu, Sericopelma, Xenesthis*
Thrigmopoeinae	Indian subcontinent	10	*Thrigmopoeus*

* Likely polyphyletic, including Acanthopelmatinae Smith, 1994 and Trichopelmatinae Raven, 1985, following Guadanucci (2014).
** Including *Annandaliella* Hirst, 1909, which different authors have placed in various subfamilies.
*** In some studies placed outside Theraphosidae, e.g. Schmidt (2002) in Barychelidae.

Ayoub et al., 2007; Bond et al., 2012). Within Theraphosidae, the poorly defined subfamily Ischnocolinae comprises several taxa of uncertain taxonomical status and has been found to be paraphyletic (Raven, 1985). Several ischnocoline genera show an unresolved phylogenetic position, e.g. *Nesiergus*, which has been removed from Ischnocolinae (sensu strictu) by Guadanucci (2014) but not given an alternative (therefore herein referred to the Ischnocolinae sensu lato). Besides the Ischnocolinae conundrum, the monophyly and validity of several other subfamilies are also under discussion. Particularly, this regards the Selenocosmiinae (West and Nunn, 2010), Avicularinae (West et al., 2008; Fukushima and Bertani, 2017) and Eumenophorinae (Guadanucci, 2011a, 2011b; Mirza et al., 2014), as well as the validity of Poecilotheriinae (Schmidt, 1995; West et al., 2008, 2012) and Psalmopoeinae (Samm and Schmidt, 2010).

Comparative molecular studies on theraphosids thus far are limited to the species to genus level, for instance in *Aphonopelma* (Hamilton et al., 2011, 2016), *Brachypelma* (Longhorn et al., 2007; Mendoza and Francke, 2017), *Grammostola* (Montes De Oca et al., 2015) and *Bonnetina* (Ortiz and Francke, 2016), whereas phylogenetic relationships among genera and subfamilies remain significantly understudied from a molecular perspective, with few exceptions (e.g., Turner et al., 2017). Several recent studies comprehensively addressed the phylogeny of spiders in general (e.g. Fernandez et al., 2014; Wheeler et al., 2016; Garrison et al., 2016), including representatives of the Theraphosidae but not addressing the phylogeny within the family.

In this study, our goal is to provide a first molecular hypothesis on the broadest scale of theraphosid evolutionary history, i.e., on the relationships among major clades at the subfamily level, serving as a baseline to direct future in-depth studies. We discuss our phylogeny in terms of genus to subfamily level classification. Furthermore, we examine the geographic distribution of major clades identified.

2. Materials and methods

2.1. Species sampling

Material was either obtained from wild caught specimens, museum collections or from the pet trade. For most samples we preserved voucher specimens, labeled with preliminary laboratory numbers (TP – Theraphosidae Project); specimens were deposited in the Zoologische Staatssammlung München (ZSM) under ZSM A20170052-A20170080, and in the Oxford University Museum of Natural History (OUMNH), plus pending deposition (Supplementary Table S1). Tissues were preserved in pure ethanol or RNAlater. DNA was obtained mostly from legs and in exceptional cases from the pedipalps or prosoma. From few specimens, only dried samples or molts from individuals collected in the wild were available.

2.2. Molecular methods

Tissues were digested by Proteinase K (10 mg/ml) at 37 °C for 24 hours, and DNA extracted following a standard salt-extraction protocol (Bruford et al., 1992). We targeted three nuclear and three mitochondrial markers using both, already available and newly designed primers. Sequences of primers used in this study are given in Table 2. A segment of the cytochrome oxidase subunit I (COI) was targeted with the primers COI.osm-for with HCO.osm-2198. A stretch comprising nearly the complete 12S and 16S rRNA genes (as well as intervening tRNAVal) was amplified and sequenced with a subset of newly developed primers, designed by aligning the complete mitochondrial genome of *Ornithoctonus huwena* (obtained from GenBank, Supplementary Table S2) to 12S and 16S sequences of other spiders and to existing primers for these genes for vertebrates (e.g., Palumbi et al., 1991) in MEGA v. 7 (Kumar et al., 2016). The existing primers were subsequently modified to fit the respective spider sequences in the alignment. For the 12S-16S stretch, we performed multiple amplifications for all samples, with all possible primer combinations to ensure successful amplification for a maximum number of samples. The newly designed primer pairs for the 12S rRNA gene were either 12SAL-Spider or 12SAL-Tarantula in combination with the reverse primer 12SBH-Spider. The 12S–tRNAVal-16S region was amplified and sequenced with 12SBL-Spider and 16SAH-Spider. The remaining part of the 16S rRNA gene was amplified with 16SAL-Spider or 16SAL-Tarantula F1 with either 16SBH-Spider, 16S-Tarantula-R1 or 16S-Tarantula-R2 as reverse primers. From the nuclear genome we amplified part of the gene for the Histone H3 Protein with H3F and H3R. We also targeted the 18S rRNA gene with 18SA and 9r, and a fragment of the 28S rRNA gene with 28SO and 28SC.

Polymerase chain reactions were performed with 0.3 µM of each primer in 0.25 mM dNTPs, 0.4 U GoTaq and 1.25x Reaction buffer (Promega) in a total volume of 10 µl. The obtained PCR products were purified with a combination of Exonuclease I and Shrimp Alkaline Phosphatase (ExoSAP) following the manufacturer's recommendations (NEB). Following purification, we directly sequenced the PCR products on an ABI 3130xl automated DNA sequencer (Applied Biosystems) with dye-labeled dideoxy terminator chemistry. We checked chromatograms in CodonCode Aligner 4.0.4 (CodonCode Corporation) to trim poor-

Table 2
Overview of Primerpairs used for amplification and sequencing of gene sequences in this study.

Gene	Location	Primer	Direction	Sequence (5→3)	Reference
12S	Mitochondrial	12SAL-Spider	FOR	AAMWAGGATTAGADACCCT	*
		12SAL-Tarantula	FOR	GACAAGGATTAGATACCCTTTTAT	*
		12SBH-Spider	REV	RAGGGTGACGGGCGATATGT	*
12S–tRNAVal-16S region	Mitochondrial	12SBL-Spider	FOR	ACATATCGCCCGTCACCCTY	*
		16SAH-Spider	REV	TARAAATGTTTTTGRTAAACAG	*
16S	Mitochondrial	16SAL-Spider	FOR	CTGTTTAYCAAAAACATTTYTA	*
		16SAL-Tarantula F1	FOR	GTGCTAAGGTAGCAYAAT	*
		16SBH-Spider	REV	CCGGTCTGAACTCAAATCATGT	*
		16S-Tarantula-R1	REV	AAAGTCGAACAGACTTTC	*
		16S-Tarantula-R2	REV	TAATTCAACATCGAGGTC	*
Cytochrome oxidase subunit I	Mitochondrial	COI.osm-for	FOR	TACTAGGAGCCCCTGATATAGC	**
		HCO.osm-2198	REV	TAAACTTCTGGATGTCCAAAGAATCA	**
18S	Nuclear	18SA	FOR	ATTAAAGTTGTTGCGGTTA	Whiting et al. (1997)
		9r	REV	GATCCTTCCGCAGGTTCACCTAC	Giribet et al. (1999)
28S	Nuclear	28SO	FOR	TCGGAAGGAACCAGCTACTA	Whiting et al. (1997)
		28SC	REV	GAAACTGCTCAAAGGTAAACGG	Hedin and Maddison (2001)
Histone H3	Nuclear	H3F	FOR	ATGGCTCGTACCAAGCAGACVG	Colgan et al. (1998)
		H3R	REV	ATGGCTCGTACCAAGCAGACVGC	Colgan et al. (1998)

* New designed primers for this study.
** Courtesy of S. Hauswaldt.

quality stretches and manually correct obvious errors. All new sequences were submitted to GenBank (accession numbers MG273461 to MG273647). For outgroups, sequences were newly generated for taxa belonging to the mygalomorph families of Dipluridae (*Linothele* sp. *"fallax"*) and Nemesiidae (*Acanthogonatus* sp.), and complemented with nemesiid sequences (*Acanthogonatus campanae*) from GenBank (see Supplementary Table S2).

2.3. Alignments, models of molecular evolution and phylogenetic reconstruction

Sequences were aligned in MEGA v. 7 (Kumar et al., 2016) manually or via the MUSCLE algorithm (Edgar, 2004a, 2004b) and exploratory phylogenetic trees reconstructed from single-gene alignments under the Maximum Likelihood criterion and the inferred best fitting model, with 100 bootstrap replicates (Supplementary Figs. S3–S7). For the ribosomal RNA genes, amplification was typically successful with the majority of primer combinations. Sequences were usually of high quality (clean chromatograms with distinct and evenly spaced peaks), and the amount of missing data for these markers is low in our data set. Amplification failure was restricted almost completely to the protein coding gene segments, H3 and COI. For these markers, manual correction and trimming of poor-quality stretches of sequences was sometimes necessary and the amount of missing data is therefore higher.

For the concatenated alignment, hypervariable regions of the ribosomal gene segments were excluded as suggested by an analysis with the Gblocks v. 0.91b online tool (Castresana, 2000) under the most stringent set of conditions. Partition schemes and models of molecular evolution were inferred using PartitionFinder v. 1.0.1 (Lanfear et al., 2012) using the Bayesian information criterion (BIC). Phylogenetic analysis of the concatenated dataset under Bayesian Inference (BI) was carried out in MrBayes v. 3.2 (Ronquist et al., 2012). We ran two simultaneous independent analyses with four Markov chains each (three heated and one cold) for 50 million generations, sampling every 1000th tree. We assessed chain mixing and stationarity by examining the standard deviation of split frequencies and by plotting –lnL per generation using Tracer 1.5 software (Rambaut and Drummond, 2007), and applied a 25% burn-in. From the remaining trees a majority rule consensus tree was constructed. In addition, we analyzed our sequence data under the Maximum Likelihood criterion (ML) in RAxML v. 8.2.9 (Stamatakis, 2014) by submission to the Cipres science gateway version 3.3 (Miller et al., 2010) with 500 rapid bootstrap iterations. Phylogenetic analysis under the Maximum Parsimony (MP) criterion was carried out with PAUP v. 4b1 (Swofford, 2002), calculating 2000 bootstrap replicates each with ten random addition sequence replicates and tree-bisection and reconnection branch swapping.

2.4. Verification of taxon identities

A large proportion of sample material used in this study was taken from the pet trade. To minimize the possibility of misidentifications, we re-evaluated or verified the taxonomic identity of all material by using appropriate morphology-based identification schemes (e.g. Smith, 1990; Schmidt, 2003; Teyssié, 2015). Furthermore, we used a DNA barcoding approach to provide additional insights about the taxonomic identity of our material: we downloaded COI sequences of theraphosid and nemesiid genera from GenBank and combined them with sequences from our material in a neighbor-joining tree calculated in MEGA 7 (Supplementary Table S2 and Fig. S8). All GenBank-mined sequences clustered with our own sequences from the same genera and/ or species (with the single exception of *Brachypelma* which has been found elsewhere as non-monophyletic, see Turner et al., 2017), thus enhancing our confidence their identification fits with current taxonomic schemes.

3. Results

Topologies of exploratory ML gene trees (see Supplementary Materials, Figs. S3–S7) did not present any major and strongly supported disagreements. The concatenated multi-gene alignment had a total length of 3498 bp for 52 taxa. Sequences of nuclear origin, consisting of the genes for 28S (522 bp), 18S (1024 bp) rRNA and Histone H3 (204 bp), represented 50% (1750 bp) of the total alignment. The remaining 50% (1748 bp) were of mitochondrial origin, including fragments of 12S/16S rRNA (1463 bp) and COI (285 bp). Combined phylogenetic analysis of the concatenated data as represented by our favoured BI tree (Fig. 1) revealed a number of major clades of which many were congruently supported by high to very high support values from all three tree reconstruction methods (BI, ML, MP), mostly according to various taxonomic subfamilies (see also Table 1) but support for many of the broader relationships among these major clades was unsatisfying. In particular, our combined analyses confirm monophyletic groups containing the included taxa, respectively, of the following theraphosid subfamilies: Harpactirinae (BI 100%, ML 98%, MP 94%), Stromatopelminae (BI 100%, ML 100%, MP 88%), Ornithoctoninae (BI 100%, ML 100%, MP 100%), Theraphosinae (BI 100%, ML

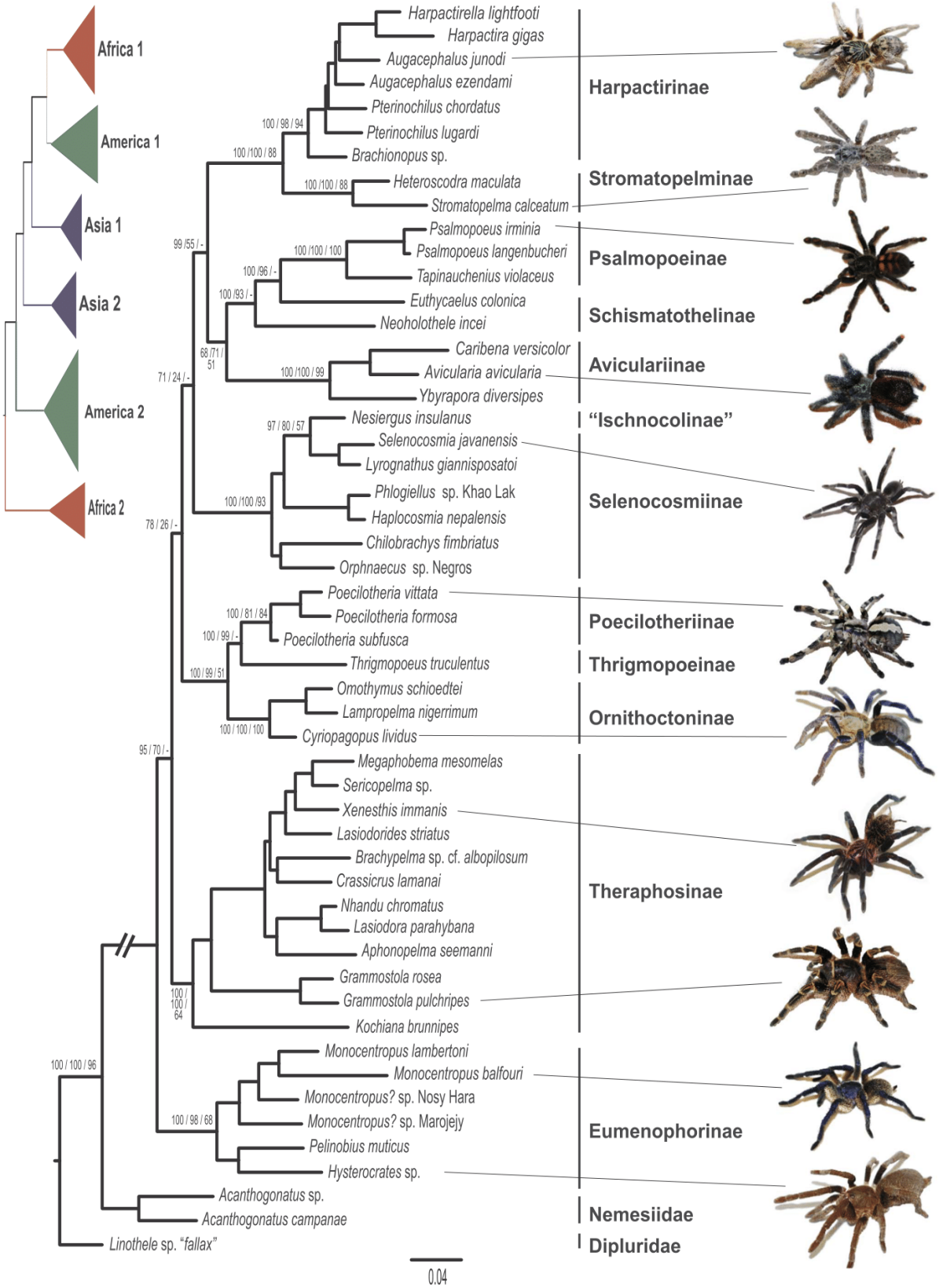


Fig. 1. Majority-rule consensus tree obtained from a Bayesian Inference (BI) analysis of a concatenated matrix of 3498 bp from three mitochondrial and three nuclear gene segments. Numbers at nodes represent support values (all in percent) from BI (posterior probabilities), and from separate ML and MP bootstrap analyses. Numbers are only shown for support values > 90% (BI) and > 50% (bootstrap). Bars indicate subfamily classification as supported by this study. The upper left inset picture indicates geographical distribution of major clades; inset photos depict typical representatives of the included subfamilies.

100%, MP 64%) and Eumenophorinae (BI 100%, ML 98%, MP 68%). Monophyly of the family Theraphosidae itself is strongly supported (BI 100%, ML 100%, MP 96%).

The three subfamilies Avicullariinae, Selenocosmiinae and Schismatothelinae were found to be potentially paraphyletic or polyphyletic. Taxa currently placed in the subfamily Avicullariinae form distinct clades in our phylogeny. The first clade contains the genera *Avicularia*, *Caribena* and *Ybyrapora* (BI 100%, ML 100%, MP 99%). Given the name bearing character of *Avicularia* for the subfamily, this clade is herein referred to as Avicullariinae. The genera *Psalmopoeus* and *Tapinauchenius* form a unique divergent clade (BI 100%, ML 100%, MP 100%) from the remaining Avicullariinae, separated by a clade containing members of the Schismatothelinae. This diverging clade of *Tapinauchenius* and *Psalmopoeus* is herein referred to as Psalmopoeinae. The members of the subfamily Schismatothelinae, which appear strongly affiliated to Psalmopoeinae and more widely to Avicullariinae, also do not form a single clade and instead successively split from the tree. The first Schismatothelinae representative, *Neoholothele incei,* is placed on a single branch splitting from Avicullariinae (BI 100%, ML 93%). The second Schismatothelinae taxon, *Euthycaelus colonica* splits off next as more closely allied to Psalmopoeinae (BI 100%, ML 96%).

Members of the traditionally recognized subfamily Selenocosmiinae also form two distinct clades. The first (BI 100%; ML 100%, MP 100%) consists of several genera, including the type genus *Selenocosmia*, and is here considered as Selenocosmiinae. The second clade (BI 100%, ML 81%, MP 84%) consists exclusively of the genus *Poecilotheria* (three taxa: *P. subfusca*, *P. vittata*, *P. formosa*) and appears to be more closely related to the subfamilies Thrigmopoeinae and Ornithoctoninae than to other Selenocosmiinae, therefore it is hereafter considered as Poecilotheriinae. Monophyly of Thrigmopoeinae remains unresolved due to only one available taxon for this study.

Interestingly, the only included taxon of the poorly defined subfamily 'Ischnocolinae' (sensu lato), *Nesiergus insulanus*, is placed inside the well-supported Selenocosmiinae, where, however, most of the internal groupings relative to *Nesiergus* are not well supported.

Although most subfamilial relationships in the tarantula phylogeny were only poorly supported, almost all of the well-supported deep clades were restricted to major geographical regions or continents. The African subfamilies Harpactirinae and Stromatopelminae were sister to each other and together formed a novel clade (here named Africa 1) (BI 100%, ML 100%, MP 88%), as did the American (mostly South American) subfamilies Schismatothelinae, Psalmopoeinae and Avicullariinae (America 1; BI 68%, ML 71%, MP 51%). These two main clades (Africa 1 and America 1) were resolved as sister clades (BI 99%, ML 55%), possibly reflecting the historical close affinities of these landmasses. On the contrary, the clades America 2 (BI 100%, ML 100%, MP 64%) consisting only of the Theraphosinae, and Africa 2 (BI 100%, ML 98%, MP 68%) consisting of the African-Madagascan Eumenophorinae, each occupied other positions in the phylogeny and split from deeper nodes. Two well supported clades from Asia were identified, corresponding to the Selenocosmiinae including the Seychellean 'Ischnocolinae' *Nesiergus insulanus* (Asia 1) separate from the second strongly supported Asian clade that groups the Poecilotheriinae, Thrigmopoeinae and Ornithoctoninae (Asia 2; BI 100%, ML 99%, MP 51%). Our results cannot exclude that these two Asian clades might be each other's sister groups as the node indicating the closer placement of Asia 1 with the clades Africa + America 1 did not receive high support from any of the three phylogenetic approaches.

4. Discussion

4.1. A preliminary molecular perspective on theraphosid phylogeny

Our molecular phylogenetic analysis reliably identified several major clades within Theraphosidae that largely agreed with the current subfamilial classification, but also provided numerous new insights. While we consider the well-supported major clades in our tree as reliable taxonomic groupings, our analysis also is affected by several restrictions.

For specimens obtained via the pet trade, although provenance was vague in several cases, all could be confidently assigned to countries of origin or at least to secure geographic regions (Supplementary Table S1). Wherever possible, we re-evaluated the taxonomic identification of our samples with careful consideration of the morphology for vouchered specimens (rather than using alleged 'trade-name identity' at face value) to ensure their utility. That said, we could not verify species-level identification with absolute certainty in all cases, nor for example could we exclude the possibility of captive-bred hybrids. However, as in this study we focus on higher-level relationships, this qualification should not affect our main conclusions since we are confident that all specimens have been correctly identified at the generic level, a conclusion further supported by our DNA barcoding approach (Supplementary Table S2 and Fig. S8).

A second restriction is a somewhat biased distribution of missing data in our analysis, potentially affecting some of the analyses. Our study included gene segments of highly different evolutionary rates, such as the often highly conserved 18S and the fast evolving COI, both of which however are missing for some taxa. This might in some cases have led to biased branch lengths in our tree (e.g. under-estimated in clades of taxa with missing faster evolving gene segments). For this reason and the lack of adequate calibration points we refrained from attempting to date the divergences in our tree. While such timetree analyses of the major theraphosid lineages would be of high biogeographical interest, in our opinion more comprehensive datasets of multiple protein-coding nuclear genes are needed to date theraphosid divergence times with confidence.

Another restriction is given by the limited taxon sampling. Besides the missing subfamily of Selenogyrinae, three other subfamilies are lacking taxa of special interest. Firstly, we highlight that our phylogeny does not include any Eumenophorinae from mainland Asia nor the recently described *Mascaraneus* from Serpent Island off Mauritius, both of which are of high biogeographic interest. Secondly, no Australian representatives of Selenocosmiinae were included in this study, but would be important for understanding Australasian lineages. Thirdly, no European/Mediterranean representatives of Ischnocolinae were included, nor several other 'ischnocolines' from diverse geographical regions. The subfamily Ischnocolinae is a problematic group (after Raven, 1985, Supplementary Fig. S9) of several genera with uncertain placement, for which we only included one genus (*Nesiergus*) currently assigned to that subfamily. The herein presented phylogeny provides new insights on its taxonomic placement. However, we point out that our two sampled Schismatothelinae were only recently transferred from Ischnocolinae to this recently erected subfamily, and here also provide further new insights about their taxonomic placement. Outside of Theraphosidae, with the exception of *Brachionopus* sp., we omitted any taxa from this study that have been associated with the family Barychelidae (a plausible sister-family, but of uncertain composition) in favour of using the relatively closely related nemesiids and a more distant diplurid as unambiguous and informative outgroups (Bond et al., 2012; Wheeler et al., 2016). Recent studies have indicated that Barychelidae might be paraphyletic with some former barychelids instead more closely related to certain ischnocoline theraphosids (e.g., Guadanucci, 2014). To address these final issues, a comprehensive sampling of ischnocolines and barychelids should be at the core of future studies.

4.2. Subfamilial classification of Theraphosidae

Our molecular phylogenetic study largely supported, in agreement with previous morphological analysis, the validity of the prevalent subfamily-level classification within the Theraphosidae. However, it

also suggests a need for revisiting the validity, composition, and interrelationships of several subfamilies.

One genus of previously uncertain placement is *Poecilotheria*, often referred to as "ornamental spiders" and placed into a monotypical subfamily Poecilotheriinae soon after its initial description (Simon, 1892). More recent studies instead had suggested a placement inside the Selenocosmiinae based on the morphology of stridulatory organs (e.g. Raven, 1985). Although Schmidt (1995) argued for validity of Simon's subfamily Poecilotheriinae, most authors refused to follow this and therefore the majority of studies treated the genus *Poecilotheria* as a member of the Selenocosmiinae. Marshall et al. (1999) even considered the genus (within Selenocosmiinae) to be sister to *Psalmopoeus*, a South American genus currently placed into the Avicullariinae by other authors (e.g. Fukushima and Bertani, 2017). In the molecular phylogeny (Fig. 1) most taxa of the Selenocosmiinae form the clade Asia 1, including the name bearing genus *Selenocosmia*, but the three included species of *Poecilotheria* (*P. vittata*, *P. formosa* and *P. subfusca*) were nested within the clade Asia 2 together with Thrigmopoeinae and Ornithoctoninae, with substantial support from BI and ML (BI 100%, ML 99%, MP 51%). From a biogeographic perspective, both *Poecilotheria* and its direct sister group Thrigmopoeinae (represented by *Thrigmopoeus truculentus* in our analysis), are South Asian taxa, all specifically from the Indian subcontinent. Given the reliable and biogeographically sound phylogenetic placement of *Poecilotheria* apart from the Selenocosmiinae, we support considering this genus as its own subfamily Poecilotheriinae in agreement with Simon (1892) and Schmidt (1995). However, future studies are necessary to understand whether inclusion of *Poecilotheria* in Thrigmopoeinae might be warranted instead.

A further enigmatic group within theraphosids is the subfamily Ischnocolinae, which has previously been found to be paraphyletic (Raven, 1985; Guadanucci, 2014). The only "Ischnocolinae" (sensu lato) taxon included in our study, *Nesiergus insulanus* from the Seychelles archipelago (Guadanucci and Gallon, 2008; Canning et al., 2013), was placed inside the Selenocosmiinae by the molecular data, although support for this grouping was low. Whether this placement reflects true relationships and also applies to other Ischnocolinae can only be decided with wider sampling. Our DNA sequences of the *Nesiergus* specimen were extracted from old preserved molts and therefore for several genes only relatively small fragments could be amplified for this species, specifically from CO1, 28S and H3. While the nested position of *Nesiergus* within Selenocosmiinae clearly requires further confirmation, both the 28S and H3 data do support the main clade of Selenocosmiinae containing this genus.

The subfamily Schismatothelinae from the Americas has recently been proposed for several genera previously classified in the Ischnocolinae (Guadanucci, 2014). The Schismatothelinae in our study are represented by two species (*Neoholothele incei* and *Euthycaelus colonica*) that do not form a monophyletic group but are paraphyletically clustered between two South American subclades. A more comprehensive sampling of these subfamilies is needed for final conclusions, but our data suggests that the systematics of Schismatothelinae is in need of further revision. The reconstructed phylogenetic placement of schismatothelines contributes to understanding the status of other South American theraphosid groups. The genera *Avicularia*, *Caribena*, *Psalmopoeus*, *Tapinauchenius* and *Ybyrapora* are currently placed in the subfamily Avicullariinae (Fukushima and Bertani, 2017), based on morphological characters. However, our molecular data implies a possible paraphyly of the Avicullariinae in which the clade of *Avicularia*, *Caribena* and *Ybyrapora* is separated from the genera *Psalmopoeus* and *Tapinauchenius* which form a clade sister to the two schismatotheline lineages discussed above. Given the key role of schismatotheline taxa for resolving avicullariine systematics, we consider the detailed analysis of this subfamily as another priority for future studies on molecular theraphosid phylogeny. That being said, we here restrict the use of Avicullariinae to the clade containing the type genus, *Avicularia*, plus *Caribena* and *Ybyrapora*. Following Samm and Schmidt (2010), we preliminarily applied the term of Psalmopoeinae for the second clade of avicullariine taxa (herein containing *Tapinauchenius* and *Psalmopoeus*), but emphasize that this phylogenetic hypothesis and classification scheme requires further testing.

The African genera *Stromatopelma* and *Heteroscodra* have also been suggested to be members of the otherwise American Avicullariinae (after West et al., 2008; see Fukushima and Bertani, 2017) whereas previous classifications considered them as Eumenophorinae (Raven, 1985; Smith, 1990) or in their own subfamily Stromatopelminae (Schmidt, 1993). In our molecular phylogeny (Fig. 1) the clade of the two genera is placed with high support as sister to the African Harpactirinae, both groups forming the clade Africa 1. This result strongly refutes a placement of *Stromatopelma* and *Heteroscodra* with Avicullariinae and we therefore suggest accepting Stromatopelminae as its own, valid subfamily. Also within the clade Africa 1, our data support the inclusion of *Brachionopus* and *Harpactirella* in Harpactirinae in agreement with Raven (1985) and Smith (1990) and contrary to their placement in Barychelidae by Schmidt (2002).

The subfamily Eumenophorinae represents one of the biogeographically most interesting clades of theraphosids, comprising species distributed all over Africa, the Arabian Peninsula, India, as well as on various Indian Ocean islands (Madagascar, Socotra and Mascarenes). Unfortunately, Asian Eumenophorinae were not available for this study, most notably several Indian taxa recently transferred to the subfamily, including some former 'Ischnocolinae' (Guadanucci, 2011a, 2011b; Mirza et al., 2014). But regardless our phylogeny provides interesting new support for the placement of the genus *Monocentropus* (occurring in Madagascar, Socotra and mainland of Yemen) as a unique clade apart from other mainland African Eumenophorinae genera. A similar Madagascar-Socotra relationship as in *Monocentropus* is also found in other animals, for instance in pseudoxyrhophiine snakes (Nagy et al., 2003). In contrast, the Seychellean genus *Nesiergus* is nested in Asian Selenocosmiinae (Fig. 1), being a further example for the predominant biogeographic relationships of the Seychelles with Asia (e.g. Biju and Bossuyt, 2003) rather than with Madagascar. The recent discovery of an endemic eumenophorine genus from Serpent Island, an islet north of the oceanic island of Mauritius (*Mascaraneus*; Gallon, 2005b) suggests that tarantulas have likely been able of overseas dispersal.

From Madagascar, only three tarantula species have been described: *Monocentropus lambertoni*, *Encyocrates raffrayi* and *Phoneyusa bouvieri* (Smith, 1990; Griswold, 2003) all of which are placed in the Eumenophorinae (Schmidt, 2003). Our tree includes three Malagasy samples, two of which showed high genetic divergences from the *Monocentropus lambertoni* sample. The two sequenced specimens from Nosy Hara (FGZC 2053) and Marojejy (FGZC 1333) have no spike or paddle setae on the coxa of legs 1–2 (see identification keys in Smith, 1990 and Schmidt, 2003), suggesting that they belong to *Monocentropus* and are not referable to the remaining two Malagasy theraphosid species (*Encyocrates raffrayi* and *Phoneyusa bouvieri*). Although tarantulas are rarely observed in much of Madagascar, it seems likely that this biodiversity hotspot harbors a larger diversity of tarantulas (especially of *Monocentropus*) than currently recognized.

4.3. Comparison with an eminent theraphosid phylogeny based on morphology

At present, the most referred phylogenetic hypothesis of theraphosid relationships is the morphological analysis of Raven (1985) (see Supplementary Fig. S9). Compared to this study, our data suggest novel relationships for several supra-generic groupings of taxa, especially for the African lineages. Raven (1985) placed the African genera *Stromatopelma* and *Heteroscodra* (today together as the subfamily Stromatopelminae) with the African subfamily Eumenophorinae. In contrast, our analysis supports the Eumenophorinae as evolutionarily distinct, instead splitting from the deepest node of the theraphosid phylogeny, and possibly forming the sister group of all remaining

theraphosids. In our tree, the Stromatopelminae in contrast are strongly supported as sister to the other major African lineage of theraphosids, the subfamily Harpactirinae. These instead were placed next to the Theraphosinae, Avicullariinae and some Ischnocolinae by Raven (1985). Finally, the morphological tree (Raven, 1985) places the South American subfamilies Avicullariinae and Theraphosinae as close relatives whereas the molecular data provide evidence for a closer relationship of our sampled Avicullariinae with the African Harpactirinae and Stromatopelminae.

4.4. Conclusion and outlook

In this study we provide a preliminary hypothesis of the deep evolutionary history of theraphosid spiders based on molecular data. We found a clear biogeographic pattern in the Theraphosidae phylogeny, with major clades each corresponding to taxa distributed on particular continents and landmasses. The African, Asian and South American regions turned out to be each colonized by two major regional endemic tarantula clades. Our phylogeny suggests that the two South American clades do not form a monophyletic group, and the same is true for the two African clades, suggesting that two separate theraphosid groups diversified on each of these continents. The situation is less clear for the two Asian clades since our tree does not provide strong support for any wider nodes concerning their relative placement. Theraphosid higher classification, so far based largely only on morphology, clearly requires revision in several aspects, although we supported many traditionally recognized subfamilies as monophyletic. In particular, the paraphyly of Schismatotheliinae and the possible nested placement of some “Ischnocolinae” within Selenocosminae require further investigation with extended sets of taxa and sequences of additional genes. Current timetrees place the split of theraphosids from their sister group into the Cretaceous (Ayoub et al., 2007; Ayoub and Hayashi, 2009; Bond et al., 2014; Garrison et al., 2016), but divergence dates within the family have not yet been reconstructed. Sampling sequences of additional nuclear protein-coding genes from representatives of the main clades identified herein, plus representatives of other arachnid orders to use fossil calibrations for the arthropod tree of life (Wolf et al., 2016), will allow formulating hypotheses on divergence dates within the Theraphosidae, and thereby contribute to further developing a biogeographic and evolutionary scenario for these fascinating animals.

Acknowledgements

We are grateful to Meike Kondermann and Marc Hoffmann for executing parts of the labwork. Jörg Bernhardt, Matthias Köhler, Richard Gallon, Volker von Wirth, Dirk Weinmann and Jean-Michel Verdez provided sample material to the study. We thank Michael Morra for assistance in dataset generation. Monty Oberländer, Aileen Schlag, Marcus Herrmann and Marlen Ziegler kindly provided and handled animals that were photographed and used to illustrate the phylogenetic tree. Susanne Hauswaldt provided COI primers originally designed for *Osmoderma* beetles. We are grateful to the Malagasy authorities for issuing research and export permits.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.10.015.

References

Ayoub, N.A., Garb, J.E., Hedin, M., Hayashi, C.Y., 2007. Utility of the nuclear protein-coding gene, elongation factor-1 gamma (*EF-1y*), for spider systematics, emphasizing family level relationships for tarantulas and their kin (Araneae: Mygalomorphae). Mol. Phylogenet. Evol. 42, 394–409.

Ayoub, N.A., Hayashi, C., 2009. Spiders (Araneae). In: Hedges, S.B., Kumar, S. (Eds.), The Timetree of Life. Oxford University Press, Oxford, pp. 255–259.

Bertani, R., 2000. Male palpal bulbs and homologous features in Theraphosinae (Araneae, Theraphosidae). J. Arachnol. 28, 29–42.

Bertani, R., Guadanucci, J.P.L., 2013. Morphology, evolution and usage of urticating setae by tarantulas (Araneae: Theraphosidae). Zoologica (Curitiba) 30, 403–418.

Biju, S.D., Bossuyt, F., 2003. New frog family from India reveals an ancient biogeographical link with the Seychelles. Nature 425, 711–714.

Bond, J.E., Opell, B.D., 2002. Phylogeny and taxonomy of the genera of south-western North American Euctenizinae trapdoor spiders and their relatives (Araneae: Mygalomorphae, Cyrtaucheniidae). Zool. J. Linn. Soc. 138, 487–534.

Bond, J.E., Hendrixson, B.E., Hamilton, C.A., Hedin, M., 2012. A reconsideration of the classification of the spider infraorder Mygalomorphae (Arachnida: Araneae) based on three nuclear genes and morphology. PLoS ONE 7, e38753.

Bond, J.E., Garrison, N.L., Hamilton, C.A., Godwin, R.L., Hedin, M., Agnarsson, I., 2014. Phylogenomics resolves a spider backbone phylogeny and rejects a prevailing paradigm for orb web evolution. Curr. Biol. 24, 1765–1771.

Bruford, M.W., Hanotte, O., Brookfield, J.F.Y., Burke, T., 1992. Single-locus and multilocus DNA fingerprinting. In: Hoelzel, A.R. (Ed.), Molecular Genetic Analysis of Populations: A Practical Approach. IRL Press, Oxford, pp. 225–270.

Canning, G., Reilly, B.K., Dippenaar-Schoeman, A.S., 2013. First description of the male of *Nesiergus insulanus* (Araneae: Theraphosidae: Ischnocolinae) from the Seychelles Archipelago. African Invertebr. 54, 241–244.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Coddington, J.A., 2005. Phylogeny and Classifcation of Spiders. In: Ubick, D., Paquin, P., Cushing, P.E., Roth, V. (Eds.), Spiders of North America: An Identification Manual. American Arachnological Society, pp. 18–24.

Colgan, D.J., McLauchlan, A., Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Cassis, G., Gray, M.R., 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. Aust. J. Zool. 46, 419–437.

Edgar, R.C., 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5, 1–19.

Edgar, R.C., 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797.

Fernandez, R., Hormiga, G., Giribet, G., 2014. Phylogenomic analysis of spiders reveals nonmonophyly of orb weavers. Curr. Biol. 24, 1772–1777.

Fukushima, C.S., Bertani, R., 2017. Taxonomic revision and cladistic analysis of *Avicularia* Lamarck, 1818 (Araneae, Theraphosidae, Avicullariinae) with description of three new avicullariine genera. ZooKeys 659, 1–185.

Gallon, R.C., 2003. A new African arboreal genus and species of theraphosid spider (Araneae, Theraphosidae, Stromatopelminae) which lacks spermathecae. Bull. Br. Arach. Soc. 12, 405–411.

Gallon, R.C., 2005a. *Encyocratella olivacea* Strand, 1907, a senior synonym of Xenodendrophila gabrieli Gallon, 2003 (Araneae: Theraphosidae: Stromatopelminae) with a description of the male. Zootaxa 1003, 45–56.

Gallon, R.C., 2005b. On a new genus and species of theraphosid spider from Serpent Island, Mauritius (Araneae, Theraphosidae, Eumenophorinae). Bull. Br. Arach. Soc. 13, 175–178.

Garrison, N.L., Rodriguez, J., Agnarsson, I., Coddington, J.A., Griswold, C.E., Hamilton, C.A., Hedin, M., Kocot, K.M., Ledford, J.M., Bond, J.E., 2016. Spider phylogenomics: untangling the Spider Tree of Life. PeerJ 4, e1719. http://dx.doi.org/10.7717/peerj.1719.

Giribet, G., Carranza, S., Riutort, M., Baguñà, J., Ribera, C., 1999. Internal phylogeny of the Chilopoda (Myriapoda, Arthropoda) using complete 18S rDNA and partial 28S rDNA sequences. Philos. Trans. R. Soc. Lon. 354, 215–222.

Goloboff, P.A., 1993. A reanalysis of mygalomorph spider families (Araneae). Am. Mus. Novit. 3056, 1–32.

Griswold, C.E., 2003. Araneae, Spiders. In: Goodman, S.M., Benstead, J.P. (Eds.), The University of Chicago Press, Chicago, pp. 579–593.

Guadanucci, J.P.L., Gallon, R.C., 2008. A revision of the spider genera *Chaetopelma* Ausserer 1871 and *Nesiergus* Simon 1903 (Araneae, Theraphosidae, Ischnocolinae). Zootaxa 1753, 34–48.

Guadanucci, J.P.L., 2011a. Cladistic analysis and biogeography of the genus *Oligoxystre* Vellard 1924 (Araneae: Mygalomorphae: Theraphosidae). J. Arachnol. 39, 320–326.

Guadanucci, J.P.L., 2011b. The genus *Plesiophrictus* Pocock and revalidation of *Heterophrictus* Pocock (Araneae: Theraphosidae). J. Arachnol. 39, 523–527.

Guadanucci, J.P.L., 2014. Theraphosidae phylogeny: relationships of the Ischnocolinaé genera (Araneae, Mygalomorphae). Zool. Scripta 43, 508–518.

Hamilton, C.A., Formanowicz, D.R., Bond, J.E., 2011. Species delimination and phylogeography of *Aphonopelma hentzi* (Araneae: Mygalomorphae: Theraphosidae): Cryptic diversity in North American tarantulas. PLoS ONE 6, e26207.

Hamilton, C.A., Hendrixson, B.E., Bond, J.E., 2016. Taxonomic revision of the tarantula genus *Aphonopelma* Pocock 1901 (Araneae, Mygalomorphae, Theraphosidae) within the United States. Zookeys 560, 1–340.

Hedin, M., Maddison, W.P., 2001. A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). Mol. Phylogenet. Evol. 18, 386–403.

Hedin, M., Bond, J.E., 2006. Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): conflict and agreement with the current system of classification. Mol. Phylogenet. Evol. 41, 454–471.

Kambas, D., 2017. Tarantupedia TM. < https://www.tarantupedia.com/ > (10.03. 2017).

Klaas, P., 1989. Vogelspinnen im Terrarium. Ulmer, Stuttgart 148 pp.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA 7: Molecular Evolution Genetics Analysis Version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol.

Biol. Evol. 29, 1695–1701.
Longhorn, S.J., Nicholas, M., Chuter, J., Vogler, A.P., 2007. The utility of molecular markers from non-lethal DNA samples of the CITES II protected "tarantula" *Brachypelma vagans* (Araneae, Theraphosidae). J. Arachnol. 35, 278–292.
Marshall, S.D., Thoms, E.M., Uetz, G.W., 1995. Setal enlargement: an undescribed method of stridulation by a Neotropical tarantula (Araneae: Theraphosidae). J. Zool. 235, 587–595.
Marshall, S.D., Raven, R.J., Hoeh, W.R., 1999. A test of alternative phylogenetic relationship of theraphosid subfamilies using cladistic analysis of morphological traits. Am. Arachnol. 60, 5.
Mendoza, J., Francke, O., 2017. Systematic revision of *Brachypelma* red-kneed tarantulas (Araneae: Theraphosidae), and the use of DNA Barcodes to assist in the identification and conservation of CITES-listed species. Invertebr. Syst. 31, 157–179.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1–8.
Mirza, Z.A., Sanap, R.V., Bhosale, H., 2014. Preliminary review of Indian Eumenophorinae (Araneae: Theraphosidae) with description of a new genus and new species from the Western Ghats. PLoS ONE 9, e87928.
Montes De Oca, L., D'Elia, G., Perez-Miles, F., 2015. An integrative approach for species delimitation in the spider genus *Grammostola* (Theraphosidae, Mygalomorphae). Zool. Scripta 45, 322–333.
Nagy, Z.T., Joger, U., Wink, M., Glaw, F., Vences, M., 2003. Multiple colonization of Madagascar and Socotra by colubrid snakes: evidence from nuclear and mitochondrial gene phylogenies. Proc. Soc. Lond. B 270, 2613–2621.
Ortiz, D., Francke, O., 2016. Two DNA barcodes and morphology for multi-method species delimitation in Bonnetina tarantulas (Araneae: Theraphosidae). Mol. Phylogenet. Evol. 101, 176–193.
Palumbi, S., Martin, A., Romano, S., McMillan, W.O., Stice, L., Grabowski, G., 1991. The Simple Fool's Guide to PCR. Version 2. University of Hawaii, Honululu (HI).
Pérez-Miles, F., Lucas, S.M., da Silva, P.I., Bertani, R., 1996. Systematic revision and cladistic analysis of Theraphosinae (Araneae: Theraphosidae). Mygalomorph 1, 33–68.
Platnick, N.I., Gertsch, W.J., 1976. The suborders of spiders: a cladistic analysis (Arachnida, Araneae). Am. Mus. Novit. 2607, 1–15.
Rambaut, A., Drummond, A.J., 2007. Tracer: Mcmc Trace Analysis Tool. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh.
Ramirez, M.J., 2014. The morphology and phylogeny of Dionychan spiders (Araneae: Araneomorphae). Bull. Am. Mus. Nat. Hist. 390, 1–374.
Raven, R.J., 1985. The spider infraorder Mygalomorphae (Araneae): cladistics and systematics. Bull. Am. Mus. Nat. Hist. 182, 1–180.
Raven, R.J., 1990. Comments on the proposed precedence of *Aphonopelma* Pocock 1901 (Arachnida, Araneae) over *Rechostica* Simon, 1892. Bull. Zool. Nomencl. 42, 126.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Samm, R., Schmidt, G., 2010. Psalmopoeinae subfamilia nov. - eine neue Unterfamilie der Theraphosidae (Araneae). Tarantulas World 142, 35–41.
Schmidt, G., 1993. Vogelspinnen, 4. Aufl. Landbuch Hannover.
Schmidt, G., 1995. Die Stellung der Gattung *Poecilotheria* im System. Arachnida 10, 1–2.
Schmidt, G., 2002. Gehören *Brachionopus* Pocock, 1897 und *Harpactirella* Purcell, 1902 zu den Theraphosiden? Arthropoda 10, 12–17.
Schmidt, G., 2003. Die Vogelspinnen: Eine weltweite Übersicht. Neue Brehm-Bücherei, Hohenwarsleben.
Simon, E., 1892. Histoire naturelle des araignées. Paris. vol. 1, pp. 1–256.
Smith, A.M., 1990. Baboon Spiders. Tarantulas of Africa and the Middle East. Fitzgerald Publishing, London.
Stamatakis, A., 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Swofford, D.L., 2002. PAUP* Phylogenetic Analysis using Parsimony/and other Methods. Version 4b10. Sinauer Associates, Sunderland, Massachusetts.
Teyssié, F., 2015. Tarantulas of the world. NAP editions, Verrières-le-Buisson.
Turner, S.P., Longhorn, S.J., Hamilton, C.A., Gabriel, R., Pèrez-Miles, F., Vogler, A.P., 2017. Re-evaluating conservation priorities of New World tarantulas (Araneae: Theraphosidae) in a molecular framework indicates non-monophyly of the genera, Aphonopelma and Brachypelma. Syst. Biodivers. doi:http://dx.doi.org/10.1080/14772000.2017.1346719.
West, R.C., Marshall, S.D., Fukushima, C.S., Bertani, R., 2008. Review and cladistic analysis of the Neotropical tarantula genus *Ephebopus* Simon 1892 (Aranae: Theraphosidae) with notes on the Aviculariinae. Zootaxa 1849, 35–58.
West, R.C., Nunn, S.C., 2010. A taxonomic revision of the tarantula spider genus *Coremiocnemis* Simon 1892 (Araneae, Theraphosidae), with further notes on the Selenocosmiinae. Zootaxa 2443, 1–64.
West, R.C., Nunn, S.C., Hogg, S., 2012. A new tarantula genus, *Psednocnemis*, from West Malaysia (Araneae: Theraphosidae) with cladistic analyses and biogeography of Selenocosmiinae. Zootaxa 3299, 1–43.
Wheeler, W.C., Coddington, J.A., Crowley, L.M., Dimitrov, D., Goloboff, P.A., Griswold, C.E., Hormiga, G., Prendini, L., Ramírez, M.J., Sierwald, P., Almeida-Silva, L., Alvarez-Padilla, F., Arnedo, M.A., Benavides Silva, L.R., Benjamin, S.P., Bond, J.E., Grismado, C.J., Hasan, E., Hedin, M., Izquierdo, M.A., Labarque, F.M., Ledford, J., Lopardo, L., Maddison, W.P., Miller, J.A., Piacentini, L.N., Platnick, N.I., Polotow, D., Silva-Dávila, D., Scharff, N., Szűts, T., Ubick, D., Vink, C.J., Wood, H.M., Zhang, J., 2016. The spider tree of life: phylogeny of Araneae based on target-gene analyses from an extensive taxon sampling. Cladistics. http://dx.doi.org/10.1111/cla.12182.
Whiting, M.F., Carpenter, J.C., Wheeler, Q.D., Wheeler, W.C., 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology treated as equally weighted. Syst. Biol. 46, 1–68.
Wolf, J.M., Daley, A.C., Legg, D.A., Edgecomb, G.D., 2016. Fossil calibrations for the arthropod Tree of Life. Earth-Sci. Rev. 160, 43–110.
World Spider Catalog, 2017. The World Spider Catalog, Natural History Museum Bern, online at http://wsc.nmbe.ch, version 18.5, accessed on october 2017.

Chapter II

Tarantula phylogenomics: A robust phylogeny of deep theraphosid clades inferred from transcriptome data sheds light on the prickly issue of urticating setae evolution

Saoirse Foley, Tim Lüddecke*, Dong-Qiang Cheng, Henrik Krehenwinkel, Sven Künzel, Stuart J. Longhorn, Ingo Wendt, Volker von Wirth, Rene´ Tänzler, Miguel Vences & William H. Piel
Molecular Phylogenetics and Evolution (2019) 140: 106573.
*shared first-authorship

Molecular Phylogenetics and Evolution 140 (2019) 106573

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Tarantula phylogenomics: A robust phylogeny of deep theraphosid clades inferred from transcriptome data sheds light on the prickly issue of urticating setae evolution

Saoirse Foley[a,c,1], Tim Lüddecke[b,*,1], Dong-Qiang Cheng[c], Henrik Krehenwinkel[d], Sven Künzel[e], Stuart J. Longhorn[f], Ingo Wendt[g], Volker von Wirth[h], Rene Tänzler[i], Miguel Vences[j], William H. Piel[a,c]

[a] National University of Singapore, Department of Biological Sciences, 16 Science Drive 4, Singapore 117558, Singapore
[b] Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Animal Venomics Research Group, Winchesterstr. 2, 35394 Gießen, Germany
[c] Yale-NUS College, 10 College Avenue West, #01-101, Singapore 138609, Singapore
[d] Department of Biogeography, Trier University, Trier, Germany
[e] Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany
[f] Arachnology Research Association, United Kingdom
[g] Staatliches Museum für Naturkunde Stuttgart, Rosenstein 1, 70191 Stuttgart, Germany
[h] Deutsche Arachnologische Gesellschaft e.V., Zeppelinstr. 28, 71672 Marbach a. N., Germany
[i] Zoologische Staatssammlung München, Münchhausenstr. 21, 81247 München, Germany
[j] Zoological Institute, Technische Universität Braunschweig, Mendelssohnstr. 4, 38106 Braunschweig, Germany

ARTICLE INFO

Keywords:
Theraphosidae
Phylogenomics
Transcriptomics
Core Ortholog Approach
Urticating Setae

ABSTRACT

Mygalomorph spiders of the family Theraphosidae, known to the broader public as tarantulas, are among the most recognizable arachnids on earth due to their large size and widespread distribution. Their use of urticating setae is a notable adaptation that has evolved exclusively in certain New World theraphosids. Thus far, the evolutionary history of Theraphosidae remains poorly understood; theraphosid systematics still largely relies on morphological datasets, which suffer from high degrees of homoplasy, and traditional Sanger sequencing of preselected genes failed to provide strong support for supra-generic clades. In this study, we provide the first robust phylogenetic hypothesis of theraphosid evolution inferred from transcriptome data. A core ortholog approach was used to generate a phylogeny from 2460 orthologous genes across 25 theraphosid genera, representing all of the major theraphosid subfamilies, except Selenogyrinae. Our phylogeny recovers an unprecedented monophyletic group that comprises the vast majority of New World theraphosid subfamilies including Avicullariinae, Schismatothelinae and Theraphosinae. Concurrently, we provide additional evidence for the integrity of questionable subfamilies, such as Poecilotheriinae and Psalmopoeinae, and support the non-monophyly of Ischnocolinae. The deeper relationships between almost all subfamilies are confidently inferred. We also used our phylogeny in tandem with published morphological data to perform ancestral state analyses on urticating setae, and contextualize our reconstructions with emphasis on the complex evolutionary history of the trait.

1. Introduction

Theraphosidae is the largest family of mygalomorph spiders whose members are more commonly known as tarantulas or bird-eating spiders. Researchers have long been intrigued by these spiders due to both their large size and the plethora of adaptive traits they display. Tarantulas were painted by early naturalists as bird-eating monstrosities, and later as enemies of humans in horror movies, which together earned them a notorious reputation among the general public. Interest in these spiders has rapidly increased in recent decades, and is still enthusiastically sustained among both hobbyists (e.g. Klaas, 1989; Von Wirth, 2005; Cleton et al., 2015; Teyssié, 2015) and researchers to this day (Molur et al., 2008; Turner et al., 2018; Lüddecke et al., 2018; Hüsser, 2018).

* Corresponding authors.
E-mail addresses: schioedtei@gmail.com (S. Foley), tim.lueddecke@outlook.com (T. Lüddecke).
[1] Shared first authorship, both authors contributed equally.

https://doi.org/10.1016/j.ympev.2019.106573
Received 8 January 2019; Received in revised form 26 April 2019; Accepted 24 July 2019
Available online 30 July 2019

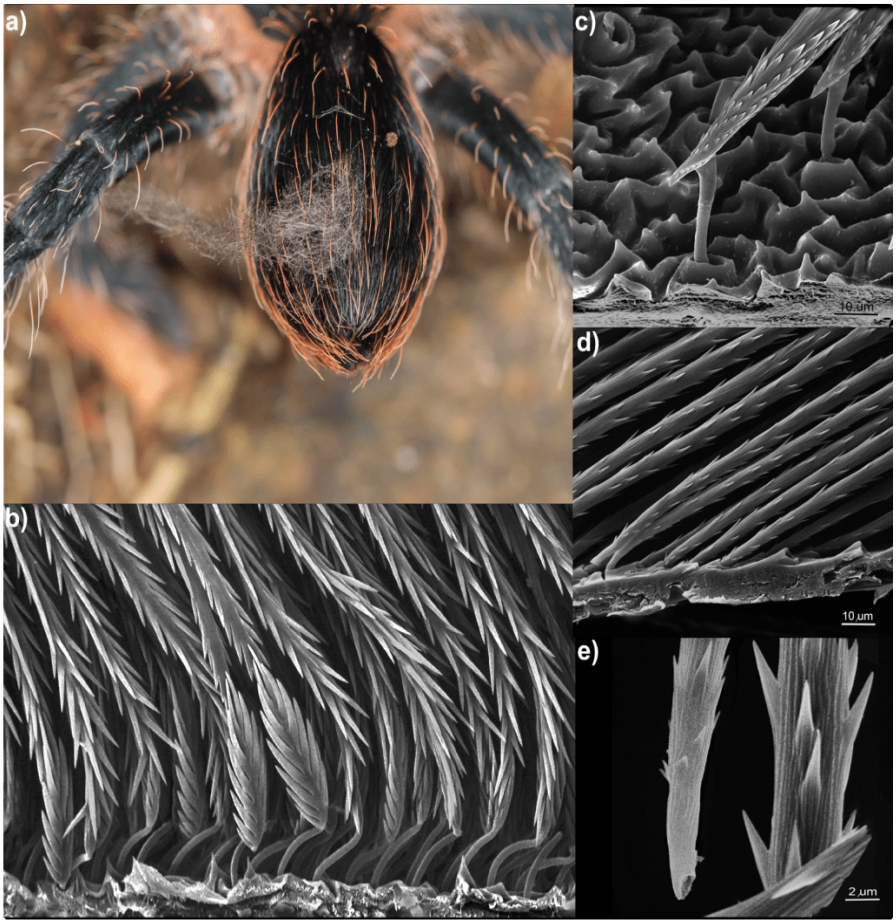


Fig. 1. Urticating setae as a means of defense in tarantulas. In most cases urticating setae are sloughed off by the spider from its dorsal or lateral opisthosoma and projected into the air, as given in (a) (photography: M. Hüsser). The barbed and harpoon-like microstructure of urticating setae in selected tarantulas is highlighted in (b)–(e) (photographs courtesy of Rainer Foelix): (b) A dense assemblage of type I & III setae in *Brachypelma* sp. from the opisthosoma, (c) type II setae from *Avicularia* sp. from the opisthosoma. (d) Type V from *Ephebopus* from the palpal femur and e) more closely detailed.

At the time of writing, Theraphosidae contains 1004 accepted species across 147 genera (World Spider Catalog, 2019) that are placed into 12 subfamilies by most authors (Kambas, 2019). This species richness, combined with their widespread distribution and hence their diversity of habitats and ecological niches, is linked to a high variability of morphological and ecological adaptations. These adaptations have been the focus of a variety of studies in recent years. For example, the mechanisms and structures that are responsible for the vast array of theraphosid colorations (Hsiung et al., 2015), venom compositions (e.g. Rodríguez-Rios et al., 2017; Santana et al., 2017), adhesive capabilities (Pérez-Miles et al., 2017) as well as urticating setae (Bertani and Guadanucci, 2013) have all received a fair amount of attention. The latter represents a special feature in Theraphosidae. These setae are exclusive to certain Neotropical tarantulas, but generally rare in the animal kingdom, although some lepidopteran caterpillars have vaguely comparable analogues (e.g. Battisti et al., 2011). Urticating setae of theraphosids are barbed, microscopic setae that are usually localized on the opisthosoma (Cooke et al., 1972; Pérez-Miles,2002; Bertani and Guadanucci, 2013; Teyssié, 2015). In a long-range defense mechanism loosely known as “hair flicking” or “bombardment” (Fig. 1a), many Neotropical species will use their legs to slough off these setae from their opisthosoma and actively disperse them into the air when a threat is perceived (Cooke et al., 1972; Bertani et al., 2003; Bertani and Guadanucci, 2013). Other species opt for an active “direct contact” approach against potential threats, in which they seek to press them onto the target (after Cooke et al., 1972; Bertani and Guadanucci, 2013; Perafán et al., 2016). Consequently, urticating setae have been known to induce painful symptoms in humans (Zilkens et al., 2012; McAnena et al., 2013). Contact with urticating setae may cause mammalian skin or eye irritations due to their harpoon-like structure, and inhalation has been known to result in respiratory problems (Schmidt, 2003; Bertani and Guadanucci, 2013; Teyssié, 2015). When directed against mammals, these defenses may provide the tarantula an opportunity to escape from predation (Cooke et al., 1972). Additionally, some species also utilize urticating setae in a passively defensive manner, where they are attached to silk that covers egg-sacs or forms molting webs. This has been shown to be effective against invertebrates, in particular countering the larvae of parasitic flies (i.e. phorids) or ants (Marshall and Uetz, 1990a; Bertani and Guadanucci, 2013).

Seven different types of urticating setae have been described and they are distinguished by characters such as their microstructure, localization, or release mechanism (Bertani and Guadanucci, 2013; Perafán et al., 2016). The types I, III, IV, VI and VII occur in members of the Theraphosinae subfamily (Fig. 1b), whereas type II is exclusive to Aviculariinae (Fig. 1c). Certain types are even genus-specific, such as type V for *Ephebopus* (Fig. 1d and e), type VI for *Hemirrhagus* (Pérez-Miles, 1998) and type VII for *Kankuamo* (Bertani and Guadanucci, 2013; Foelix et al., 2009; Marshall and Uetz, 1990b; Mendoza and Francke, 2018; Perafán et al., 2016). This conservation and specificity to certain groups of tarantulas has led to their widespread use as diagnostic taxonomic characters (e.g. Schmidt, 2003; Turner et al., 2018).

Given the species richness, charismatic nature and the breadth of scientific work on their adaptations, it is rather surprising and unfortunate that the taxonomic status and evolutionary history of theraphosids in general remains poorly understood. Most previous works on theraphosid evolution and systematics have been based exclusively upon the analysis of morphological characters (e.g. West et al., 2008; West and Nunn, 2010; Guadanucci, 2014; Fukushima and Bertani, 2017), which have often yielded inconsistent results. Subsequent studies recognized the limitations of purely morphological analyses due to high degrees of homoplasy (e.g. Ortiz et al., 2018). The use of molecular characters to study theraphosid evolution is relatively recent (e.g.

Longhorn et al., 2007; Petersen et al., 2007; Hamilton et al., 2016a; Hamilton et al., 2016b; Ortiz and Francke, 2016; Mendoza and Francke, 2017; Turner et al., 2018; Ortiz et al., 2018; Hüsser, 2018), and a first comprehensive analysis of major theraphosid lineages was able to support the monophyly of Theraphosidae itself as well as of many subfamilies within (Lüddecke et al., 2018). Further, it provided molecular evidence for the validity of certain subfamilies that were previously uncertain, and therefore highlighted the potential for molecular data to resolve the theraphosid phylogeny. Despite this first study targeting several commonly used nuclear and mitochondrial genes and revealing several well-supported suprageneric clades within theraphosids, the deeper phylogenetic relationships among many of these clades were poorly supported and remained uncertain.

Well-supported phylogenies are key to making meaningful inferences about evolution, phylogeography, adaptation, and beyond. In this study, we aim to address this deficiency by providing a phylogenomic perspective on theraphosid evolution with a robust backbone inferred from an extensive molecular dataset. We sequenced 29 transcriptomes from a comprehensive range of disparate species and combined these with 4 other published transcriptomes to generate a phylogenetic hypothesis that includes representatives of all major subgroups within the family (except Selenogyrinae). We analysed several subsets of orthologous genes (OGs), selected under varying measures of stringency with the intent to (i) resolve deeper clades within the phylogeny of Theraphosidae, (ii) test whether strongly supported nodes will be recovered in congruence with previous morphological and molecular studies as a means to assess their validity, and to (iii) reconstruct the ancestral states of urticating setae, an excellent model for clade-specific traits with important biological functions. Further, this phylogeny will serve as the groundwork for future evolutionary studies to examine other morphological adaptations in Theraphosidae, as well as other aspects such as evolution of venom components and biogeographical analyses.

2. Materials and methods

2.1. Sampling, RNA extraction and sequencing

Detailed information regarding the particulars of each sample is available in supplementary Table S1, along with voucher deposition numbers. Sample material for this study was obtained either from the pet trade, private breeders, or collected in the field. Each specimen was identified using published identification keys, at least to genus level (Schmidt, 2003; Teyssié, 2015). We collected material from 26 members of Theraphosidae for this study, covering most of the subfamilies within the group. Publicly available transcriptome data from *Damarchus* sp. (Nemesiidae) was included as the closest currently available secure outgroup taxon to Theraphosidae (Fernández et al., 2018) and the wider Theraphosoidina clade (Opatova et al., 2019 in prep). However, we also generated transcriptome data from the diplurid spider *Linothele* sp. as an additional and more distant outgroup taxon due to ongoing debates surrounding mygalomorph familial relationships.

Whenever possible, whole bodies of each specimen were sampled for this study, but in some cases only autotomized legs were available. Samples were either stored in RNAlater (coded as TPx in Table S1), or were pre-frozen in liquid nitrogen prior to freezing at −80 °C (coded as SFx or RSWx in Table S1). RNA extractions were performed either via the TRIzol total RNA extraction protocol or traditional phenol chloroform extractions (Simms et al., 1993). Purified RNA was processed in the Max Planck Institute in Plön, Germany, or sent to commercial sequencing companies who prepared cDNA libraries, and the samples sequenced using Illumina HiSeq and NextSeq technologies. These transcriptomes were complemented with publicly available data for four additional species acquired from the NCBI SRA database (http://www.ncbi.nlm.nih.gov/sra), one of which being the aforementioned *Damarchus* sp. sample from the family Nemesiidae. We endeavoured to confirm that any variance from using different sequencing technologies and transcripts from different tissues was minimized by including pairs of conspecific (2x *Neoholothele incei* and 2x *Poecilotheria vittata*) and congeneric (2x *Caribena*, 2x *Psalmopoeus*, 2x *Pterinochilus* and 2x *Cyriopagopus*) taxa under the expectation that they would group together. Both the sequencing method and tissue sample for each member per pair differed from the other. All newly obtained transcriptome raw data were submitted to the Sequence Read Archive (SRA) under Bioproject number PRJNA534037. Assembly data for newly sequenced transcriptomes is available in supplementary Table S2.

2.2. Core ortholog analysis pipeline for generating the phylogeny

Transcriptomes were assembled using Trinity (Grabherr et al., 2011). Prediction of protein coding regions was performed in TransDecoder (Haas et al., 2013). A core ortholog approach based on Garrison et al. (2016) and modified by Cheng and Piel (2018) was used for putative ortholog selection, and aims to select only the genes that are orthologous across our taxon set. Following from Cheng and Piel (2018), a core ortholog set was derived from transcripts of *Damon variegatus, Acanthoscurria geniculata, Dolomedes triton, Ero leonina, Hypochilus pococki, Leucauge venusta, Liphistius malayanus, Megahexura fulva, Neoscona arabesca, Stegodyphus mimosarum*, and *Uloborus* sp., all of which are publicly available from SRA, and a set of 4446 spider-specific profile hidden Markov models (pHMMs) was then built following from Cheng and Piel (2018). HaMStR v.13.1 (Ebersberger et al., 2009) was used to infer orthology between the pHMMs and sequences from each taxon analysed here (see Supp. Tab. S1). Groups of orthologous genes (OGs) corresponding to the pHMMs were subsequently pooled and subjected to several rounds of filtering: Sequences shorter than 75 amino acids and OGs sampled for fewer than 17 specimens were discarded. MAFFT (Katoh, 2005) was used to align each OG, with the settings “auto”, “localpair” and “maxiterate 1000” selected. Alignments were trimmed by ALISCORE (Misof and Misof, 2009; Kück et al., 2010). Ambiguous regions were excised in ALICUT (Kück, 2009), and Infoalign (Rice et al., 2000) was used to build consensus sequences for each alignment. In an effort to remove paralogs, sequences that were far from the consensus (i.e. with an infoalign *change* value between the sequence and consensus exceeding 75) were removed. Sequences with > 9 gaps flanking a region with 20 or fewer mistranslated amino acids were excluded, as were alignment columns with < 4 non-gap characters. Any sequences that were shorter than 75 amino acids after these filtering steps were also removed. Finally, sequences that did not overlap with all other sequences in the alignment by at least 20 amino acids were removed, and OGs sampled for fewer than 17 taxa were once again discarded. This yielded our “full set” of OGs, which constitute our first dataset (matrix 1).

Following Cheng and Piel (2018), data for four additional OG matrices with varying degrees of conservation was derived via matrix reduction and optimization. These resulted from custom scripts (Cheng and Piel, 2018). The matrix 1 dataset was sorted first by gene occupancy, and then by gene length. OGs composing matrix 2 (“first reduce”) were obtained by only retaining the larger half of the sorted matrix 1 dataset. OGs composing matrix 3 (“second reduce”) were obtained by only retaining the larger half of matrix 2 dataset. OGs composing matrix 4 were selected using BaCoCa (Kück and Struck, 2014), which optimised the full matrix 1 dataset to retain only 50% composed of the most phylogenetically informative sites. OGs for matrix 5 were selected using MARE (Meyer et al., 2011), which assessed the matrix 1 dataset partitions, provided a measure of tree-likeness for each gene, and finally optimized the matrix for information content with an alpha value of 5. FASconCAT (Kück and Meusemann, 2010) was used to concatenate each of the five OG datasets to yield the five data matrices. Metrics for all matrices, including numbers of OGs and amino acids per dataset, are available in Table 1.

ExaML v.3.0.2 (Kozlov et al., 2015) was used for maximum likelihood searches, and optimal trees were calculated for each matrix, and

Table 1
Metrics of the five OG matrices, including the number of OGs per matrix, corresponding amino acid numbers (AAs), the percentage of missing data, and the log likelihood of the best tree for each matrix.

Analysis	#OGs	#AAs	% Missing	Log Likelihood
Matrix 1: All genes	2460	1,096,124	16.85%	−9703757.281168
Matrix 2: First reduce	1230	628,946	7.92%	−5534973.862171
Matrix 3: Second reduce	615	400,656	5.36%	−2883038.747349
Matrix 4: BaCoCa	1230	712,868	14.34%	−6099109.419374
Matrix 5: MARE	1548	723,749	9.80%	−5554777.432207
ASTRAL	2460	–	–	–

optimal models for amino acid substitutions were determined using the AUTO command. As in Cheng and Piel (2018), the gamma parameter was estimated using a model of rate heterogeneity with four discrete rates. Parsimony trees constructed by RAxML v.8.2.11 were used to initiate the searches. Bootstrapping was performed in ExaML v.3.0.2 with the same parameters as above, with 200 replicates generated per dataset.

The bootstrapped set of trees corresponding to the matrix 1 dataset served as an input for RogueNaRok (Aberer et al., 2012) to identify potential rogue taxa. RogueNaRok was ran under default conditions, and no taxa were marked as too essential to remove.

ASTRAL multispecies coalescent analysis (Mirabab et al., 2014) was also used as a sixth approach, but as it is particularly sensitive to paralogy, we implemented an additional method of ortholog filtering. Rooted gene trees for each OG were generated by RAxML v.8.2.11 (Stamatakis, 2014), and these trees were pruned to further excise putatively paralogous sequences using a custom script. This script, a modification from the original pipeline, uses the rooted trees to group each taxon into bins based on the distribution of root-to-tip lengths. The script prunes small subsets of taxa with root-to-tip lengths that greatly exceed the bulk of other taxa on the expectation that truly orthologous genes approximate a molecular clock. These pruned gene trees served as inputs for ASTRAL, which estimated the coalescent species tree.

2.3. Gene jackknifing

Analysis of large phylogenomic data sets can result in maximum support for nodes even if the actual support is rather spurious in the data (Philippe et al., 2011). In such cases, wrong topologies might be inferred with high confidence due to the effects of even very minor cross-contamination, paralogy, unevenly distributed missing data, or alignment errors remaining in the data set despite stringent filtering. To better understand if any of the nodes in our phylogenomic analyses may be affected by such phenomena, gene jackknifing was carried out following Irisarri et al. (2017). We generated 100 gene jackknife replicates for concatenated OG subsample alignments of 10,000, 100,000, 500,000 and 1,000,000 amino acids in length. Maximum likelihood trees for each replicate were estimated and majority rule consensus trees for each of the four differently sized jackknife matrices were computed from these replicates using RAxML v.8.2.11. From these files, we examined selected nodes in our phylogeny to see whether they stabilized with high support values with smaller data subsets (strong phylogenetic signal), or whether they required a larger proportion of our original data (limited phylogenetic signal).

2.4. Ancestral state reconstruction of urticating setae

Two separate character matrices were constructed for urticating setae among New World species in our phylogeny in accordance with the literature (e.g. Bertani and Guadanucci, 2013; Hüsser, 2018). We first constructed a multistate character matrix, and assigned urticating setae of types I, III and IV to different taxa from the subfamily Theraphosinae (after Bertani and Guadanucci, 2013), type II to Aviculariinae (Fukushima and Bertani, 2017), and type V solely to the genus *Ephebopus* (Foelix et al., 2009; Marshall and Uetz, 1990b) under the assumption that the different types of setae are homologous. We also constructed a binary character matrix to detail the presence or absence of any urticating setae in a taxon, regardless of the type of setae. Urticating setae types VI and VII, each known from only *Hemirrhagus* and *Kankuamo* respectively, were not included in this analysis due to a lack of sample material.

Our multistate character matrix was used to perform ancestral state reconstruction in Mesquite v3.51 (Maddison and Maddison, 2018) under parsimony with default parameters. Character states were treated as unordered. Our binary character matrix was used to perform stochastic ancestral state reconstructions via the "make.simmap" function in the R package "phytools" (Revell, 2012) with 10,000 simulations under an equal rates model.

3. Results

3.1. Metrics of tarantula phylogenomics

New transcriptomes were generated from 28 samples of 26 theraphosid species, plus another from the family Dipluridae. Combined with existing transcriptomic data from four additional species, we generated five OG matrices that ranged in size from 615 to 2460 OGs, with a total length of 1,096,124 amino acids for the largest "all genes" matrix of all OGs combined (Table 1). Additional derivative matrices "first reduce" and "second reduce," which decreased the matrix size by half each time through removal of genes with the greatest proportion of missing data, improved completion from 16.85% missing data in the initial matrix down only to 7.92% in 1230 OGs and 5.36% in 615 OGs respectively. Matrix 4 (BaCoCa) had 14.43% missing data despite containing the same number of OGs as matrix 2 (1203). Matrix 5 (MARE) had 9.80% missing data and contained a large subset of OGs (1548). The number of OGs for each taxon that matched to the initial spider-specific set of 4,446 pHMMs ("SPIDs") ranged from 2611 (*Psalmopoeus irminia*) to 3578 (*Thrigmopoeus* sp.). The total numbers of OGs per taxon in the "full matrix" (matrix 1) ranged from 1691 (*Linothele* sp.) to 2348 (*Thrigmopoeus* sp.). A full list of SPIDs, total OGs in matrix 1, and assembly data for all newly sequenced transcriptomes are available in Table S2. As expected by us, all conspecifics and congenerics emerged as more closely related to each other than to other taxa across all analyses.

Bootstrap values of 100% were recovered for all but two nodes (see in Fig. 2) across all six of our phylogenomic analyses with maximum likelihood. Overall, our phylogeny includes 29 theraphosid species, representing 25 genera and 11 subfamilies. Members of two other mygalomorph families, Dipluridae (*Linothele* sp.) and Nemesiidae (*Damarchus* sp.) were included as outgroups to root the phylogenetic trees.

3.2. The deep "Tarantula Tree of Life" inferred from transcriptome data

Our combined results (Fig. 2) recovered Theraphosidae as monophyletic with respect to *Linothele* and *Damarchus*. The first clade in our phylogeny, sister to the remaining theraphosids, consists of African representatives of the subfamily Eumenophorinae (*Hysterocrates* sp., *Monocentropus balfouri* and *Pelinobius muticus*), and an American member of the Ischnocolinae (*Catumiri* sp.). Next, we recover a clade that is sister to our non-Eumenophorinae (plus *Catumiri* sp.) theraphosids, consisting of Indian Thrigmopoeinae (represented by a single taxon, *Thrigmopoeus* sp.), and Selenocosmiinae from across many parts of Asia and Australasia (*Haplocosmia* sp., *Phlogiellus inermis* and *Selenocosmia javanensis*).

The next node is the only one that did not receive consistently high bootstrap support across all analyses, and is comprised of two main clades. The first of these main clades includes four subfamilies across two subclades. This first subclade includes two African subfamilies placed sister to one another — Stromatopelminae (*Heteroscodra*

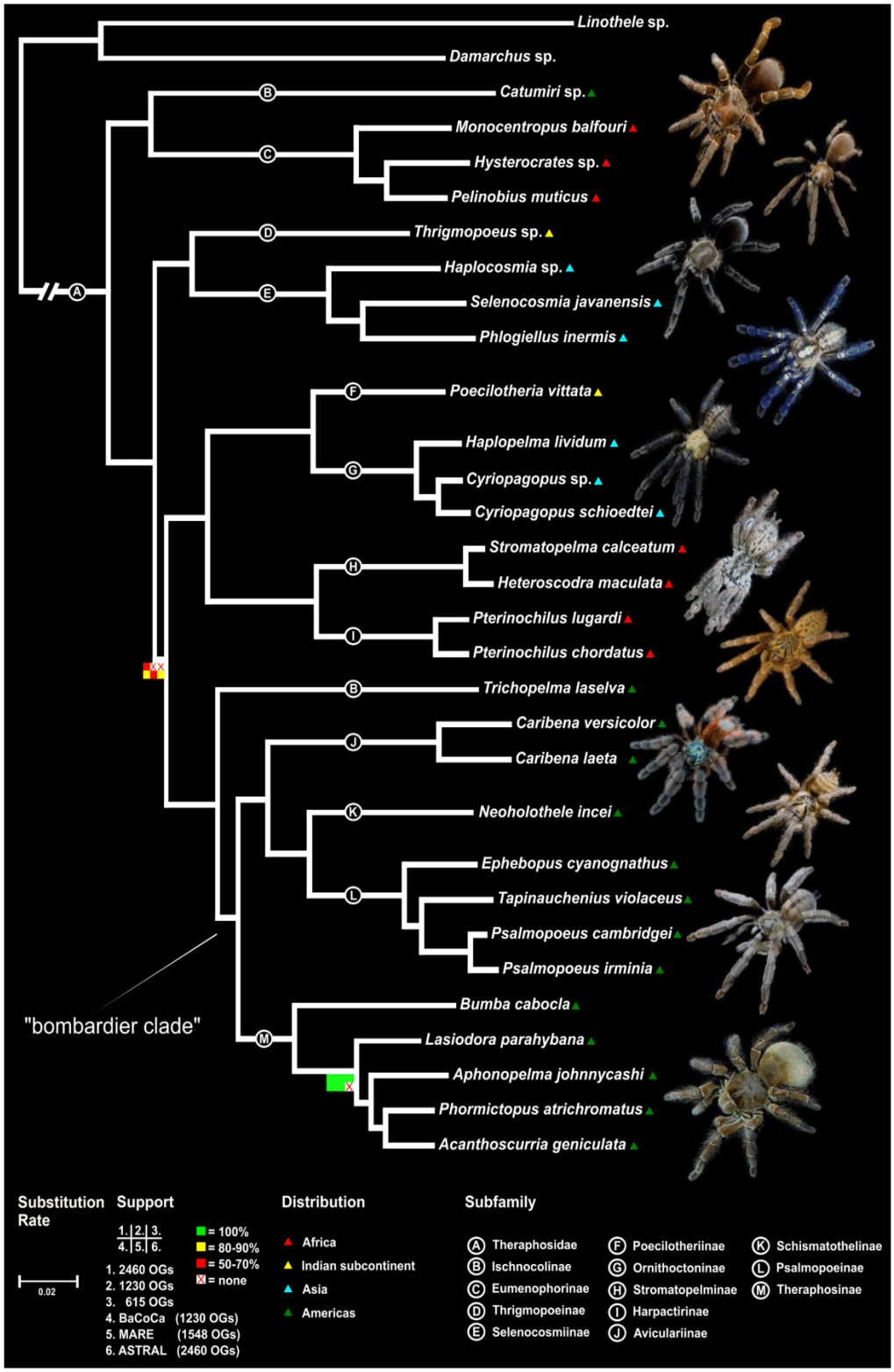


Fig. 2. Summary result of our maximum likelihood phylogenomic tree of Theraphosidae generated using ExaML. The tree obtained from matrix 1 is shown to serve as the background tree, upon which node support values for each of the other five analyses are superimposed. Subfamilial designations, support values, substitution rate, and geographic range are provided. Bootstrap support values are 100% unless otherwise indicated. Images depict representative taxa of included subfamilies. Photographs courtesy of Bastian Rast.

maculata and *Stromatopelma calceatum*) and Harpactirinae (*Pterinochilus chordatus* and *Pterinochilus lugardi*). The second subclade consists of Poecilotheriinae (two replicates of *Poecilotheria vittata* generated from different individuals), which are endemic to the Indian subcontinent, and their sister group Ornithoctoninae (*Haplopelma lividum*, *Cyriopagopus* sp. and *Cyriopagopus schioedtei*), members of which are

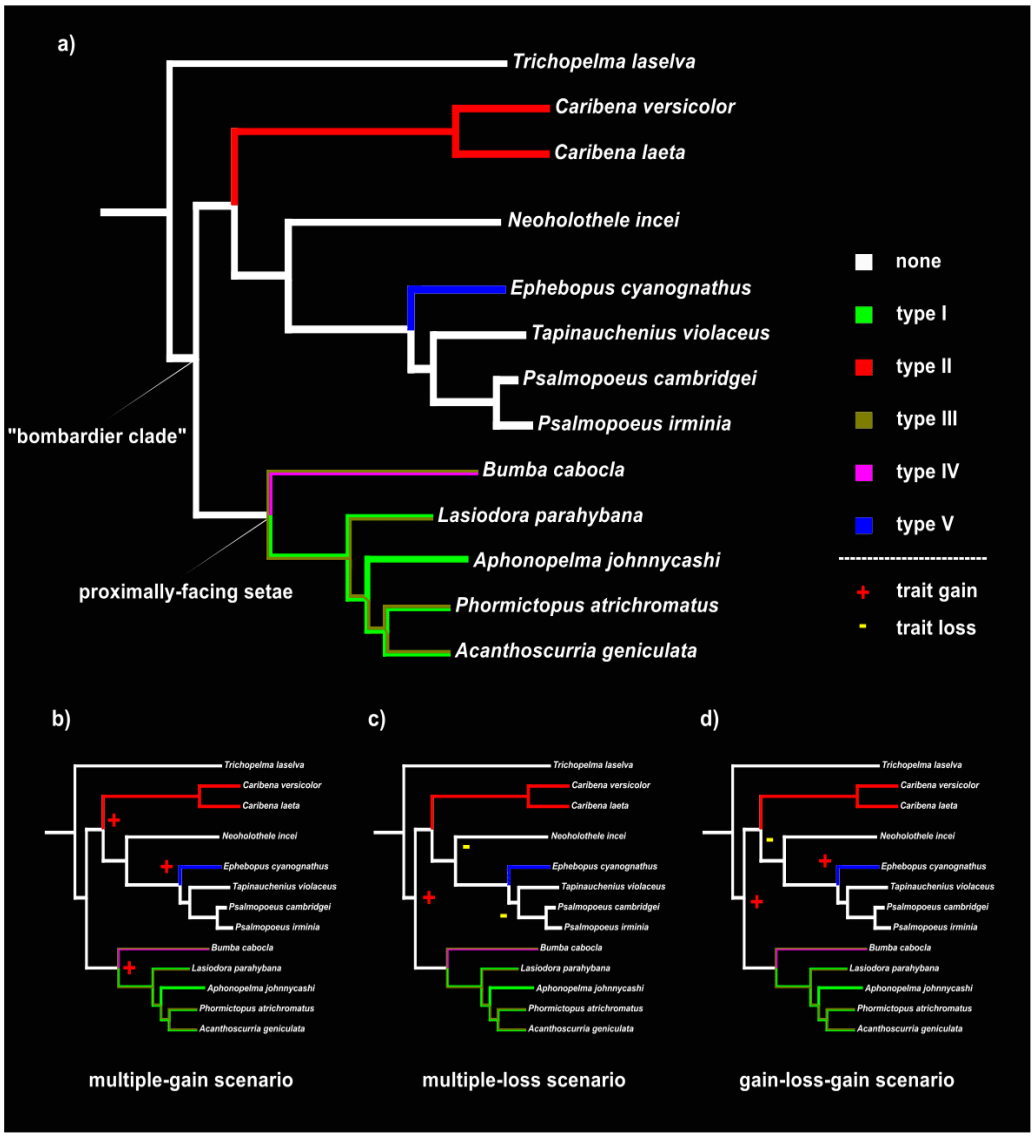


Fig. 3. Overview of urticating setae evolution among theraphosids, with three different hypotheses. (a) The ancestral state reconstruction of different urticating setae types among our "bombardier" clade, and the shared ancestry of "proximally-facing" setae types I, III and IV. (b) Prevailing hypothesis whereby urticating setae evolved three separate times independently in Theraphosinae, Avicullariinae and *Ephebopus* (Psalmopoeinae). (c) Hypothesis of a single origin for urticating setae with one origin leading to types I, II, III and IV in Theraphosinae and Avicullariinae, then subsequently being lost in both Schismatothelinae and then Psalmopoeinae (after the divergence of *Ephebopus*). (d) Hypothesis of two potential independent origins for urticating setae, with one origin again, then subsequently being lost when the common ancestor for Schismatothelinae and Psalmopoeinae diverged from Avicullariinae, before re-emerging in Psalmopoeinae via *Ephebopus* as type V setae.

found throughout Southeast Asia and parts of China. The second main clade exclusively harbors representatives from five subfamilies of Theraphosidae from the Americas, of which a member of Ischnocolinae (*Trichopelma laselva*) emerged as sister to the remaining four subfamilies. The rest of this American clade is further divided in two main subclades. The first of these includes only members of the subfamily Theraphosinae (*Acanthoscurria geniculata*, *Aphonopelma johnnycashi*, *Bumba cabocla*, *Lasiodora parahybana* and *Phormictopus atrichomatus*). The second clade consists of three subfamilies, with Avicullariinae (*Caribena versicolor* and *Caribena laeta*) as sister group of both Schismatothelinae (two replicates of *Neoholothele incei*) and Psalmopoeinae, with *Ephebopus* (*Ephebopus cyanognathus*), as sister to the remaining species (*Psalmopoeus cambridgei*, *Psalmopoeus irminia* and *Tapinauchenius violaceus*).

With the exception of the "ischnocoline" species *Catumiri* sp., we have recovered all New World theraphosids in our phylogeny as a monophyletic group. The subfamilies Eumenophorinae, Harpactirinae, Stromatopelminae, Theraphosinae, Selenocosmiinae, Ornithoctoninae, Psalmopoeinae and Avicullariinae, included in our analyses with at least two species each, were recovered as monophyletic with 100% support across all analyses. The monophyly of Thrigmopoeinae, Poecilotheriinae and Schismatothelinae cannot be confirmed due to the availability of only a single species for each subfamily, and the known-to-be paraphyletic Ischnocolinae were recovered as such.

Contrary to prior studies, our phylogeny is distinguished by high bootstrap support values across almost all nodes, with all but two nodes recovering 100% support across all six analyses. The placement of *Lasiodora parahybana* as sister taxon to three other Theraphosinae (*Aphonopelma johnnycashi*, *Acanthoscurria geniculata* and *Phormictopus atrichomatus*) received 100% bootstrap support in most analyses. *Aphonopelma johnnycashi* was absent from matrix 5 (MARE), although adjacent nodes nonetheless received 100% bootstrap support. In the ASTRAL topology, *Aphonopelma johnnycashi* was instead recovered as the sister taxon to *Lasiodora parahybana*, *Acanthoscuria geniculata* and *Phormictopus atrichomatus* and received 100% bootstrap support. The node defining the clade placed sister to Selenocosmiinae + Thrigmopoeinae

received bootstrap values of 55% (matrix 1), 90% (BaCoCa), 69% (MARE) and 82% (ASTRAL). Matrices 2 and 3 recovered an alternative topology where Selenocosmiinae + Thrigmopoeinae branched alongside Poecilotheriinae/Ornithoctoninae/Harpactirinae/Stromatopelminae, with bootstrap values of 81% and 56% respectively. Despite this, RogueNaRok reported a relative bipartition information content value (RBIC) of 0.985 and did not recommend that any taxa should be excluded.

Gene jackknifing confirmed high support (Gene Jackknife Proportions, GJP > 70%) already with gene sampling at 10 Kbp, and GJP = 99–100% with 100 Kbp for the majority of nodes (Supplementary Table S4). This suggests an extraordinarily strong phylogenetic signal in our data set. As the only two exceptions, (i) the node joining *Lasiodora parahybana* + *Aphonopelma johnnycashi* received GJP < 50% at 10 and 100 Kbp and only stabilized (GJP = 71% and 100%) with 500 Kbp and 1 Mbp), and (ii) the node defining the clade sister to *Thrigmopoeus* + Selenocosminae (which also received poor bootstrap support in several analyses; Fig. 2) received poor GJP around or below 50% up to inclusion of 500 Kbp, and only started stabilizing at 72% with 1 Mbp data sets.

3.3. Three possible scenarios for the evolution of urticating setae

We identified a clade in our phylogeny that includes the members of four New World theraphosid subfamilies: Avicullariinae, Psalmopoeinae (including *Ephebopus*), Schismatothelinae and Theraphosinae. All the species in our phylogeny that possess urticating setae, and who have been observed to engage in hair-bombardment behavior, are included in these subfamilies. Hence, we will refer to this clade as the "bombardier clade." While urticating setae types I, III and IV are present in Theraphosinae, one of the two clades arising from the earliest internal split, type II setae are only found in some members of its other clade, corresponding to the Avicullariinae. Schismatothelinae lack urticating setae, as do most Psalmopoeinae genera (e.g. *Psalmopoeus* and *Tapinauchenius*). *Ephebopus* (represented in our phylogeny by *Ephebopus cyanognathus*) is the only known genus within Psalmopoeinae that possesses urticating setae (type V). This array of urticating setae types found in our sampled taxa are coded in supplementary Table S3. Our ancestral state reconstruction of urticating setae evolution within the bombardier clade under parsimony fit the characters onto our topology with 7 steps. The result (Fig. 3a) illustrates the different urticating setae types traced through the bombardier clade. Based on this, we infer three potential evolutionary histories for urticating setae – (i) "multiple gain" scenario (Fig. 3b), (ii) "multiple loss" scenario (Fig. 3c), and (iii) "gain-loss-gain" scenario (Fig. 3d). The output from our stochastic character mapping of urticating setae can be found in supplementary Fig. S5, and appears to support the "multiple gain" scenario in Fig. 3b.

4. Discussion

4.1. A first phylogenomic overview of Theraphosidae

This study marks the first time in which phylogenomic data has been applied to study the evolution of theraphosid spiders. We opted for a transcriptomic approach because low to moderate coverage can be an efficient means to obtain sequences of large numbers of protein-coding ortholog genes from the nuclear genome (e.g. Irisarri et al., 2017) and therefore constitutes a promising method for resolving deeper nodes in the tarantula tree of life. In our case, all but two nodes were recovered consistently with 100% bootstrap support in all phylogenomic analyses. As a result, we confidently infer for the first time many of the relationships between subfamilies and their placements within the overall theraphosid tree. Although both the number of SPIDs and total number of OGs recovered in matrix 1 was generally greater in taxa from which whole-body samples were taken as compared to taxa from which only leg material was sequenced, this did not prevent conspecifics and congenerics from emerging as more closely related to each other than to other taxa across all analyses. Despite the consistently high bootstrap support across our phylogeny, it is important to be mindful of the fact that Theraphosidae is a large and diverse family. While we have representative taxa for all major theraphosid groups minus Selenogyrinae, our taxon sampling is quite limited in the context of the entire family.

Consistent with previous molecular studies, such as Lüddecke et al. (2018), we recovered the subfamily Eumenophorinae close to the base of our phylogeny. However, we also recovered *Catumiri* sp. as its sister group — a member of the taxonomically unresolved Ischnocolinae, and not included in Lüddecke et al. (2018). It has been suggested that members of Ischnocolinae may represent some of the most primitive members of Theraphosidae (after Raven, 1985; Schmidt, 2003; Guadanucci, 2014), which is at least partially supported by our results given that we found *Catumiri* placed near the base of our phylogenetic tree. Interestingly, despite differences in outgroup choice, our topology is consistent with the subclade of Theraphosidae in the Opatova et al. (2019, in prep) phylogeny, which placed *Catumiri parvum* as sister to their eight other theraphosid species, further suggesting an early separation for this genus.

The next earliest split leads to another clade composed of two Asian subfamilies: Selenocosmiinae and Thrigmopoeinae. The former is found in Southeast Asia (with some found on the Indian subcontinent and China) and Australasia, while the latter is found exclusively in India, and has received little attention in previous comprehensive studies of theraphosids. Yet, Thrigmopoeinae has a long history of taxonomic revisions and many synonyms exist in this group (e.g. Prasanth and Sunil Jose, 2014; Sanap and Mirza, 2014; Sankaran and Sebastian, 2018). In previous morphological analyses (e.g. Raven, 1985; Guadanucci, 2014), Ornithoctoninae were found to be the closest relatives to Thrigmopoeinae. The molecular analysis of Lüddecke et al. (2018) was also in line with this relationship, placing Thrigmopoeinae and Poecilotheriinae together as a clade sister to Ornithoctoninae. This hypothesis was further supported from the perspective of biogeography, as the placement of Thrigmopoeinae with Poecilotheriinae results in a clade of taxa endemic to the Indian subcontinent. Contrary to all previously mentioned studies, our phylogenomic hypothesis supports a closer relationship of Thrigmopoeinae with Selenocosmiinae, which also has some representatives from Indian subcontinent as part of this subfamily's broad geographic distribution. Therefore it rejects a placement of Thrigmopoeinae near Ornithoctoninae which are more exclusively restricted to Southeast Asia and parts of China.

Next is a clade that contains four subfamilies across two subclades. The two African subfamilies (Harpactirinae and Stromatopelminae) form the first subclade, and two Asian subfamilies (Ornithoctoninae and Poecilotheriinae) form the second. Although the molecular results of Lüddecke et al. (2018) already recovered the close relationship of Harpactirinae and Stromatopelminae, this connection has been controversial in the past. With morphological data, Stromatopelminae have historically been placed with Eumenophorinae (Raven, 1985), or inside Avicullariinae (West et al., 2008; Guadanucci, 2014; Fukushima and Bertani, 2017). Harpactirinae were otherwise thought to be more closely related to Theraphosinae and Avicullariinae (Raven, 1985), or even Schismatothelinae (Guadanucci, 2014). However, these hypotheses are rejected by our phylogenomic analysis, and we instead validate the close relationship of Harpactirinae and Stromatopelminae, supported by some morphological data (Gallon, 2003). Regarding the sister clade, the placement of Poecilotheriinae in the theraphosid tree has long been the subject of debate, with its members having most recently been included in Selenocosmiinae (e.g. Kambas, 2019, following Raven, 1985). The closer relationship between Poecilotheriinae and Ornithoctoninae was only recently revealed by molecular data in Lüddecke et al. (2018), and is here further supported by our far more comprehensive transcriptomic approach.

A further major clade identified in our study consists exclusively of taxa from the Americas (our "bombardier" clade), and includes another member of Ischnocolinae, the Central American *Trichopelma laselva*,

emerging as sister to the rest. The clade is then divided into the subfamily Theraphosinae, and three other subfamilies Avicullariinae, Psalmopoeinae and Schismatothelinae that together form another clade. In previous studies, the subfamilies of our newly established bombardier clade were scattered across the tarantula phylogeny, or formed unresolved polytomies with others (e.g. Raven, 1985; Guadanucci, 2014). In our study, for the first time, we recover the majority of New World theraphosids (excluding *Catumiri* sp.) together as a well-supported monophyletic group. Pérez-Miles et al. (1996) suggested that Theraphosinae and Avicullariinae are probably sister to one another based on the fact that defensive displays via abdominal movements are only known for these groups, which is consistent with our results in that Theraphosinae is sister to the clade containing Avicullariinae, Schismatothelinae and Psalmopoeinae. The placement of Theraphosinae is remarkable, as it differs from a wide range of previously published studies (e.g. Raven, 1985; Guadanucci, 2014; Lüddecke et al., 2018), all of which placed Theraphosinae closer to the base of their respective trees, although that alternative placement did not receive strong support in the molecular study of Lüddecke et al. (2018). The placements of Avicullariinae, Psalmopoeinae and Schismatothelinae together far from the root of our tree is consistent with the multigene results of Lüddecke et al. (2018), which also recovered a corresponding clade for these, in line with other results from Sanger-sequenced gene fragments (Turner et al., 2018; Hüsser, 2018).

4.2. Implications for subfamilial validity in Theraphosidae

While Theraphosidae itself and most of its well-established subfamilies, namely Avicullariinae, Eumenophorinae, Harpactirinae, Ornithoctoninae, Selenocosmiinae, Stromatopelminae and Theraphosinae, have emerged as monophyletic from our analysis, there are nevertheless some novel results that encourage us to reconsider the validity of certain subfamilies.

The subfamily of Poecilotheriinae from India and Sri Lanka has long been one of the most controversial from a taxonomic perspective. It was initially established as a unique lineage for just two genera including the genus *Poecilotheria*, which was later transferred to the Selenocosmiinae subfamily primarily due to the morphology of their stridulatory organs (e.g. Raven, 1985). However, this placement has always been controversial (e.g. Schmidt, 1995, who argued instead for a monotypic Poecilotheriinae), and no universally accepted consensus surrounding the placement of *Poecilotheria* has so far been reached by morphological analysis. Molecular studies, on the other hand, recently demonstrated that the genus is genetically very distinct from Selenocosmiinae, and instead seems to share a more recent common ancestor with Ornithoctoninae. It has therefore been suggested that the placement of *Poecilotheria* inside Selenocosmiinae is not justified, and that the validity of Poecilotheriinae as a subfamily should be considered (Lüddecke et al., 2018). In agreement with this, our phylogenomic data also supports the distinction of *Poecilotheria* from Selenocosmiinae and instead supports their closer relationship with Ornithoctoninae as previously suggested by much smaller molecular datasets (Lüddecke et al., 2018; Turner et al., 2018). However, *Poecilotheria* differs from Ornithoctoninae not only genetically but further in respect to morphology: They lack plumose bristles on the retrolateral surface of chelicerae, which are diagnostic for Ornithoctoninae (Von Wirth and Striffler, 2005; West and Nunn, 2010). Given this genetic and phenotypic distinction of *Poecilotheria*, we follow Schmidt (1995) to emphasize in agreement with Lüddecke et al. (2018) that the unique subfamilial status Poecilotheriinae Simon, 1892 stat. rev. should be adopted, with *Poecilotheria* Simon, 1885 as its type genus. This group will require a more detailed revision, ideally based on morphological as well as molecular data from a larger set of species, in the near future.

Among all theraphosid subfamilies, the Ischnocolinae are recognized as the taxonomically most problematic group. These are mostly relatively small spiders, often referred to as “dwarf tarantulas.” that have the widest distribution among all theraphosids, with species being found on many continents, such as the Americas, Asia, Africa and even Europe. Originally, Ischnocolinae was described by Simon (1892) with all tarsi divided being the main diagnostic morphological character. However, it later became apparent that this trait is not unique to ischnocolines (e.g. Schmidt, 2003). In addition, other key tarantula traits such as urticating setae, stridulatory organs or basally connected spermathecae were never found in Ischnocolinae, resulting in a paucity of other informative characters (after Raven, 1985; Schmidt, 2003; Guadanucci, 2014). Facing this lack of diagnostic information, morphological studies have struggled to characterise Ischnocolinae relationships. It is hence not surprising that the subfamily as a whole has earned a reputation for being something akin to a “holding place” for theraphosids that cannot be placed into other lineages. Consequently, morphology-based studies of Ischnocolinae recovered the group as paraphyletic. In his broad morphological study, Raven (1985) identified at least four independent clades within Ischnocolinae, with three of them proposed as the earliest branching theraphosids in his phylogeny. Later, Guadanucci (2014) performed phylogenetic analyses on all known genera of Ischnocolinae, plus representatives of the other subfamilies, again based on morphology. He identified the “core group” (Ischnocolinae *sensu stricto*) to include the type genus of the subfamily (*Ischnocolus*) as well as *Acanthopelma*, *Holothele*, *Reichlingia* and *Trichopelma* (Guadanucci, 2014). Some other Ischnocolinae (*sensu lato*) were placed in a separate clade, including *Schismatothele* plus others including several former *Holothele* later transferred to other genera (including *Neoholothele*), as the newly erected subfamily of Schismatothelinae (Guadanucci, 2014; Guadanucci & Weinmann, 2015). Remaining ischnocoline genera, such as *Heterothele*, *Nesiergus* and *Catumiri*, were dispersed across the phylogeny, and their phylogenetic affinities remained largely unresolved. Our phylogenomic analysis, for the first time, represents a non-morphology based multigene hypothesis for theraphosid subfamily relationships that includes more than one member of Ischnocolinae. We included *Trichopelma laselva* (Ischnocolinae *sensu stricto*) *Neoholothele incei* (a former ischnocoline now in Schismatothelinae), and *Catumiri* sp. (representing Ischnocolinae *sensu lato*). In agreement with Guadanucci (2014), we detected no obvious close phylogenetic affinity between these three “ischnocoline” taxa, and instead found that *Trichopelma laselva* (supposedly Ischnocolinae *sensu stricto*) shares a more recent common ancestor with other Neotropical subfamilies, including Avicullariinae and Theraphosinae, than it does with *Catumiri* sp. (Ischnocolinae *sensu lato*). Interestingly, members of Schismatothelinae were also found inside this “bombardier clade,” nested between Avicullariinae and Psalmopoeinae, in a placement that sufficiently distinguishes them from other sampled “ischnocolines.” In contrast, *Catumiri* sp. was found as one of the earliest branching theraphosids in our tree as a sister group to Eumenophorinae. In the recently published molecular phylogeny of Lüddecke et al. (2018), only a single member of Ischnocolinae was included (*Nesiergus insulanus* from the Seychelles), which was seemingly most closely related with Asian Selenocosmiinae. Put together, we conclude that, as implied by Guadanucci (2014), the “Ischnocolinae” as currently defined represent multiple independent theraphosid lineages. We emphasize the need for a future large scale integrative studies that are inclusive of problematic members from across the tarantula family, with a particular focus including multiple “ischnocoline” from diverse geographical areas, plus several representatives of Schismatothelinae, which has also been shown to likely be paraphyletic (Lüddecke et al., 2018; Hüsser, 2018). Unfortunately, we were unable to include additional Schismatothelinae taxa in this study, and therefore cannot test the monophyly or paraphyly of the subfamily.

Finally, *Ephebopus* is one of the most controversially placed genera of Theraphosidae in recent times worthy of mention. This genus has only recently been recognized as part of the Psalmopoeinae which otherwise consists of the genera *Psalmopoeus*, *Pseudoclamoris* and *Tapinauchenius* (after Hüsser, 2018). This subfamily is represented in our study by four species: *Ephebopus cyanognathus*, *Psalmopoeus cambridgei*, *Psalmopoeus irminia* and *Tapinauchenius violaceus*. Various members of this clade have commonly been considered as members of

the Avicularíinae (e.g. Schmidt, 2003; West et al., 2008; Fukushima and Bertani, 2017), or Selenocosmiinae (e.g. Guadanucci, 2014), although some were also placed in a separate subfamily, "Psalmopoeinae" by other authors (Samm and Schmidt, 2010). Recently, Lüddecke et al. (2018) and Turner et al. (2018) found molecular evidence for the validity of the subfamily, which shortly thereafter was supported by both morphological and molecular data in Hüsser (2018). Consistent with these studies, our phylogenomic analysis further supports the distinction between Avicularíinae and Psalmopoeinae (including *Ephebopus*), as evidenced by both the high bootstrap support, and by the fact that, once again, Schismatothelinae were found to be closer to Psalmopoeinae than to Avicularíinae, which was placed sister to both of these other subfamilies.

4.3. Morphology meets phylogenomics: Insights into the evolution of urticating setae and defense strategies in tarantulas

Based on morphological similarities, and the existence of intermediate forms of urticating setae, Bertani and Guadanucci (2013) proposed that "primitive," short setae subsequently became modified to urticating setae, and that type III represents the ancestral state that gave rise to types I and IV (see also Pérez-Miles, 2002). This appears to make sense, as types I, III and IV are all found in the subfamily Theraphosinae and share several morphological similarities — including attachment via a stalk, a distal penetrating tip, and most notably the numerous and sometimes prominent "proximally-facing" barbs along the shaft of the setae. Types II (in Avicularíinae) and V (only known to be present in *Ephebopus*, Psalmopoeinae) in contrast possess barbs that point distally, but beyond this neither type shows much structural similarity with each other, nor with those in Theraphosinae. This led Bertani and Guadanucci (2013) to conclude that urticating setae evolved at least three times in Theraphosidae, however, their evolutionary history remained speculative. Our data supports the hypothesis that the "proximally-facing" types I, III and IV share a common origin in the subfamily Theraphosinae (Fig. 3a). Further, our ancestral state reconstructions under both parsimony and stochastic criteria indicate that, as in Bertani and Guadanucci (2013), urticating setae in Theraphosinae, Avicularíinae and *Ephebopus* (Psalmopoeinae) could each represent an independent origin for urticating setae, regardless of the type (multiple-gain scenario, Fig. 3b). Since the outputs from both our parsimony and stochastic reconstructions support this, Fig. 3b appears to represent the most probable scenario.

Type V urticating setae are remarkable in that they are not localized on the opisthosoma, but are instead found on the palpal femora (Marshall and Uetz, 1990b; Foelix et al., 2009). Based on the position criterion of type V setae, they may not be homologous to other types of urticating setae, and both Bertani and Guadanucci (2013) and Hüsser (2018) considered type V setae in *Ephebopus* to be an autapomorphy for the genus. Combined with our robust placement of *Ephebopus* within Psalmopoeinae and the absence of any urticating setae in Schismatothelinae and the remaining Psalmopoeinae, Fig. 3b supports the idea that type V setae are autapomorphic in *Ephebopus* and could represent a separate evolutionary event for the gain of urticating setae.

Although Fig. 3b appears to be the most probable scenario, our parsimony reconstruction also recovered two additional alternatives. Urticating setae diversity could be explained by modifications upon an ancestral structure that originated once before the divergence of Avicularíinae and Theraphosinae (i.e. earlier than was previously assumed), and then diversified throughout these clades. In this scenario, the trait was subsequently lost twice — once in Schismatothelinae, and again later in some Psalmopoeinae (multiple-loss scenario, Fig. 3c). Another scenario was recovered where urticating setae could have again first originated before the divergence of Avicularíinae and Theraphosinae, but were then lost in the ancestor to both Schismatothelinae and Psalmopoeinae when they diverged, before subsequently being regained in *Ephebopus* (gain-loss-gain scenario, Fig. 3d). Given the structural differences between different types, the conjunction of multiple types on the abdomens of certain lineages, and the fact that type V are localized to a different area of the body than other types, we suggest that some or all types (and particularly the latter) are better treated as separate morphological characters in future analyses. While Fig. 3b represents our preferred hypothesis, the scenarios presented in Fig. 3c and d should also be considered given that two additional types of urticating setae have been described for taxa not represented in our analyses.

Perafán et al. (2016) described a monotypic genus *Kankuamo* from the subfamily Theraphosinae, which most notably has a novel type of urticating setae (type VII). None of the previously described types displayed much morphological similarity with type II beyond the position criterion (Bertani and Guadanucci, 2013). However, type VII shares several characteristics with type II setae (Perafán et al., 2016). It has even been suggested that urticating setae in *Kankuamo* share a similar "direct contact" method with type II setae — i.e. where they are rubbed directly from the opisthosoma onto the target instead of projected into the air by "bombardment" (with the exception of *Caribena versicolor* — a species which has been observed to flick type II hairs, as in Bertani et al., 2003). However, the release mechanism in *Kankuamo* has not been directly observed. Unfortunately, no phylogenetic study has included members of *Kankuamo*, but its placement will likely be crucial in resolving the evolutionary history of urticating hairs. Given their remarkable structural similarity, attachment mechanism and potentially similar method of release, it is plausible that the type II setae in Avicularíinae are derived from type VII in *Kankuamo* (or vice versa), providing a link between the type II of Avicularíinae and types I, III & IV in Theraphosinae. This could serve to challenge the prevailing hypothesis presented in Fig. 3b (i.e. that urticating setae independently emerged three times), and potentially lend more support to either of the alternative hypotheses illustrated in Fig. 3c and Fig. 3d respectively.

Some members of the theraphosine genus *Hemirrhagus* possess type VI urticating setae (as described by Pérez-Miles, 1998), which are morphologically very similar to type V (Bertani and Guadanucci, 2013), but are localized on the opisthosoma. *Hemirrhagus* contains some species notable for their unusual troglobitic lifestyles (Mendoza and Francke, 2018). The occupation of subterranean ecological niches, and the switch to a cave-dwelling lifestyle, has probably contributed to the evolution of unique and unusual morphological traits in *Hemirrhagus*. On the other hand, other traits have been lost in some species, including eye pigmentation and urticating setae, which have been interpreted as derived reversals (Mendoza and Francke, 2018). If members of this genus can be included in future studies, it may pave the way to study general patterns of ecological adaptation, adaptive potential and trait evolution in Theraphosidae. Unfortunately, we were unable to obtain members of *Hemirrhagus* for this work, but in our view, the genus represents an interesting taxon that will be vital to consider in further studies on the evolution of tarantulas. We emphasize that future studies specifically addressing the evolution of urticating setae within the bombardier clade should focus on any additional insights that either *Kankuamo* or *Hemirrhagus* may provide.

Having provided support for the close relationship between Theraphosinae and Avicularíinae, as initially alluded to by Pérez-Miles et al. (1996), we suggest that the evolution of urticating setae in Theraphosidae represents a key innovation that could have facilitated their rapid adaptive radiation in the New World. This hypothesis is further supported by the fact that the Theraphosinae subfamily, in which the diversity and abundance of urticating setae is most conspicuous, represents the largest radiation of known genera and species within Theraphosidae. In fact, this subfamily accounts for about 50% of the known theraphosid diversity (Kambas, 2019), with several types of urticating setae (e.g. Schmidt, 2003). It appears reasonable to assume that this high success in the New World might be facilitated via the invention of defensive urticating setae. Like all spiders, members of the "bombardier clade" are venomous, and can deliver defensive bites to would-be predators. Generally, venom serves the purposes of subduing prey, deterring competitors or defense against predators (Casewell

et al., 2013). However, venom is foremost a costly resource, and any utilization that conserves as much venom as possible should be considered a competitive advantage (Morgenstern and King, 2013). Although tarantulas in general are not particularly venomous to vertebrates such as humans, a small fraction of species seem to be capable of delivering medically significant bites. As scientific literature currently only treats members of *Poecilotheria* as potentially dangerous (Hauke and Herzig, 2017), bites from other species of African and Asian theraphosids have anecdotally been reported to cause relatively severe and painful symptoms (Escoubas and Rash, 2004). It is intriguing that all species from which medically significant bites and potent venom are reliably reported lack the defensive mechanism of urticating setae. The study of venom evolution elsewhere has taught us that venom components in ancient lineages, such as spiders, often evolve under strong purifying selection pressures, which predicts that over the course of evolution, only indispensable components remain present in venom cocktails (Sunagar and Moran, 2015). In light of this, it becomes possible to imagine a scenario in which the evolution of inexpensive urticating setae may have led to the loss of costly venom components, possibly as an adaptive response to a new environment.

4.4. Contentious nodes

Despite the overall high support and strong phylogenetic signal in our study, as revealed from maximum bootstrap and high gene jackknife support for most nodes, two nodes were not strongly supported in all analyses and require more detailed discussion.

The first of these corresponds to the *Lasiodora parahybana* + *Aphonopelma johnnycashi* node, which was not recovered by ASTRAL. In matrices 1–4, *Lasiodora parahybana* was recovered as the sister taxon to *Aphonopelma johnnycashi*, *Acanthoscurria geniculata* and *Phormictopus atrichomatus*. Matrix 5 (MARE) did not include *Aphonopelma johnnycashi*, but adjacent nodes still received 100% bootstrap support. ASTRAL instead placed *Aphonopelma johnnycashi* as sister to *Lasiodora parahybana*, *Acanthoscurria geniculata* and *Phormictopus atrichomatus*. Although this means the relationship did not receive 100% support across all analyses, the general relationship between the two was recovered across all six analyses, and we are quite confident in the validity of this node based on the strong bootstrap support from matrices 1–5, which each recovered *Lasiodora parahybana* as sister to *Aphonopelma johnnycashi*. Of course, it should be stressed that this relationship exists from the perspective of our dataset, which contains relatively few Theraphosinae. The addition of further theraphosine taxa would almost certainly result in a slightly more distant relationship between *Lasiodora* and *Aphonopelma*.

Our second poorly supported node may instead highlight some key limitations of the core-ortholog approach. Our phylogeny shows Selenocosmiinae + Thrigmopoeinae as early branching subfamilies, but they were found to branch alongside Poecilotheriinae/Ornithoctoninae/Harpactirinae/Stromatopelminae in matrices 2 and 3. This topology was not strongly supported in any of our analyses. Prior to the addition of the gene tree pruning step to further minimise paralogy, this node had even weaker support — it was only recovered in 3 of our 6 topologies (as opposed to 4 of 6 post-pruning), with matrix 3 yielding the greatest highest value of 62% for that node despite no longer recovering it post-pruning. Having performed a rogue taxon analysis with RogueNaRok (Aberer et al., 2012) which did not recommend that any of our taxa be excluded, we caution future analyses that use a core-ortholog approach that this part of the tarantula phylogeny may be particularly susceptible to some influence of paralogous genes, and that careful filtering is warranted. Yet, the poor support of this node remains somewhat of an enigma, although the very short length of the branch leading to it (Fig. 2) suggests that the respective evolutionary divergences might have taken place in a very short time with few phylogenetically informative substitutions.

Possibly, the support values for this node could be improved with a greater breadth of taxon sampling. In particular, we believe that the addition of a member of Selenogyrinae (the only subfamily absent from our phylogeny) could be key to breaking up these long branches and resolving this node. Selenogyrinae are found throughout west / central Africa, as well as India, though the placement of the latter in this subfamily may be questionable. Given the strong biogeographic links we found across the phylogeny, it is reasonable to suppose that Selenogyrinae would emerge alongside other subfamilies found in these regions., but it seems plausible that Selenogyrinae could help resolve this uncertainty based on their geographic spread. Unfortunately, we recognize that these taxa are very difficult to procure, with new field-collection likely to be the only source for future studies.

4.5. Future perspectives

Although our phylogeny provides a robust framework for future evolutionary studies on theraphosids, there are still some potential sources of bias that should be discussed. It has been shown that taxon selection and the density of sampled taxa represent important factors that influence tree topologies, and therefore the inferred hypothesis (e.g. Pollock et al., 2002; Zwickl and Hillis, 2002). Increased taxon coverage from both within Theraphosidae and from other mygalomorph families can benefit our understanding of theraphosid evolution. Relatively little is known about the family Barychelidae, commonly known as brushed-foot trapdoor spiders. Results from other molecular datasets show that Barychelidae can be the sister group to Theraphosidae (after Hedin and Bond, 2006; e.g. Hamilton et al., 2016b; Garrison et al., 2016; Hedin et al., 2018; Opatova et al., 2019 in prep), but no transcriptome data is currently available for any members of Barychelidae. The relationships between these two families are tenuous with morphology, and some theraphosids have a long history of being treated as barychelids. For example Guadanucci (2014) determined that many genera, including *Trichopelma* (represented in our study by *T. laselva*), actually belong in Theraphosidae and not Barychelidae as was previously thought. Similarly the assignment of *Brachionopus* and *Harpactirella* as members of Barychelidae (based on Schmidt, 2002) was recently demonstrated as invalid by molecular work and instead a placement of both genera within Theraphosidae was proposed (Lüddecke et al., 2018). These examples should signal a need to more carefully re-examine the relationships between theraphosids and barychelids using molecular analyses.

Our phylogenomic analysis, despite including representatives of most major theraphosid lineages, is based on data from only 25 of the 146 currently accepted genera (i.e. 18%), and it is apparent that our dataset only captures a subset of the whole radiation of this incredibly species-rich family (World Spider Catalog, 2019). It should therefore be a high priority for future studies to expand the inherent taxon coverage within theraphosids to provide a more complete picture of the family's evolution. We suggest that the expansion of taxon sampling in future research should further be economized and focused on lineages of uncertain taxonomic status, such as Ischnocolinae, Selenogyrinae and Thrigmopoeinae, or taxa from currently neglected geographic areas such as the Indian subcontinent, Europe or Australasia. Efforts of collecting representatives of these groups are currently underway. Moreover, as previously mentioned, the addition of further taxa with unique urticating setae types will be required to capture the full gamut of diversity of this unusual character, which in turn will be invaluable to determine exactly how this trait has evolved. That said, the study of venom evolution, in particular the relationship between toxicity and urticating setae possession, might be possible in the future based on our robust phylogenomic reconstruction, especially if more taxa from the above mentioned groups are included.

5. Conclusion

Our research provides the first phylogenomic hypothesis on the evolution of tarantula spiders and helps clarify possible evolutionary scenarios pertaining to their urticating setae. Our phylogeny differs from a wide variety of previously published morphological and

molecular studies. Deeper nodes illustrating the relationships between subfamilial lineages have been recovered with strong support for the first time on a robust phylogeny, which can serve as a sturdy foundation for a diverse range of subsequent studies on these enigmatic animals.

Acknowledgements

We are grateful to the German Arachnological Society (Deutsche Arachnologische Gesellschaft, DeArGe e.V.) for providing Tim Lüddecke with funds for this study, and to the National Parks Board of Singapore for permission to collect *Phlogiellus inermis* (permit number NP/RP18-046). Support for this research was also provided by the South East Asian Biodiversity Genomics (SEABIG) center, and the Singapore Ministry of Education (grant number MOE2016-T2-2-137, and by Yale-NUS College (grant numbers IG14-SI002, IG15-SI101 R-607-265-052-121). We thank Iker Irrisari for generating the initial jackknife replicates, and David Court, Fernando Pérez-Miles, Martin Hüsser and Björn Marcus von Reumont for valuable discussions on the manuscript. We are indebted to Bastian Rast, who provided picture material that has been used to illustrate the phylogenetic tree. Martin Hüsser and Rainer Foelix kindly provided pictures that were used to assemble our figure on urticating setae. The graphical abstract was built with picture material that was taken with support of Monty Oberländer and Aileen Schlag.

Author contributions

SF and TL designed the study and acquired sample material. SF, TL and SK performed laboratory work. TL, MV and WHP attracted funding for the study. Bioinformatic work was conducted by SF, WHP, MV and DQC. HK, SJL, IW, VvW and RT contributed to experimental design and assured valid taxonomical as well as morphological assignments. SF and TL wrote the manuscript with substantial input from MV, HK, SJL, IW, VvW and RT. All authors reviewed and agreed on the manuscript. The authors declare no competing interests.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.106573.

References

Aberer, A.J., Krompass, D., Stamatakis, A., 2012. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. System. Biol. 62, 162–166. https://doi.org/10.1093/sysbio/sys078.

Battisti, A., Holm, G., Fagrell, B., Larsson, S., 2011. Urticating hairs in arthropods: their nature and medical significance. Annu. Rev. Entomol. 56, 203–220. https://doi.org/10.1146/annurev-ento-120709-144844.

Bertani, R., Boston, T., Evenou, Y., Guadanucci, J.P.L., 2003. Release of urticating hairs by *Avicularia versicolor* (Walckenaer, 1837) (Araneae, Theraphosidae). Bull. British Arachnol. Soc. 12 (9), 395–398.

Bertani, R., Guadanucci, J.P.L., 2013. Morphology, evolution and usage of urticating setae by tarantulas (Araneae: Theraphosidae). Zoologia (Curitiba) 30, 403–418. https://doi.org/10.1590/s1984-46702013000400006.

Casewell, N.R., Wüster, W., Vonk, F.J., Harrison, R.A., Fry, B.G., 2013. Complex cocktails: the evolutionary novelty of venoms. Trends Ecol. Evol. 28, 219–229. https://doi.org/10.1016/j.tree.2012.10.020.

Cooke, J.A.L., Roth, V.D., Miller, F.H., 1972. The urticating hairs of theraphosid spider. Am. Mus. Novit. 2498, 1–43.

Cheng, D.-Q., Piel, W.H., 2018. The origins of the Psechridae: web-building lycosoid spiders. Mol. Phylogenet. Evol. 125, 213–219. https://doi.org/10.1016/j.ympev.2018.03.035.

Cléton, F., Sigwalt, Y., Verdez, J.M., 2015. Tarantulas Breeding Experience & Wildlife. Edition Chimaira, Frankfurt am Main.

Ebersberger, I., Strauss, S., Haeseler, A.V., 2009. HaMStR: Profile hidden markov model based search for orthologs in ESTs. BMC Evol. Biol. 9, 157. https://doi.org/10.1186/1471-2148-9-157.

Escoubas, P., Rash, L., 2004. Tarantulas: eight-legged pharmacists and combinatorial chemists. Toxicon 43, 555–574. https://doi.org/10.1016/j.toxicon.2004.02.007.

Fernández, R., Kallal, R.J., Dimitrov, D., Ballesteros, J.A., Arnedo, M.A., Giribet, G., Hormiga, G., 2018. Phylogenomics, diversification dynamics, and comparative transcriptomics across the spider tree of life. Curr. Biol. 28, 1489–1497.

Foelix, R.F., Bastian, R., Erb, B., 2009. Palpal urticating hairs in the tarantula *Ephebopus*: fine structure and mechanism of release. J. Arachnol. 37, 292–298. https://doi.org/10.1636/sh08-106.1.

Fukushima, C.S., Bertani, R., 2017. Taxonomic revision and cladistic analysis of *Avicularia* Lamarck, 1818 (Araneae, Theraphosidae, Aviculariinae) with description of three new aviculariine genera. ZooKeys 659, 1–185. https://doi.org/10.3897/zookeys.659.10717.

Gallon, R.C., 2003. A new African arboreal genus and species of theraphosid spider (Araneae, Theraphosidae, Stromatopelminae) which lacks spermathecae. Bull. Br. Arach. 12 (9), 405–411.

Garrison, N.L., Rodriguez, J., Agnarsson, I., Coddington, J.A., Griswold, C.E., Hamilton, C.A., Hedin, M., Kocot, K.M., Ledford, J.M., Bond, J.E., 2016. Spider phylogenomics: untangling the Spider Tree of Life. PeerJ 4, e1719. https://doi.org/10.7717/peerj.1719.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Palma, F.D., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883.

Guadanucci, J.P.L., 2014. Theraphosidae phylogeny: relationships of the '*Ischnocolinae*' genera (Araneae, Mygalomorphae). Zoolog. Scr. 43, 508–518. https://doi.org/10.1111/zsc.12065.

Guadanucci, J.P.L., Weinmann, D., 2015. Description of *Neoholothele* gen. nov. (Araneae, Theraphosidae, Schismatothelinae). Stud. Neotropical Fauna Environ. 50, 221–228. https://doi.org/10.1080/01650521.2015.1110309.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., Macmanes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512. https://doi.org/10.1038/nprot.2013.084.

Hamilton, C.A., Hendrixson, B.E., Bond, J.E., 2016a. Taxonomic revision of the tarantula genus *Aphonopelma* Pocock, 1901 (Araneae, Mygalomorphae, Theraphosidae) within the United States. ZooKeys 560, 1–340. https://doi.org/10.3897/zookeys.560.6264.

Hamilton, C.A., Lemmon, A.R., Lemmon, E.M., Bond, J.E., 2016b. Expanding anchored hybrid enrichment to resolve both deep and shallow relationships within the spider tree of life. BMC Evol. Biol. 16. https://doi.org/10.1186/s12862-016-0769-y.

Hauke, T.J., Herzig, V., 2017. Dangerous arachnids—fake news or reality? Toxicon 138, 173–183. https://doi.org/10.1016/j.toxicon.2017.08.024.

Hedin, M., Derkarabetian, S., Ramirez, M.J., Vink, C., Bond, J.E., 2018. Phylogenomic reclassification of the worldśs most venomous spiders (Mygalomorphae, Atracinae), with implications for venom evolution. Sci. Rep. 8, 1636.

Hedin, M., Bond, J.E., 2006. Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): conflict and agreement with the current system of classification. Mol. Phylogenet. Evol. 41, 454–471. https://doi.org/10.1016/j.ympev.2006.05.017.

Hsiung, B.-K., Deheyn, D.D., Shawkey, M.D., Blackledge, T.A., 2015. Blue reflectance in tarantulas is evolutionarily conserved despite nanostructural diversity. Sci. Adv. 1. https://doi.org/10.1126/sciadv.1500709.

Hüsser, M., 2018. A first phylogenetic analysis reveals a new arboreal tarantula genus from South America with description of a new species and two new species of *Tapinauchenius* Ausserer, 1871 (Araneae, Mygalomorphae, Theraphosidae). ZooKeys 784, 59–93. https://doi.org/10.3897/zookeys.784.26521.

Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J.-Y., Kupfer, A., Petersen, J., Jarek, M., Meyer, A., Vences, M., Philippe, H., 2017. Phylotranscriptomic consolidation of the jawed vertebrate timetree. Nat. Ecol. Evol. 1, 1370–1378. https://doi.org/10.1038/s41559-017-0240-5.

Kambas, D., 2019. Tarantupedia TM. https://www.tarantupedia.com/ (accessed at 01 April 2019).

Katoh, K., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucl. Acids Res. 33, 511–518. https://doi.org/10.1093/nar/gki198.

Klaas, P., 1989. Vogelspinnen im Terrarium. Ulmer, Stuttgart.

Kozlov, A.M., Aberer, A.J., Stamatakis, A., 2015. ExaML version 3: a tool for phylogenomic analyses on supercomputers. Bioinformatics 31, 2577–2579. https://doi.org/10.1093/bioinformatics/btv184.

Kück, P., Struck, T.H., 2014. BaCoCa – a heuristic software tool for the parallel assessment of sequence biases in hundreds of gene and taxon partitions. Mol. Phylogenet. Evol. 70, 94–98. https://doi.org/10.1016/j.ympev.2013.09.011.

Kück, P., Meusemann, K., 2010. FASconCAT: convenient handling of data matrices. Mol. Phylogenet. Evol. 56, 1115–1118. https://doi.org/10.1016/j.ympev.2010.04.024.

Kück, P., Meusemann, K., Dambach, J., Thormann, B., Reumont, B.M.V., Wägele, J.W., Misof, B., 2010. Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. Front. Zool. 7, 10. https://doi.org/10.1186/1742-9994-7-10.

Kück, P., 2009. ALICUT: a Perlscript which cuts ALISCORE identified RSS. Version 2. Department of Bioinformatics, Zoologisches Forschungsmuseum A.Koenig (ZFMK), Bonn (Germany) (cited 2018 Jun 22).

Longhorn, S.J., Nicholas, M., Chuter, J., Vogler, A.P., 2007. The utility of molecular markers from non-lethal DNA samples of the Cites protected "tarantula" *Brachypelma vagans* (Araneae, Theraphosidae). J. Arachnol. 35, 278–292. https://doi.org/10.1636/s05-62.1.

Lüddecke, T., Krehenwinkel, H., Canning, G., Glaw, F., Longhorn, S.J., Tänzler, R., Wendt, I., Vences, M., 2018. Discovering the silk road: Nuclear and mitochondrial sequence data resolve the phylogenetic relationships among theraphosid spider subfamilies. Mol. Phylogenet. Evol. 119, 63–70. https://doi.org/10.1016/j.ympev.2017.10.015.

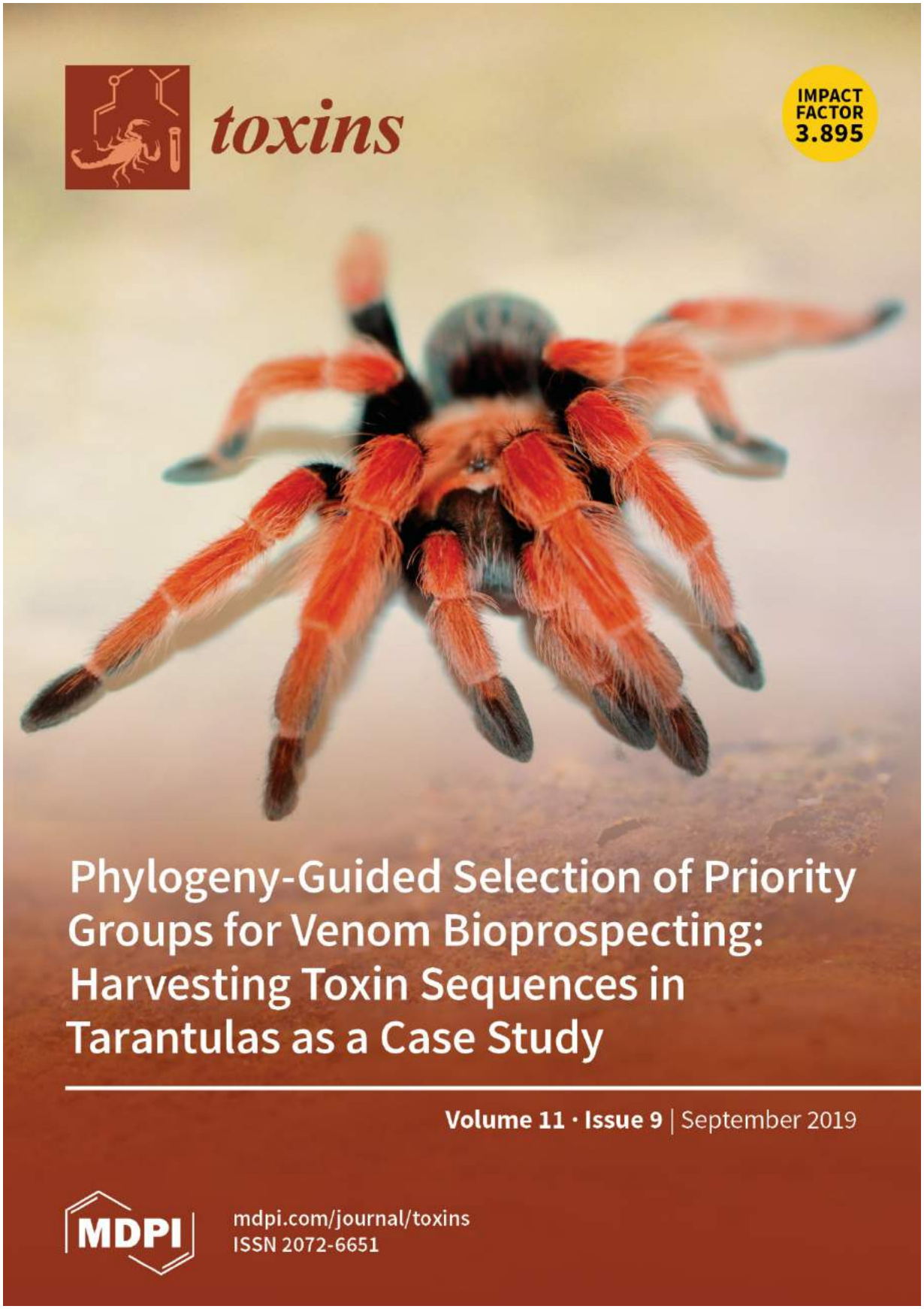
Maddison, W.P., Maddison, D.R., 2018. Mesquite: a modular system for evolutionary analysis. Version 3.51. http://www.mesquiteproject.org.

Marshall, S.D., Uetz, G.W., 1990a. Incorporation of urticating hairs into silk: a novel defense mechanism in two Neotropical tarantulas (Araneae, Theraphosidae). J. Arachnol. 18, 143–149.
Marshall, S.D., Uetz, G.W., 1990b. The pedipalpal brush of *Ephebopus* sp. (Araneae, Theraphosidae): evidence of a new site for urticating hairs. Bull. Br. Arachnol. Soc. 8 (4), 122–124.
McAnena, L., Murphy, C., O'Connor, J., 2013. "Tarantula Keratitis" a case report. Ir. J. Med. Sci. 182 (3), 349–350.
Mendoza, J., Francke, O., 2017. Systematic revision of Brachypelma red-kneed tarantulas (Araneae: Theraphosidae), and the use of DNA barcodes to assist in the identification and conservation of CITES-listed species. Invertebr. System. 31, 157–179.
Mendoza, J.I., Francke, O.F., 2018. Five new cave-dwelling species of *Hemirrhagus* Simon 1903 (Araneae, Theraphosidae, Theraphosinae), with notes on the generic distribution and novel morphological features. Zootaxa 4407, 451. https://doi.org/10.11646/zootaxa.4407.4.1.
Meyer, B., Meusemann, K., Misof, B., 2011. MARE: MAtrix REduction—a tool to select optimized data subsets from supermatrices for phylogenetic inference. Zentrum für molekulare Biodiversitätsforschung (zmb) am ZFMK, Bonn (Germany).
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541–i548. https://doi.org/10.1093/bioinformatics/btu462.
Misof, B., Misof, K., 2009. A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. Syst. Biol. 58, 21–34. https://doi.org/10.1093/sysbio/syp006.
Molur, S., Silliwal, M., Daniel, B.A., 2008. At last! Indian tarantulas on IUCN red list. Zoós Print 23, 1–3.
Morgenstern, D., King, G.F., 2013. The venom optimization hypothesis revisited. Toxicon 63, 120–128. https://doi.org/10.1016/j.toxicon.2012.11.022.
Opatova, V., Hamilton, C., Hedin, M., Montes de Oca, L., Kral, J., Bond, J.E., 2019. Phylogenetic systematics and evolution of the spider infraorder Mygalomorphae using genomic scale data. bioRxiv 531756. https://doi.org/10.1101/531756.
Ortiz, D., Francke, O.F., 2016. Two DNA barcodes and morphology for multi-method species delimitation in *Bonnetina* tarantulas (Araneae: Theraphosidae). Mol. Phylogenet. Evol. 101, 176–193. https://doi.org/10.1016/j.ympev.2016.05.003.
Ortiz, D., Francke, O.F., Bond, J.E., 2018. A tangle of forms and phylogeny: extensive morphological homoplasy and molecular clock heterogeneity in *Bonnetina* and related tarantulas. Mol. Phylogenet. Evol. 127, 55–73. https://doi.org/10.1016/j.ympev.2018.05.013.
Pérez-Miles, F., Lucas, S.M., Da Silva Jr., P.I., Bertani, R., 1996. Systematic revision and cladistic analysis of Theraphosinae (Araneae: Theraphosidae). Mygalomorph 1, 33–68.
Pérez-Miles, F., 1998. Notes on the systematics of the little known theraphosid spider *Hemirrhagus cervinus*, with a description of a new type of urticating hair. J. Arachnol. 26 (1), 120–123.
Pérez-Miles, F., 2002. The occurrence of abdominal urticating hairs during development in Theraphosinae (Araneae: Theraphosidae): phylogenetic implications. J. Arachnol. 30, 316–320.
Pérez-Miles, F., Guadanucci, J.P.L., Jurgilas, J.P., Becco, R., Perafán, C., 2017. Morphology and evolution of scopula, pseudoscopula and claw tufts in Mygalomorphae (Araneae). Zoomorphology 136 (4), 435–459.
Petersen, S.D., Mason, T., Akber, S., West, R., White, B., Wilson, P., 2007. Species identification of tarantulas using exuviae for international wildlife law enforcement. Conserv. Genet. 8 (2), 497–502.
Prasanth, M.T., Sunil Jose, K., 2014. A new species of the genus *Haploclastus* from Western Ghats, India (Araneae: Theraphosidae). Munis Entomol. Zool. 9 (1), 494–500.
Perafán, C., Galvis, W., Gutiérrez, M., Pérez-Miles, F., 2016. *Kankuamo*, a new theraphosid genus from Colombia (Araneae, Mygalomorphae), with a new type of urticating setae and divergent male genitalia. ZooKeys 601, 89–109. https://doi.org/10.3897/zookeys.601.7704.
Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9, e1000602.
Pollock, D.D., Zwickl, D.J., Mcguire, J.A., Hillis, D.M., 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol. 51, 664–671. https://doi.org/10.1080/10635150290102357.
Raven, R.J., 1985. The spider infraorder Mygalomorphae (Araneae): cladistics and systematics. Bull. Am. Mus. Nat. Hist. 182, 1–180.
Revell, L.J., 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x. v.0.6-60.
Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. 16 (6), 276277. https://doi.org/10.1016/S0168-9525(00)02024-2.
Rodríguez-Rios, L., Díaz-Peña, L.F., Lazcano-Pérez, F., Arreguín-Espinosa, R., Rojas-Molina, A., García-Arredondo, A., 2017. Hyaluronidase-like enzymes are a frequent component of venoms from theraphosid spiders. Toxicon 136, 34–43.
Samm, R., Schmidt, G., 2010. Psalmopoeinae subfamilia nov. - eine neue Unterfamilie der Theraphosidae (Araneae). Tarantulas World 142, 35–41.
Sanap, R.V., Mirza, Z.A., 2014. A new iridescent tarantula of the genus *Thrigmopoeus* Pocock, 1899 from Western Ghats, India. C.R. Biol. 337, 480–486. https://doi.org/10.1016/j.crvi.2014.06.003.
Sankaran, P.M., Sebastian, P.A., 2018. A new synonym in the subfamily Thrigmopoeinae Pocock, 1900 (Araneae, Theraphosidae). ZooKeys 749, 81–86. https://doi.org/10.3897/zookeys.749.23414.
Santana, R., Perez, D., Dobson, J., Panagides, N., Raven, R., Nouwens, A., Jones, A., King, G., Fry, B., 2017. Venom profiling of a population of the theraphosid spider *Phlogius crassipes* reveals continuous ontogenetic changes from juveniles through adulthood. Toxins 9, 116. https://doi.org/10.3390/toxins9040116.
Schmidt, G., 1995. Die Stellung der Gattung *Poecilotheria* im System. Arachnida 10, 1–2.
Schmidt, G., 2002. Gehören *Brachionopus* Pocock, 1897 und *Harpactirella* Purcell, 1902 zu den Theraphosiden? Arthropoda 10 (1), 12–17.
Schmidt, G., 2003. Die Vogelspinnen: eine weltweite Übersicht. Die Neue Brehm-Bücherei Bd. 641, Westarp Wissenschaften, Hohenwarsleben.
Simms, D., Cizdziel, P.E., Chomczynski, P., 1993. TRIzol: A new reagent for optimal single-step isolation of RNA. Focus 15, 532–535.
Simon, E., 1892. Histoire naturelle des araignées. Paris. vol. 1, pp. 1–256.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
Sunagar, K., Moran, Y., 2015. The rise and fall of an evolutionary innovation: contrasting strategies of venom evolution in ancient and young animals. PLoS Genet. 11. https://doi.org/10.1371/journal.pgen.1005596.
Teyssié, F., 2015. Tarantulas of the World – Theraphosidae. N.A.P. Editions, Verrieres-le-Buisson, France.
Turner, S.P., Longhorn, S.J., Hamilton, C.A., Gabriel, R., Pèrez-Miles, F., Vogler, A.P., 2018. Re-evaluating conservation priorities of New World tarantulas (Araneae: Theraphosidae) in a molecular framework indicates non-monophyly of the genera, Aphonopelma and Brachypelma. Syst. Biodivers. 16 (1), 89–107. https://doi.org/10.1080/14772000.2017.1346719.
Von Wirth, V., 2005. Vogelspinnen faszinierend & exotisch. Gräfe & Unzer, München.
Von Wirth, V., Striffler, B.F., 2005. Neue Erkenntnisse zur Vogelspinnen – Unterfamilie Ornithoctoninae, mit Beschreibung von Ornithoctonus aureotibialis sp. n. und Haplopelma longipes sp. n (Araneae, Theraphosidae). Arthropoda 13 (2), 2–27.
West, R.C., Marshall, S.D., Fukushima, C.S., Bertani, R., 2008. Review and cladistic analysis of the Neotropical tarantula genus *Ephebopus* Simon 1892 (Araneae: Theraphosidae) with notes on the Avicylariinae. Zootaxa 1849, 35–58.
West, R.C., Nunn, S.C., 2010. A taxonomic revision of the tarantula spider genus *Coremiocnemis* Simon 1892 (Araneae, Theraphosidae), with further notes on theSelenocosmiinae. Zootaxa 2443, 1–64.
World Spider Catalog, 2018. The World Spider Catalog, Natural History Museum Bern, online at http://wsc.nmbe.ch, version 20.0 (accessed on 01 April 2019).
Zilkens, K.M., Bastian, N., Löffler, K.U., Holz, F.G., 2012. "Line of defence" durchbrochen. Der Ophtalmologe 109 (8), 798–800.
Zwickl, D.J., Hillis, D.M., 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. 51, 588–598. https://doi.org/10.1080/10635150290102339.

Chapter III

Phylogeny-Guided Selection of Priority Groups for Venom Bioprospecting: Harvesting Toxin Sequences in Tarantulas as a Case Study

Tim Lüddecke, Andreas Vilcinskas & Sarah Lemke
Toxins (2019) 11(9): E488.


toxins
IMPACT FACTOR 3.895
Phylogeny-Guided Selection of Priority Groups for Venom Bioprospecting: Harvesting Toxin Sequences in Tarantulas as a Case Study
Volume 11 · Issue 9 | September 2019
MDPI
mdpi.com/journal/toxins
ISSN 2072-6651

toxins

MDPI

Communication

Phylogeny-Guided Selection of Priority Groups for Venom Bioprospecting: Harvesting Toxin Sequences in Tarantulas as a Case Study

Tim Lüddecke [1,*], Andreas Vilcinskas [1,2,3] and Sarah Lemke [2]

1 Department of Bioresources, Fraunhofer Institute for Molecular Biology and Applied Ecology, Winchesterstr. 2, 35394 Gießen, Germany
2 Institute for Insect Biotechnology, Justus-Liebig-University of Gießen, Heinrich-Buff-Ring 26-32, 35392 Gießen, Germany
3 LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Senckenberganlage 25, 60325 Frankfurt, Germany
* Correspondence: tim.lueddecke@outlook.com

Received: 19 July 2019; Accepted: 22 August 2019; Published: 25 August 2019

Abstract: Animal venoms are promising sources of novel drug leads, but their translational potential is hampered by the low success rate of earlier biodiscovery programs, in part reflecting the narrow selection of targets for investigation. To increase the number of lead candidates, here we discuss a phylogeny-guided approach for the rational selection of venomous taxa, using tarantulas (family Theraphosidae) as a case study. We found that previous biodiscovery programs have prioritized the three subfamilies Ornithoctoninae, Selenocosmiinae, and Theraphosinae, which provide almost all of the toxin sequences currently available in public databases. The remaining subfamilies are poorly represented, if at all. These overlooked subfamilies include several that form entire clades of the theraphosid life tree, such as the subfamilies Eumenophorinae, Harpactirinae, and Stromatopelminae, indicating that biodiversity space has not been covered effectively for venom biodiscovery in Theraphosidae. Focusing on these underrepresented taxa will increase the likelihood that promising candidates with novel structures and mechanisms of action can be identified in future bioprospecting programs.

Keywords: spiders; Theraphosidae; phylogenetics; venomics; bioprospecting; taxonomic bias

Key Contribution: Research on tarantula venom focuses only on a small fraction of the overall biodiversity. By combining phylogenetic and venomic data, we herein highlight the key tarantula lineages that need to be studied for optimizing bioprospecting within the family. This approach might contribute to streamline and economized venom biodiscovery from diverse animal groups in the future.

1. Introduction

Nature abounds with bioactive molecules synthesized by species that interact with each other, either competitively or cooperatively. These species have evolved the ability to produce chemical components that increase their fitness and favor their survival, for example by antagonizing competitors, predators, prey, and pathogens or by attracting symbionts and commensals. In the search for drugs against infectious and acquired diseases, humans have often turned to such natural bioactive molecules because they have acquired outstanding pharmacologies through millions of years of subsequent evolutive optimization towards potent bioactivity for their natural function. Accordingly, many of our current drugs are natural chemical entities or their derivatives [1].

Bioactive molecules are often sourced from microbes and plants, but attention has turned more recently to animal venoms. These evolved for hunting prey, defense against predation, and intraspecific competition [2]. The in-depth survey of venoms and their components has already led to the development of several important drugs, such as the analgesic ziconotide from the cone snail *Conus magus*, the antidiabetic exenatide, a synthetic derivative of exendin-4 from venom of the beaded-lizard *Heloderma suspectum*, and the antihypertensive captopril from the lancehead viper *Bothrops jararaca* [3]. That said, all venom-derived drugs have been isolated from a small and unrepresentative minority of venomous species, in particular from the largest and most dangerous taxa. However, venom evolved convergently in metazoans multiple times [2]. Several of the lineages that successfully evolved venom systems are additionally quite diverse on the species level with fish, insects, or arachnids being some examples. Interestingly, most representatives of these groups have not yet been studied for their venom in more detail. This means that the vast majority of venomous species remain virtually unexploited [4].

Spiders (order Araneae) provide an informative example of the problem discussed above. There are currently 48,249 recognized species of spiders, almost all of which produce venom [5], but the overwhelming majority of species that have been investigated in the search for venom-derived drugs are again either the larger or more dangerous members (Figure 1), especially from the genera *Atrax*, *Hadronyche*, *Missulena*, *Sicarius*, *Latrodectus*, *Hexophtalma*, and *Phoneutria*. This phenomenon has been coined as "taxonomic bias" and was subject to critical discussion in the recent past [6].

(A)

Family	Dangerous species	Larger species
Barychelidae		
Scytodidae		
Pisauridae		
Agelenidae		
Hexathelidae*	+	+
Ctenidae	+	+
Araneidae		(+)
Theraphosidae	(+)	+
Sicariidae	+	
Lycosidae		
Other	(+)	(+)

*) recently split into multiple families but not yet updated in databases

(B)

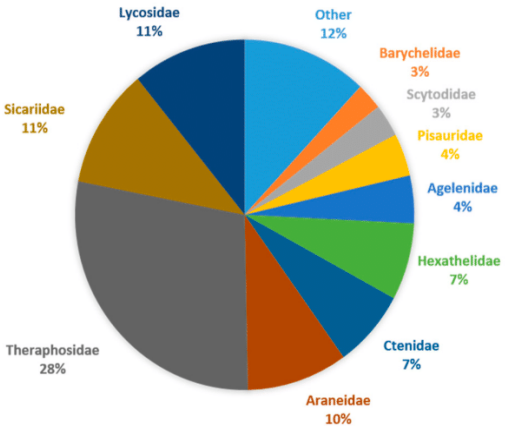


Figure 1. Taxonomic bias in spider venom research. Given in (**A**) are spider families which have been studied for their venom so far, together with an assignment of threat potential and size. Brackets indicate that only a small fraction of included species are either dangerous or large. Further note that the grouping "other" in reality represents the remaining 109 spider families, thus "other" contains the majority of spider biodiversity. (**B**) Visualizes the percentage of deposited toxin sequences per family. Current knowledge on spider venom is mostly based on data from those larger and more dangerous lineages and therefore is taxonomically biased. Data from [5,7,8]; see [6] for an in-depth discussion on taxonomic bias in spider venom research.

For biodiversity, spiders are the most successful lineage of venomous animals. The exploitation of all spiders could yield about 10 million different proteinaceous venom components, yet we have only just begun to tap this resource, with only 0.02% of such components identified thus far [8,9]. Spider venom is a promising target for bioprospecting because it is largely composed of small peptides that exhibit very specific and potent bioactivity against neuronal targets and further share an inhibitor cystine knot (ICK or knottin) motif. This structure confers a remarkable level of resistance against heat, chemicals, and proteases by structural stabilization via cysteine cross-linking, so drug candidates derived from these peptides are likely to be extremely stable in vivo [10]. Given that stability in vivo

and target specificity are major constraints for suitable drug candidates, facing the sheer diversity of peptides in spider venom, it is clear that it likely harbors several yet to be discovered biologics that will almost certainly serve as drug leads in the future.

2. Novel Strategies in Venom Bioprospecting

The translation of natural molecules to market-ready drugs is time consuming and expensive because most drug candidates fail, and the later this failure occurs in the development pipeline the greater the cost. The risk-averse pharmaceutical industry has largely abandoned such bioprospecting studies, and the burden now falls on research organizations, which work under tight financial constraints. This issue could potentially be addressed by optimizing bioprospecting strategies, for example by introducing a rational approach for the selection of taxa for investigation. Currently, bioprospecting is biased towards species that are considered to be medically significant or according to size, accessibility, and abundance [6]. This exclusion of large swathes of biodiversity stacks the odds against the discovery of promising new leads. It would be better to include neglected venomous lineages as priority groups based on rational factors, such as phylogeny, a spread-betting approach that would improve the likelihood of discovering promising candidates by attempting to cover 'biodiversity space'. Phylogenetic distance is, besides the species ecology, among the main drivers acting on venom evolution in terms of compound diversity, and it has been suggested before that phylogenetic distance should be acknowledged by scientists who aim to gain a holistic understanding on venom compositions within taxonomic groups [11]. The conceptual basis for using phylogenetic data as a roadmap in bioprospecting is that distantly related species are likely to evolve rather different venom profiles than closely related species and, therefore, are better candidates for yielding novel biologics [11].

Therefore, researchers should include diverse genetic lineages in their investigation to maximize the likelihood of finding such novel compounds [11]. This strategy is facilitated by the increasing availability of phylogenetic trees for animal lineages [12–15], providing quantitative data that will help with the selection of target species that represent the available biodiversity.

Bioprospecting from animal venoms was predominantly performed via pharmacological screenings, in which crude venoms or isolated toxins were subjected to specialized bioassays for each respective drug-target [16]. Although this approach was successfully applied in the past to identify promising drug leads from several species of reptiles, cone snails, and larger arachnids, among others [3], it relies on the ability to obtain meaningful amounts of venom from such organisms. Thus, the pharmacology driven strategy in venom bioprospecting is somewhat restricted to animals that are either large, easy to collect/breed, or otherwise deliver high venom yields. For the vast majority of venomous animals that are quite small, rare to find in nature, or difficult to sample for their crude venom, this approach is inappropriate [17]. However, based on the recent advances in mass spectrometry and next generation sequencing, it became possible to study even those critical taxa by means of the "methodological triad" of venomics (proteomics, transcriptomics, and genomics) [4,17,18]. The increased sensitivity and depth of instruments involved in such studies combined with significant cost reductions over the last decades allows us to identify a plethora of toxin sequences from such groups, including geophilomorph centipedes, remipede crustaceans, and pseudoscorpions [19–23]. In order to exploit these identified sequences for pharmaceutical applications and feed them into the value chain, they need to be synthesized or recombinantly expressed prior to extensive bioactivity tests. Facing the fact that this approach has its own drawbacks and disadvantages, it is, however, currently the method-to-choose for the study of venoms from taxa where pharmacological driven surveys fail, although it is important to note that none of these approaches will probably be able to study all venoms from all taxa alone, and rather a strategy that applies both strategies in tandem might be the most fruitful.

As a direct consequence of the ~omics-based approach, which has extensively been used for the study of spider venoms in the recent past [24–26], several sequence databases were erected to manage

the bulk of sequence data that is created by such ~omics-based studies, with the Arachnoserver and Venomzone being two examples.

3. Phylogeny-Guided Selection of Priority Groups for Venom Bioprospecting: Tarantulas as a Case Study

Among spiders, the family Theraphosidae (commonly known as tarantulas) has been the subject of recent detailed phylogenetic and phylogenomic studies, revealing for the first time the deep evolutionary relationships among 12 of the 14 currently accepted subfamilies within theraphosids [15,27–29]. Tarantulas are particularly suitable as a case study for rational selection in the context of venom bioprospecting because many tarantula-derived toxin sequences, obtained by the application of the previously explained ~omics-based venom bioprospecting strategy, are already available in protein databases [8,19]. This unique framework means that sequence and phylogenetic data can be combined to develop, test, and validate an optimized sampling strategy based on phylogenetic distance.

Accordingly, we sourced the available data from two manually curated venom databases, specifically Arachnoserver (AS) and Venomzone (VZ) for tarantula toxin sequences, as well as the World Spider Catalog (WSC) and Tarantupedia taxonomic databases [5,8,30,31]. We also inferred phylogenetic relationships based on recently published studies of tarantula evolution [15,29]. These data were used to identify the subfamilies and genera of the family Theraphosidae that are currently underrepresented in terms of the quantity of deposited venom peptide sequences, and which should therefore be targeted in future bioprospecting studies.

At the time of writing, the two venom databases were not identical in terms of the number of deposited toxin sequences (450 sequences in AS, 532 in VZ), probably reflecting differences in the stringency of criteria for data deposition and topicality. However, the databases followed similar trends in terms of the distribution of toxin sequences among the 14 recognized Theraphosidae subfamilies (Table 1). Most sequences represented subfamily Ornithoctoninae (247 sequences in AS, 339 in VZ), followed by Selenocosmiinae (76 sequences in AS, 120 in VZ) and Theraphosinae (95 sequences in AS, 51 in VZ). The remaining subfamilies were scarcely represented (e.g., Eumenophorinae and Psalmopoeinae), or no toxin sequences were present at all (e.g., Poecilotheriinae and Thrigmopoeinae). This shows that tarantula research is strongly biased towards the Ornithoctoninae, Selenocosmiinae, and Theraphosinae, whereas the other subfamilies are left behind as a biological "black box". The reasons for such a taxonomically-biased picture may be either the size of the respective spider, which influences the venom sampling, the availability, the ability to securely identify the spider, or a combination thereof. For an in-depth discussion about taxonomic bias in spider venom research and for solutions to the problem see [6].

Table 1. Species richness in the family Theraphosidae and the number of toxin sequences deposited in the databases Arachnoserver (AS) and Venomzone (VZ) plus species diversity from World Spider Catalog (WSC) for each subfamily. Although the number of deposited sequences differs between the databases, they follow the same trend with the majority of sequences representing subfamilies Ornithoctoninae, Selenocosmiinae, and Theraphosinae. The remaining subfamilies are poorly represented, if at all. Subfamilies with unclear phylogenetic placement are marked with astersisks.

Subfamily	Number of Species WSC	Number of Toxins AS	Number of Toxins VZ
Acanthopelminae [1]	2	0	0
Avicylariinae	31	0	1
Eumenophorinae	62	10	1
Harpactirinae	62	8	7
Ischnocolinae [2]	85	3	0
Ornithoctoninae	27	247	339

Table 1. *Cont.*

Subfamily	Number of Species WSC	Number of Toxins AS	Number of Toxins VZ
Poecilotheriinae [3]	14	0	0
Psalmopoeinae	27	6	7
Schismatothelinae [2]	21	0	0
Selenocosmiinae	114	76	120
Selenogyrinae [1]	10	0	0
Stromatopelminae	10	5	6
Theraphosinae	526	95	51
Thrigmopoeinae	9	0	0

[1] subfamily with unclear phylogenetic placement., [2] paraphyletic, [3] toxins described but not yet added to databases.

A priori, one could hypothesize that the predominance of toxin sequences representing particular subfamilies may reflect the species richness within these taxonomic groups, but this turns out not to be the case. For example, the subfamily Ornithoctoninae was the most abundantly represented group in the venom databases, but accounts for only ~3% of tarantulas, whereas the subfamily Theraphosinae ranked third for venom but accounts for more than 50% of all known species in tarantulas [31]. Having discarded the hypothesis that the abundance of toxin data correlates with species richness, we considered the possibility that the most "dangerous" species have been prioritized for investigation. Although tarantulas are generally not considered dangerous to humans, anecdotal reports suggest that species from Asia and Africa deliver more intense bites and cause more painful envenoming effects compared to species from the Americas [32]. The subfamilies Ornithoctoninae and Selenocosmiinae are African and Asian groups, providing some evidence to support the focus on "dangerous" species, but if this is the case, it remains unclear why other African and Asian subfamilies, such as Thrigmopoeinae, Stromatopelminae, Eumenophorinae. and Harpactirinae, have been largely overlooked. Most pertinently, the subfamily Poecilotheriinae is the only group of tarantulas considered medically significant for humans [7], and yet this is currently one of the least represented groups in terms of toxin sequences deposited to the herein analyzed databases (Table 1) (representatives of over- and understudied tarantula groups are depicted in Figure 2). However, we are aware of the fact that especially the situation of Poecilotheriinae and the described toxin sequences from this subfamily is a difficult case, which will be necessary to be evaluated again soon: A bioprospecting study from 2017 had *Poecilotheria formosa*, a representative of Poecilotheriinae, among the studied taxa and described over 100 toxin sequences from its venom, using a proteotranscriptomic approach [33]. Unfortunately, these toxins were not yet added to any of the herein utilized databases and were omitted by us for consistency reasons.

Recent evolutionary analysis on 12 out of the total 14 subfamilies within Theraphosidae revealed that the so far phylogenetically determined subfamilies form five major clades, representing distinct genetic lineages (Figure 3). If the distribution of toxin data for each subfamily is mapped onto this phylogeny, it becomes clear that several of these major clades have been neglected in previous studies: This is true for the clade formed by subfamily Eumenophorinae and parts of the paraphyletic Ischnocolinae, the clade formed by the African subfamilies Harpactirinae and Stromatopelminae, and the non-theraphosine members of the clade, comprising American subfamilies Avicullariinae, Schismatothelinae, and Psalmopoeinae, plus some species of Ischnocolinae. Given the clear bias in the coverage of toxin sequences and the plethora of toxins anticipated in these three underrepresented clades, plus their evolutionary distance from other tarantula subfamilies, we propose that the members of these clades should be prioritized in future bioprospecting studies. The subfamilies Poecilotheriinae and Thrigmopoeinae are likewise underrepresented, but given their closer relationship to Ornithoctoninae and Selenocosmiinae, they are probably less likely to harbor really novel toxins.

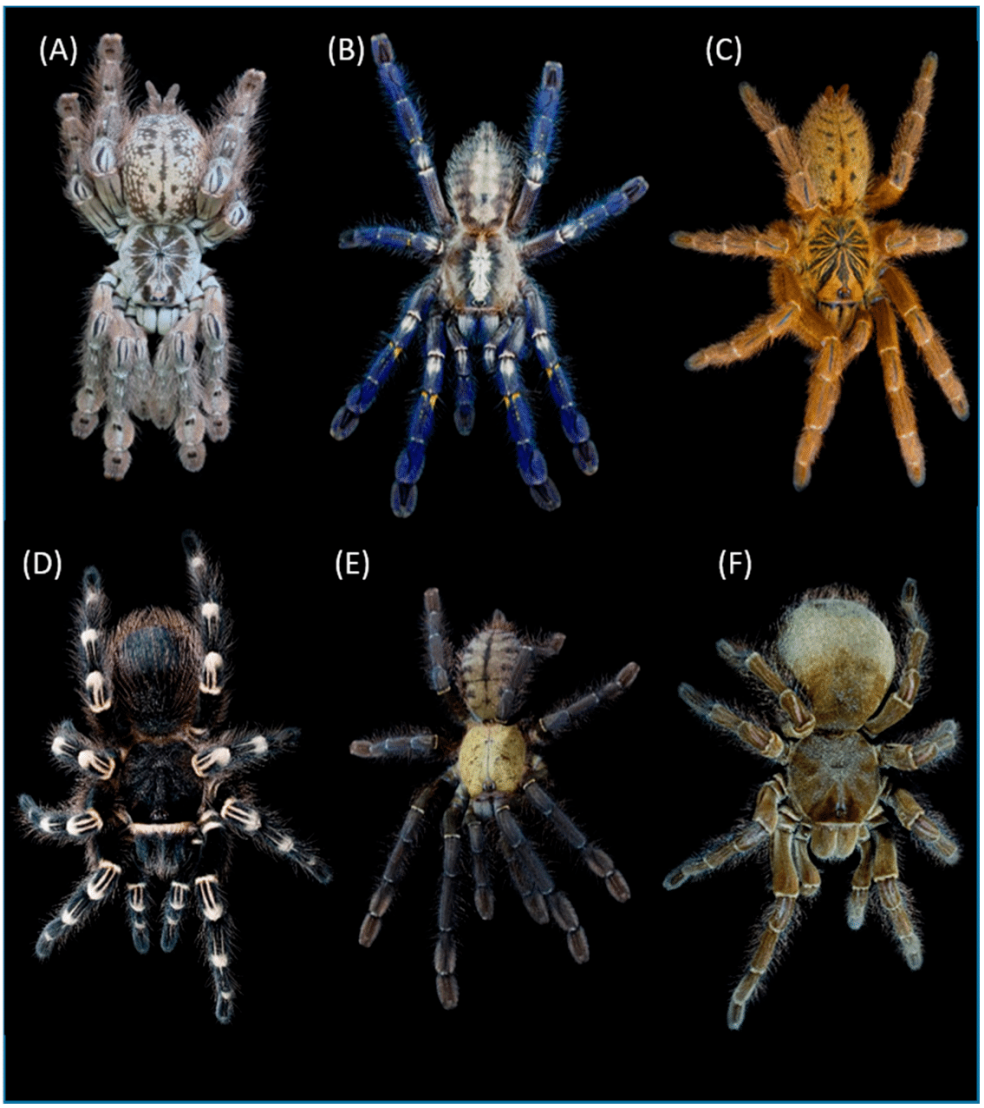


Figure 2. Venom-wise understudied and over-studied members of Theraphosidae. Several lineages within the theraphosid radiation were mostly neglected in the past. The upper row includes representatives of such lineages, reflecting some of the herein evaluated priority groups: (**A**) *Heteroscodra maculata* (Stromatopelminae), (**B**) *Poecilotheria metallica* (Poecilotheriinae), and (**C**) *Pterinochilus murinus* (Harpactirinae). On the other hand, some lineages are responsible for the bulk of knowledge that is available on tarantula venom: (**D**) *Acanthoscurria geniculata* (Theraphosinae), (**E**) *Cyriopagopus schioedtei* (Ornithotoctoninae), and (**F**) *Theraphosa stirmi* (again, Theraphosinae). Photography is courtesy of Bastian Rast, Switzerland.

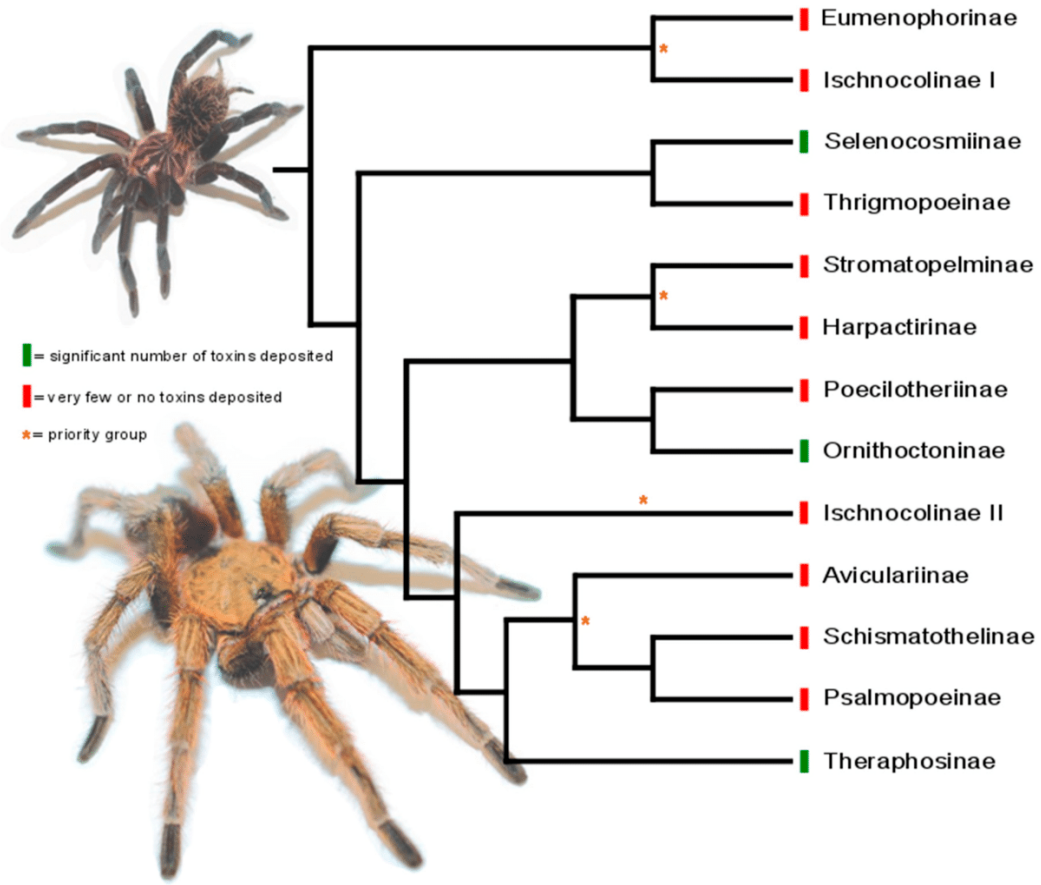


Figure 3. Priority groups for venom-based biodiscovery programs assigned to the tarantula tree of life, encompassing the 12 so far phylogenetically assessed subfamilies within Theraphosidae (cladogram based on [29]). The subfamilies Ornithoctoninae, Selenocosmiinae, and Theraphosinae dominate in terms of deposited toxin sequences (green), whereas little or no information is available for the other subfamilies (red). Entire theraphosid clades, representing major radiations within the family, are almost completely unrepresented and are therefore considered priority groups (red asterisk), including the clade of Eumenophorinae and Ischnocolinae, the African clade of Harpactirinae and Stromatopelminae, and the non-theraphosine new-world tarantulas. Note that the subfamily Ischnocolinae is paraphyletic and therefore appears twice in the phylogeny.

4. Concluding Remarks

The advent of venom-focused genomics, transcriptomics, and proteomics has provided the means to study venoms in a high-throughput and cost-effective manner [34,35]. We are confident that the application of venomics to the priority groups we have identified will contribute to the understanding of theraphosid venoms and will help to accelerate venom-based biodiscovery programs focusing on these intriguing and charismatic spiders. Apart from the family Theraphosidae, the phylogeny-guided bioprospecting approach herein discussed might further accelerate biodiscovery from very diverse venomous lineages in general. For example, across the spider tree of life alone, multiple clades have been phylogenetically resolved recently but remain either completely unstudied for their venom so far or all available information on venoms from these clades is derived only from one or two species [8,36–40], thus reflecting a very narrow fraction of these lineages. Gaining a thorough understanding of the venom composition for those groups is a major challenge if a complete idea upon arachnid venoms wants to be achieved, may it be for bioprospecting or for basic research on the biology of spiders itself.

Independent of studied taxa or research aims, the study of venoms from understudied groups needs to be performed more rapidly. On one hand, this is important for the aforementioned streamlining and economization of bioprospecting programs, but on the other hand, this is of pivotal importance in order to create something akin to a "library of bioresources" from venomous animals. We are currently living in the sixth mass extinction event, and the dramatic loss of global biodiversity likely affects many toxin-producing species, as powerfully highlighted by the dramatic biodiversity loss in amphibians [41–46]. Consequently, it is a real concern that valuable bioresources which could be found in such animals are getting lost forever, in case the respective species goes extinct. Therefore, it is a major task for the toxinological community to enhance the studies of venoms and create such a "library of bioresources" in order to save the genetic information of venom proteins for the future.

The comprehensive study of venom as a bioresource suffers from a variety of problems that affect its success rate. For example, many of the unstudied venomous species are rather small, and it is notoriously difficult to obtain meaningful amounts of venom for bioactivity screens or proteomic studies from these. Furthermore, some of these species are difficult to study because of their secretive lifestyle or their natural habitat being cumbersome to explore [4,35]. Additionally, political restrictions, such as those imposed by the Nagoya Protocol, are major impediments that somewhat negatively affect venom bioprospecting (see [47] for a discussion on the topic using microbiology as an example). Beyond these major problems, which certainly need to be solved in the future, the rational selection of taxa by means of phylogenetic distance could drastically improve any research efforts in this direction and contribute to the achievement of such a goal.

Author Contributions: T.L., S.L. and A.V. together designed the study. T.L. performed data collection and analysis. T.L. wrote the manuscript with substantial input of A.V. and S.L.

Funding: This research received no external funding.

Acknowledgments: The authors wish to thank the German Arachnological Society (Deutsche Arachnologische Gesellschaft–DeArGe e.V) and its members for supporting the study logistically and intellectually. Bastian Rast gratefully provided picture material of tarantulas that was used to assemble Figure 3. We thank Richard M. Twyman for his comments on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cragg, G.M.; Newman, D.J. Natural products: A continuing source of novel drug leads. *Biochim. Biophys. Acta* **2013**, *1830*, 3670–3695. [CrossRef] [PubMed]
2. Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.; King, G.F.; Nevalainen, T.J.; Norman, J.A.; Lewis, R.J.; Norton, R.S.; et al. The toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* **2009**, *10*, 483–511. [CrossRef] [PubMed]
3. Holford, M.; Daly, M.; King, G.F.; Norton, R.S. Venoms to the rescue. *Science* **2018**, *6405*, 842–844. [CrossRef] [PubMed]
4. Von Reumont, B.M.; Campbell, L.I.; Jenner, R.A. Quo vadis venomics? A roadmap to neglected venomous invertebrates. *Toxins* **2014**, *6*, 3488–3551. [CrossRef] [PubMed]
5. World Spider Catalog V.20.0. Natural History Museum Bern. 2019. Available online: www.wsc.nmbe.ch (accessed on 25 June 2019).
6. Herzig, V.; King, G.F.; Undheim, E.A.B. Can we resolve the taxonomic bias in spider venom research? *Toxicon X* **2019**, *1*, 100005. [CrossRef]
7. Hauke, T.J.; Herzig, V. Dangerous arachnids- Fake news or reality? *Toxicon* **2017**, *138*, 173–183. [CrossRef] [PubMed]
8. Pineda, S.S.; Chaumeil, P.A.; Kunert, A.; Kaas, Q.; Thang, M.W.C.; Le, L.; Nuhn, M.; Herzig, V.; Saez, N.J.; Cristofori-Armstrong, B.; et al. ArachnoServer 3.0: An online resource for automated discovery, analysis and annotation of spider toxins. *Bioinformatics* **2018**, *34*, 1074–1076. [CrossRef] [PubMed]
9. Herzig, V.; Wood, D.L.A.; Newell, F.; Chaumeil, P.; Kaas, Q.; Binford, G.J.; Nicholson, G.M.; Gorse, D.; King, G.F. Arachnoserver 2.0, an updated online resource for spider toxin sequences and structures. *Nucl. Acids Res.* **2011**, *39*, D653–D657. [CrossRef] [PubMed]

10. Pineda, S.S.; Undheim, E.A.; Rupasinghe, B.E.; Ikonomopoulou, M.P.; King, G.F. Spider venomics: Implications for drug discovery. *Future Med. Chem.* **2014**, *6*, 1699–1714. [CrossRef]
11. Fry, B.G.; Koludarov, I.; Jackson, T.N.W.; Holford, M.; Terrat, Y.; Casewell, N.R.; Undheim, E.A.B.; Vetter, I.; Ali, S.A.; Dolyce, H.W.; et al. Seeing the Woods for the Trees: Understanding Venom Evolution as a Guide for Biodiscovery. In *Venoms to Drugs: Venom as a Source for the Development of Human Therapeutics*; RSC Drug Discovery Series No.42; King, G.F., Ed.; The Royal Society of Chemistry: Cambridge, UK, 2015; pp. 1–36.
12. Vidal, N.; Hedges, S.B. The phylogeny of squamate reptiles (lizards, snakes and amphisbaenians) inferred from nine nuclear protein-coding genes. *Comptes Rendus Biol.* **2005**, *328*, 1000–1008. [CrossRef]
13. Bond, J.E.; Garrison, N.L.; Hamilton, C.A.; Godwin, R.L.; Hedin, M.; Agnarsson, I. Phylogenomics resolve a spider backbone phylogeny and rejects a prevailing paradigm for orb web evolution. *Curr. Biol.* **2014**, *24*, 1765–1771. [CrossRef] [PubMed]
14. Vences, M.; Sanchez, E.; Hauswaldt, J.S.; Eikelmann, D.; Rodriguez, A.; Carranza, S.; Donaire, D.; Gehara, M.; Helfer, V.; Lötters, S.; et al. Nuclear and mitochondrial multilocus phylogeny and survey of alkaloid content in true salamanders of the genus *Salamandra* (Salamandridae). *Mol. Phylogenet. Evol.* **2014**, *73*, 208–216. [CrossRef] [PubMed]
15. Lüddecke, T.; Krehenwinkel, H.; Canning, G.; Glaw, F.; Longhorn, S.J.; Tänzler, R.; Wendt, I.; Vences, M. Discovering the silk road: Nuclear and mitochondrial data resolve the phylogenetic relationships among theraphosid spider subfamilies. *Mol. Phylogenet. Evol.* **2018**, *119*, 63–70. [CrossRef] [PubMed]
16. Pennington, M.W.; Czerwinski, A.; Norton, R.S. Peptide therapeutics from venom: Current status and potential. *Bioorg. Med. Chem.* **2018**, *26*, 2738–2758. [CrossRef] [PubMed]
17. Von Reumont, B.M. Studying Smaller and Neglected Organisms in Modern Evolutionary Venomics Implementing RNASeq (Transcriptomics)—A Critical Guide. *Toxins* **2018**, *10*, 292. [CrossRef] [PubMed]
18. Sunagar, K.; Morgenstern, D.; Reitzel, A.M.; Moran, Y. Ecological venomics: How genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom. *J. Proteom.* **2016**, *135*, 62–72. [CrossRef]
19. Jenner, R.A.; von Reumont, B.M.; Campbell, L.; Undheim, E.A.B. Parallel evolution of complex centipede venoms revealed by comparative proteotranscriptomic analyses. *Mol. Biol. Evol.* **2019**. [CrossRef] [PubMed]
20. Von Reumont, B.M.; Blanke, A.; Richter, S.; Alvarez, F.; Bleidorn, C.; Jenner, R.A. The first venomous crustacean revealed by transcriptomics and functional morphology: Remipede venom glands express a unique toxin cocktail dominated by enzymes and a neurotoxin. *Mol. Biol. Evol.* **2014**, *31*, 48–58. [CrossRef]
21. Von Reumont, B.M.; Undheim, E.A.B.; Jauss, R.T.; Jenner, R.A. Venomics of Remipede Crustaceans Reveals Novel Peptide Diversity and Illuminates the Venom's Biological Role. *Toxins* **2017**, *9*, 234. [CrossRef]
22. Santibáñez-López, C.E.; Ontano, A.Z.; Harvey, M.S.; Sharma, P.P. Transcriptomic Analysis of Pseudoscorpion Venom Reveals a Unique Cocktail Dominated by Enzymes and Protease Inhibitors. *Toxins* **2018**, *10*, 207. [CrossRef]
23. Krämer, J.; Pohl, H.; Predel, R. Venom collection and analysis in the pseudoscorpion *Chelifer cancroides* (Pseudoscorpiones: Cheliferidae). *Toxicon* **2019**, *162*, 15–23. [CrossRef] [PubMed]
24. Paiva, A.L.B.; Mudadu, M.A.; Pereira, E.H.T.; Marri, C.A.; Guerra-Duarte, C.; Diniz, M.R.V. Transcriptome analysis of the spider *Phoneutria reidyi* venom gland reveals novel venom components for the genus *Phoneutria*. *Toxicon* **2019**, *163*, 59–69. [CrossRef] [PubMed]
25. Kuhn-Nentwig, L.; Langenegger, N.; Heller, M.; Koua, D.; Nentwig, W. The Dual Prey-Inactivation Strategy of Spiders-In-Depth Venomic Analysis of *Cupiennius Salei*. *Toxins* **2019**, *11*, 167. [CrossRef] [PubMed]
26. Hu, Z.; Chen, B.; Xiao, Z.; Zhou, X.; Liu, Z. Transcriptomic Analysis of the Spider venom Gland Reveals Venom Diversity and Species Consanguinity. *Toxins* **2019**, *11*, 68. [CrossRef] [PubMed]
27. Ortiz, D.; Francke, O.F.; Bond, J.E. A tangle of forms and phylogeny: Extensive morphological homoplasy and molecular clock heterogeneity in *Bonnetina* and related tarantulas. *Mol. Phylogenet. Evol.* **2018**, *127*, 55–73. [CrossRef]
28. Hüsser, M. A first phylogenetic analysis reveals a new arboreal tarantula genus from South America with description of a new species and two new species of *Tapinauchenius* Ausserer, 1871 (Araneae, Mygalomorphae, Theraphosidae). *Zookeys* **2018**, *784*, 59–93. [CrossRef]

29. Foley, S.; Lüddecke, T.; Dong-Chiang, C.; Krehenwinkel, H.; Künzel, S.; Longhorn, S.J.; Wendt, I.; von Wirth, V.; Tänzler, R.; Vences, M.; et al. Tarantula phylogenomics: A robust phylogeny of multiple tarantula lineages inferred from transcriptome data sheds light on the prickly issue of urticating setae evolution. *Mol. Phylogenet. Evol.* **2019**, *140*. [CrossRef]
30. Venomzone. SIB Swiss Institute of Bioinformatics. Available online: https://www.venomzone.expasy.org (accessed on 25 June 2019).
31. Tarantupedia: An online Taxonomic Database for the World Largest Spiders. Available online: www.tarantupedia.com (accessed on 25 June 2019).
32. Escoubas, P.; Rash, L. Tarantulas: Eight-legged pharmacists and combinatorial chemists. *Toxicon* **2004**, *43*, 555–574. [CrossRef]
33. Oldrati, V.; Koua, D.; Allard, P.M.; Hulo, N.; Arrell, M.; Nentwig, W.; Lisacek, F.; Wolfender, J.L.; Kuhn-Nentwig, L.; Stöcklin, R. Peptidomic and transcriptomic profiling of four distinct spider venoms. *PLoS ONE* **2017**, *12*, e0172966. [CrossRef]
34. Oldrati, V.; Arrell, M.; Violette, A.; Perret, F.; Sprüngli, X.; Wolfender, J.L.; Stöcklin, R. Advances in Venomics. *Mol. Biosyst.* **2016**, *12*, 3530–3543. [CrossRef]
35. Drukewitz, S.H.; von Reumont, B.M. The significance of comparative genomics in modern evolutionary venomics. *Front. Ecol. Evol.* **2019**. [CrossRef]
36. Piacentini, L.N.; Ramirez, M.J. Hunting the wolf: A molecular phylogeny of the wolf spiders (Araneae, Lycosidae). *Mol. Phylogenet. Evol.* **2019**, *136*, 227–240. [CrossRef] [PubMed]
37. Huber, B.A.; Eberle, J.; Dimitrov, D. The phylogeny of pholcid spiders: A critical evaluation of relationships suggested by molecular data (Araneae, Pholcidae). *Zookeys* **2018**, *789*, 51–101. [CrossRef] [PubMed]
38. Godwin, R.L.; Opatova, V.; Garrison, N.L.; Hamilton, C.A.; Bond, J.E. Phylogeny of a cosmopolitan family of morphologically conserved trapdoor spiders (Mygalomorphae, Ctenizidae) using Anchored Hybrid Enrichment, with a description of the family, Halonoproctidae Pocock 1901. *Mol. Phylogenet. Evol.* **2018**, *126*, 303–313. [CrossRef] [PubMed]
39. Kallal, R.J.; Fernandez, R.; Giribet, G.; Hormiga, G. A phylotranscriptomic backbone of the orb-weaving spider family Araneidae (Arachnida, Araneae) supported by multiple methodological approaches. *Mol. Phylogenet. Evol.* **2018**, *126*, 129–140. [CrossRef] [PubMed]
40. Cheng, D.Q.; Piel, W.H. The origins of Psechridae: Web-building lycosoid spiders. *Mol. Phylogenet. Evol.* **2018**, *125*, 213–219. [CrossRef] [PubMed]
41. Barnosky, A.D.; Matzke, N.; Tomiya, S.; Wogan, G.O.; Swartz, B.; Quental, T.B.; Marshall, C.; McGuire, J.L.; Lindsey, E.L.; Maguire, K.C.; et al. Has the Earth's sixth mass extinction already arrived? *Nature* **2011**, *471*, 51–57. [CrossRef] [PubMed]
42. Dirzo, R.; Young, H.S.; Galetti, M.; Ceballos, G.; Isaac, N.J.; Collen, B. Defaunation in the Anthropocene. *Science* **2014**, *345*, 401–406. [CrossRef] [PubMed]
43. Régnier, C.; Achaz, G.; Lambert, A.; Cowie, R.H.; Bouchet, P.; Fontaine, B. Mass extinction in poorly known taxa. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7761–7766.
44. Wake, D.B.; Koo, M.S. Amphibians. *Curr. Biol.* **2018**, *28*, R1237–R1241. [CrossRef]
45. Collins, J.P. Amphibian decline and extinction: What we know and what we need to learn. *Dis. Aquat. Organ.* **2010**, *92*, 93–99. [CrossRef] [PubMed]
46. Wake, D.B.; Vredenburg, V.T. Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11466–11473. [CrossRef] [PubMed]
47. Overmann, J.; Hartmann Scholz, A. Microbiological Research Under the Nagoya Protocol: Facts and Fiction. *TIMI* **2016**, *1385*. [CrossRef] [PubMed]

Chapter IV

Microscopic Analyses of the Wasp Spider (*Argiope bruennichi*) Venom System – Insights into the Architecture of an Arachnid Armory

Henrike Schmidtberg, Sarah Lemke, Björn M. von Reumont, Andreas Vilcinskas & Tim Lüddecke

Microscopic Analyses of the Wasp Spider (*Argiope bruennichi*) Venom System – Insights into the Architecture of an Arachnid Armory

Henrike Schmidberg [1], Björn M. von Reumont [1,2], Sarah Lemke [1], Andreas Vilcinskas [1, 2,3] and Tim Lüddecke [2,3] *

[1] Justus Liebig University of Gießen, Institute for Insect Biotechnology, Heinrich-Buff-Ring 26-32, 35392 Gießen
[2] LOEWE Centre for Translational Biodiversity Genomics (TBG), Senckenberganlage 25, 60325 Frankfurt, Germany
[3] Fraunhofer Institute for Molecular Biology and Applied Ecology, Department of Bioresources, Ohlebergsweg 12, 35392 Gießen, Germany

*Corresponding author: tim.lueddecke@ime.fraunhofer.de

Simple Summary: Despite that spiders are the most successful group of venomous animals, the venom systems of most taxa remain unexplored for their biochemistry and morphological organization. Here, we analyzed the morphology of the venom system of the wasp spider *Argiope bruennichi*, a member of the Araneidae family. Its architecture largely resembles venom systems of other spider lineages. Wasp spiders pursue an uncommon trophic strategy as their hunting is almost exclusively based on the silk apparatus and previous research found, that they evolved an arachno-atypical venom as a likely consequence of this adaption. Interestingly, this unusual character of the venom is not reflected in the wasp spider's venom system morphology. Also, the herein performed analysis revealed that the venom duct, that connects the spiders fang with the venom gland, is composed of four differentiated subunits and thus is more complex than previously acknowledged. These results suggest that the venom duct in spiders may be involved in biosynthesis of venom components as known from pitvipers and cone snails.

Abstract: Spiders are one of the most successful groups of venomous animals, but only a fraction has been examined for the structural basis of their venom system. Among the neglected spiders, Araneidae as one of the most diverse families is of particular interest. For *Argiope bruennichi*, known as the wasp spider, it was recently shown that it features an arachno-atypical venom mostly composed of CAP proteins. Thus, we studied the morphology its venom apparatus by thorough microscopic investigation. Further, we explored if the arachno-atypical nature of this spiders venom is also reflected in an unusual architecture of its venom system. We find, that the venom system of wasp spiders morphologically largely confers to those found in already studied species. A comparison with other spider venom systems across different families, ecological niches and taxonomic infraorders revealed that large swathes of the spider venom system architectures are conserved between studied taxa and little alteration is present. However, a detailed analysis of the wasp spiders venom duct revealed, that this element of the venom system is composed of four structurally different subunits. As similar substructures in pitvipers and cone snails were previously found to be involved in toxin biosynthesis, we propose

that parts of the venom duct may likewise shape spider venom profiles and represent previously underestimated components for the venom system.

Keywords: Araneidae; *Argiope bruennichi*; Venom glands; Morphology

1. Introduction

Across the metazoan tree of life, venoms convergently evolved more than 100 times in all major animal lineages [1, 2]. Beyond the three principal biological functions of venom (predation, defense and competition) at least eleven different ecological functions for venoms can be described [1,3]. The functional diversity of venoms and its application ways is inherently linked to the organization of the venom delivery apparatus, hence its functional morphology may constrain on venom usage. For example the ability to modulate venom compositions and to shape venom profiles can be influenced by the possession of a centralized or decentralized venom system, or the degree of gland-complexity [1]. A thorough description of the morphological organization of venom delivery systems in venomous animals is one of the prerequisite to understand their venom biology.

Except of Uloboridae, all spiders utilize a venom system. This reasons that spiders are commonly acknowledged as the worlds most successful group of venomous animals comprising 48,464 extant species from 4157 genera and 120 families [4, 5]. Spider venoms are chemically complex entities and composed of small molecules, proteins and peptides of which numerous representatives were isolated and pharmacologically examined in the past [5–10]. The current knowledge on spider venoms, however, comes from only a small non-representative fraction of species, while the vast majority of spiders remain unstudied [11, 12]. Surprisingly, even less is known about the morphological organization of most spider venom delivery systems. Although some detailed studies exist, these previous works selected again an unrepresentative minority of taxa that do not reflect overall spider diversity. In particular, the venom systems of potentially dangerous species such as black widows (genus *Latrodectus*), recluse spiders (genus *Loxosceles*) and wandering spiders (genus *Phoneutria*) were subject to closer investigations [13–18]. Recently, further taxa were studied such as lynx spiders (Oxyopidae), wolf spiders (genus *Lycosa*), furrow orb weavers (genus *Larinioides*), tarantulas (genus *Vitalius*), and tube web spiders (genus *Segestria*) [19–24]. However, for a more comprehensive understanding of the morphological organization of venom delivery systems in spiders, it is indispensible to study further taxa.

With 3,058 commonly accepted species, the family Araneidae, also known as orb-weavers, represents the third most speciose spider family [4]. Araneidae are general predators on insects that utilize complex and conspicuous orb-shaped foraging webs for hunting [25]. These provided them with a widespread recognition in the general public and drew a significant research interest to uncover their natural history and ecology [26–29]. In particular, *Argiope bruennichi*, which is often referred to as the wasp spider due to its characteristic black-yellow coloration that resembles some hymenopterans, became a frequently studied model taxon in this context (Figure 1). As it displays an array of unique ethological characters, *A. bruennichi* was exhaustively studied for its mating, silk-spinning and predatory behavior [30–36]. Moreover, the genetic basis of its recent, outstandingly

successful, pole-ward range expansion as well as its microbiome was of particular interest [37–40].

Apart from these studies, the venom of *A. bruennichi* has recently been investigated by proteo-transcriptomics guided venomics [41]. In this study it was revealed, that wasp spider venom appears as arachno-atypical: Spider venoms are generally thought to be mostly composed of small neurotoxic peptides with an inhibitor cysteine knot (ICK) motif and to be highly complex cocktails [41]. Large proteins, on the other hand, were considered minor components of spider venoms. The venom of wasp spiders, however, is primarily composed of CAP proteins instead of ICKs and contains predominantly large proteins [41]. Most interestingly, the wasp spider venom abounds with an astonishing simplicity as it is dominated by only a single protein class [41]. It has been proposed that this remarkable venom evolved as an outcome of an economic competition between the venom system and the silk apparatus during hunting [41]. Contrary to most other spiders that rely on venomous bite as a first means to overpower their prey, wasp spiders instead favor their silk apparatus to immobilize prey [32]. Here, the victim is spun in and only bitten when fully trapped in silk, thereby providing the wasp spider an increased success rate when capturing well defended prey [32].

Figure 1. The wasp spider *Argiope bruennichi* displays a characteristic black-yellow banding pattern.

Facing the biological importance of Araneidae as one of the most diverse and derived spider families [4], it is rather unfortunate that their venom systems have been largely neglected so far. The venom of *Araneus ventricosus* is, besides *A. bruennichi*, the only araneid venom that has been studied in detail [42]. The morphological organization of venom systems in Araneidae, on the other hand, has not been subject to closer investigation.

This work fills the gap in lack of knowledge on venom system morphology in Araneidae by studying that of *A. bruennichi* via microscopic techniques. In several cases throughout the animal kingdom, venom system morphology impose evolutionary frameworks upon the venom itself – thus both systems are evolutionary tightly connected [1, 43]. In perspective to the arachno-atypical nature of wasp spider venom, we asked if the morphological organization of its venom system likewise displays unusual adaptions. Our research supplies data on the morphology behind spider venom systems and provides a framework

on which subsequent exploratory studies of the remaining majority of spider venom systems can be informed upon.

2. Material and Methods

2.1. Collection of spiders

Adult female wasp spiders were collected in August 2019 on a meadow in Gießen, Germany. The collection site has been used previously for the collection of wasp spiders for venomic analysis. The collected animals were kept in plastic enclosures (20 x 20 x 20 cm) until further processing. Prior to any analyses, the respective wasp spiders were anesthetized and killed with CO_2.

2.2. Preparation of histological sections

For structural analyses, the prosomata of collected wasp spiders with both chelicera and their associated venom glands were removed from the opisthosoma. The tissues were submerged in ice-cold PBS (pH 7.2) before they were fixed in 2.5% glutaraldehyde in 0.2 M phosphate buffer (pH 7.4) for 2 h. After washing in phosphate buffer, the samples were post-fixed with 1% OsO_4 in the same buffer at room temperature for 1 h, followed by washing in tap water, dehydrating through a graded ethanol series, and embedding in Araldite. Semi-thin sections (1 µm) were prepared using a Leica Reichert Om/U3 ultra-microtome (Leica Microsystems, Wetzlar, Germany). Staining was performed with 0.5% toluidine blue in 0.5% sodium borate. The positioning of cross-sections is given in Figure S1.

2.3. Microscopy

The external morphology of wasp spiders was examined via light microscopy on a Keyence VHX stereomicroscope. Histological sections were examined with a DM 4 B microscope (Leica Microsystems, Wetzlar, Germany).

3. Results

3.1. The external morphology of the wasp spider venom system

The external morphology of the wasp spider venom system is composed of a pair of chelicerae consisting of an enlarged basal segment and a smaller curved fang proximal to the basal segment (Figure 2). The orientation of the chelicerae is labidognathous, thus both parts are directed towards each other. Some sensory hairs are distributed across the basal segments. The ventral side of the chelicerae harbor a series of cheliceral teeth. These effectively form an U-shaped pocket into which the fang is enfolded when in resting position.

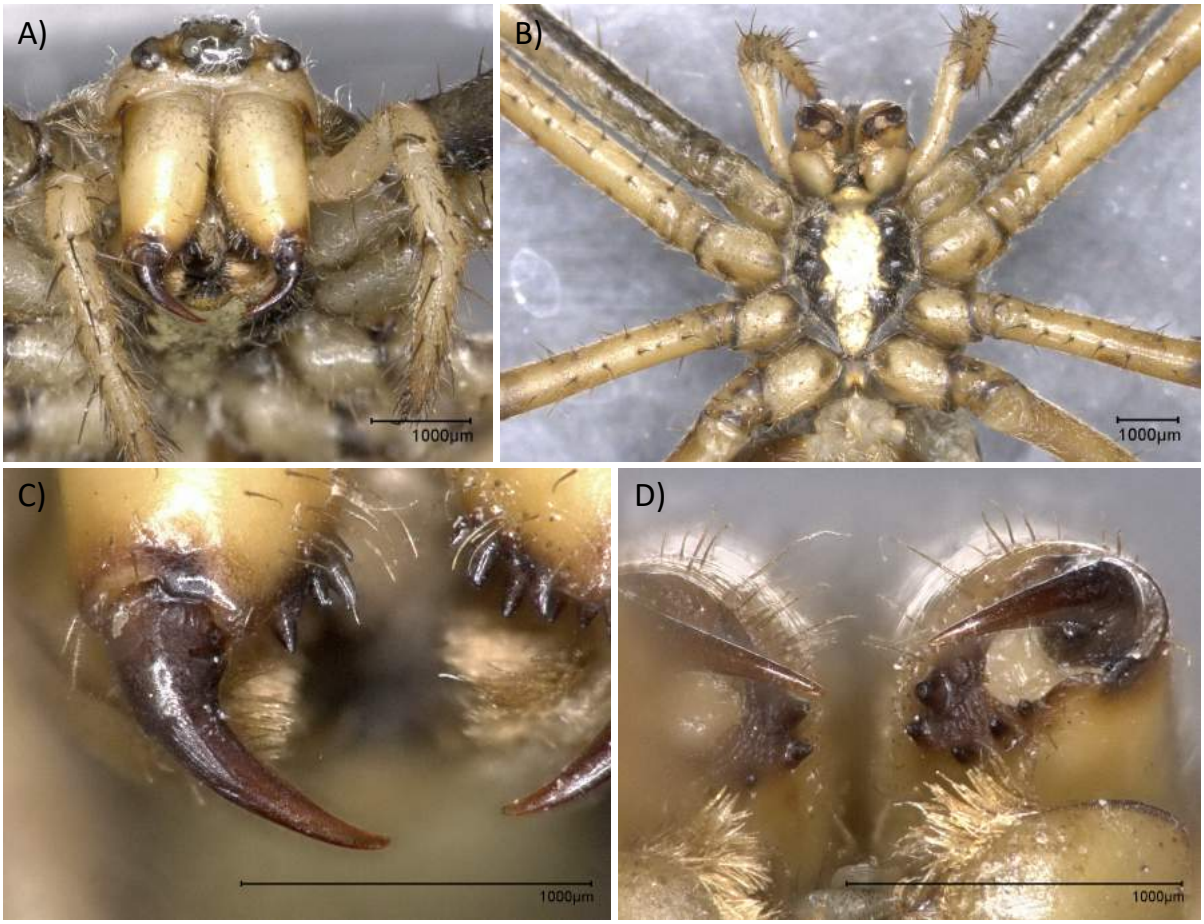


Figure 2. The external morphology of the wasp spider venom apparatus. Anterior (a) and ventral view (b) on the chelicerae of *A. bruennichi* illustrates the labidognathous orientation of the venom apparatus. It is composed of an enlarged cheliceral basal segment and a small curved fang at the tip. Magnification of the right fang is given in (c). A magnified ventral view on the cheliceral basal segment (d) highlights the assembly of cheliceral teeth (arrows), forming a pocket into which the fangs can be enfolded.

3.2. From fang to reservoir: Structural analysis of the venom apparatus

The analysis of the serial semi-thin sections throughout the prosoma illustrated the course of the venom system in *A. bruennichi*. Starting from the distal part of the paired venom ducts in the cheliceral fangs, the venom apparatus ends with the posterior parts of the glands within the prosoma. According to the location and the cellular structure, the venom ducts can be differentiated into four discrete parts: I) the orifical venom duct (ovd) in the distal parts of the fangs which opens with a pore, II) a distal venom duct (dvd) in the proximal part of the fangs and in the cheliceral basal segment, III) a central venom duct (cvd) in the basal segment, and IV) a proximal venom duct (pvd) in immediate contact to the venom gland in the prosoma. All parts show different structural characteristics.

The orifice of the venom duct opens nearby the fang tips (Figures 3, 4a). The fangs are surrounded by an outer- and inner endocuticle formed by a hypodermis consisting of cuboidal cells (Figures 4b, c). The orificidal venom duct proceeding in the distal part of the fang is arranged by a flattened epithelium (Figure 3c) lying on a small layer of connective tissue. Apically, the epithelium comprises a thin cuticle layer (Figures 4b, c). As the cellular organization in this part of the venom duct differs from the following part, we refer to both

as two different units (Figure 3d). Thus, cross sections of the distal venom ducts within the proximal fang and the basal segments reveal cuboidal cells with a slightly fringed apical region surrounding a broad luminal cavity (Figure 4). A few droplet-like granules are frequently distributed in the luminal cavity in this part of the venom duct. Its epithelium is also lying on a thin layer of connective tissue.

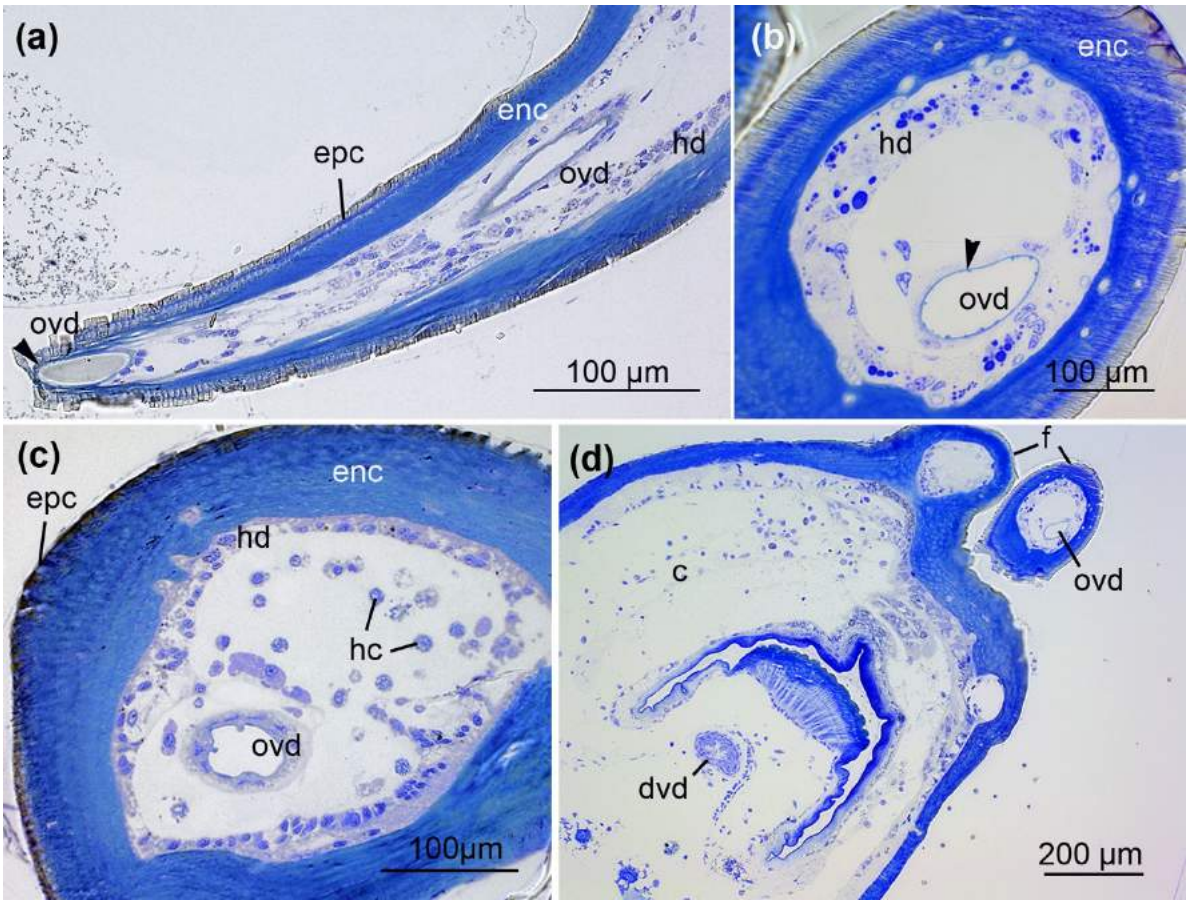


Figure 3. Course of the venom duct in the fang (S1). (a) Longitudinal section of a fang nearby the orifice of the venom duct (arrowhead). (b) and (c) Cross sections of a fang with orificidal venom duct (ovd). The fang is surrounded by an outer epicuticle (epc) and inner endocuticle (enc) formed by a hypodermis (hd) consisting of cuboidal epithelial cells. Hemocytes (hc) are present. The epithelium of the orificidal venom duct (ovd) comprises a thin cuticle layer (arrowhead). (d) Cross section of the chelicera (c) reveals two different venom ducts: the orificidal venom duct (ovd) within the fang and the distal venom duct (dvd) in the basal segment.

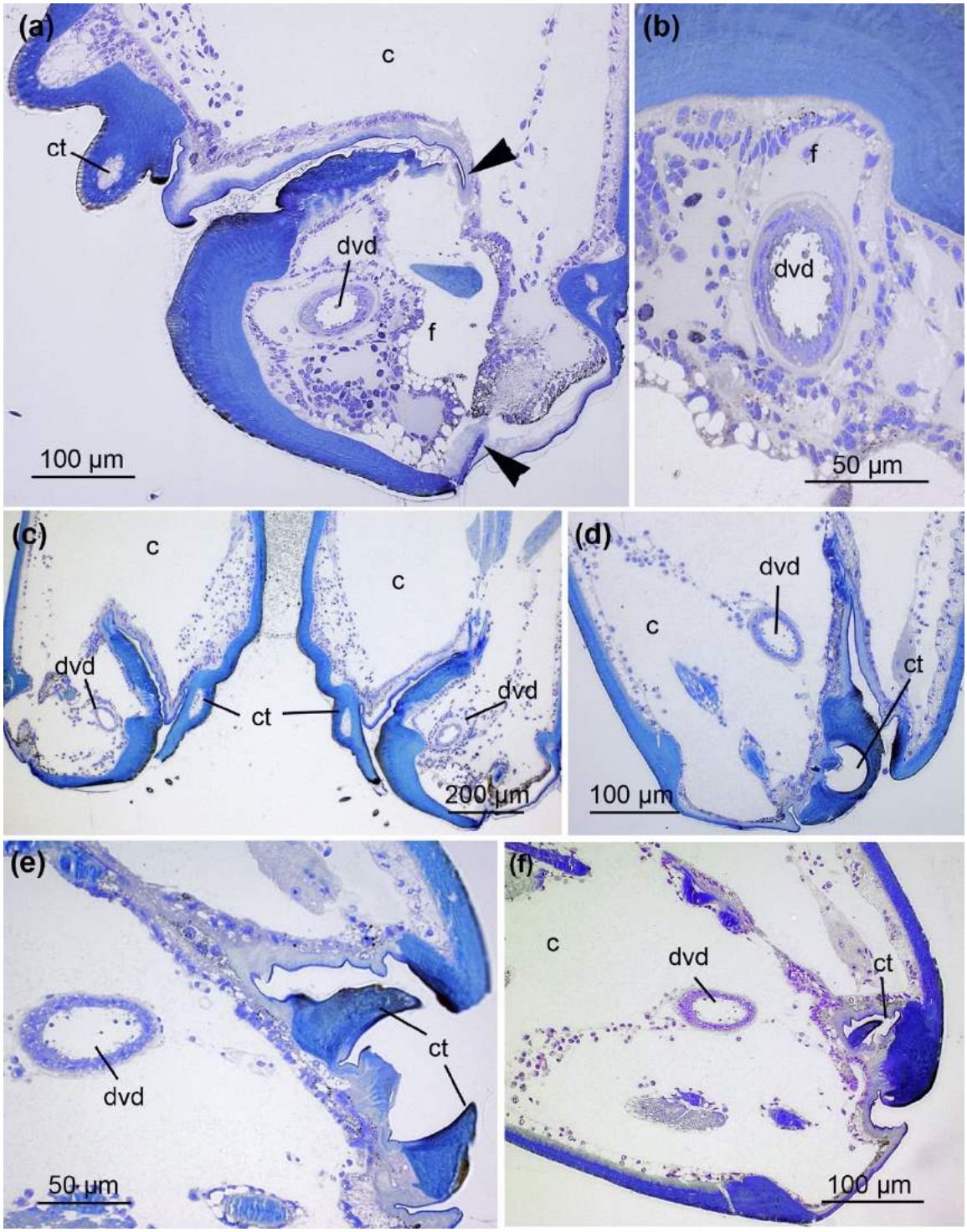


Figure 4. The section (S2) through the chelicera illustrates the course of the distal venom duct (dvd). (a) The fang (f) with the distal venom duct separates (arrowheads) from the cheliceral basal segment (c). At the ventral surface of the chelicera cheliceral teeth (ct) are localized. (b) Magnification of the distal venom duct (dvd) at the base of the fang reveals a cuboidal epithelium with its slightly fringed apical region surrounding a broad luminal cavity. (c) - (d) Serial sections of basal segment showing the assembly of cheliceral teeth (ct).

The anterior parts of the prosoma and the cheliceral basal segments are packed with muscles inoculating the pharynx in caudal parts of the prosoma and the chelicera, respectively (Figure 5a). At this level the central venom duct transverses the cheliceral basal segment. It consists of a single layer of flat cells with slender nuclei supported by a small layer of connective tissue (Figure 5b). It proceeds dorsally towards the region of the

emerging anterior part of the venom gland (Figures 6c, d). The duct again structurally alters close to the venom gland, thus forming a fourth section of the venom duct, henceforth referred to as the proximal venom duct (Figure 5e). Here, a thick layer of connective tissue with flattened nuclei is surrounding the single layer of columnar epithelial cells. The apical parts of the cells form thin projections that enclose large granules towards the ducts lumen (Figure 5f). The proximal venom duct joins ventrally the anterior part of the venom gland in the prosoma (Figure 6). The transition from proximal venom duct to venom gland is marked by changes in the cellular organization. The epithelium of the proximal venom duct shows basally located nuclei and apically almost fringed projections. The epithelial cells of the venom gland are loaded with secretory granules of different sizes and small vesicles that increase towards the gland lumen. In contrast to the venom gland, the proximal venom duct lacks surrounding muscle tissue.

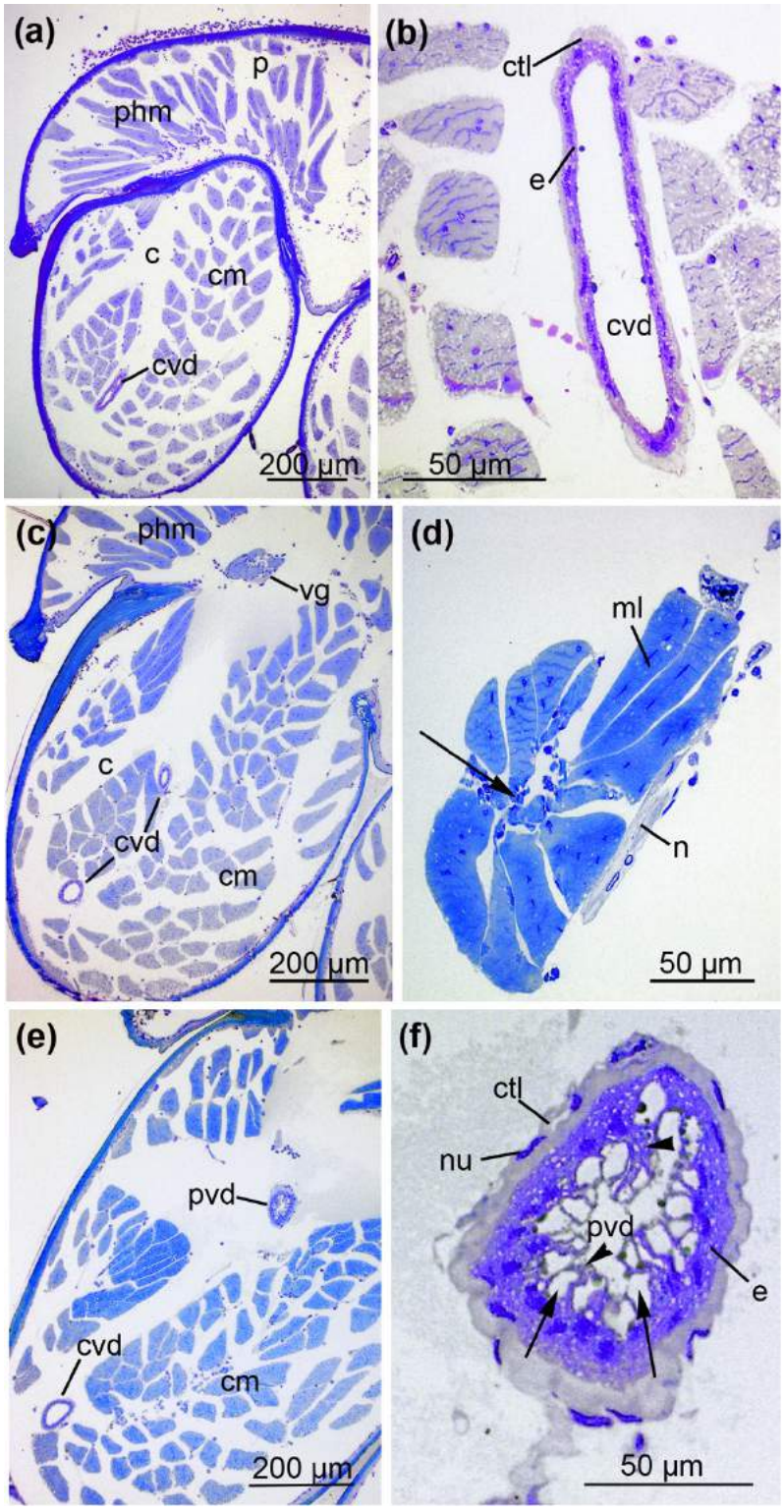


Figure 5. Location of central and proximal venom duct and anterior part of the venom gland. (a) Cross-section S3 through the prosoma (p) and cheliceral basal segment (c) including the central venom duct (cvd) plus pharynx (phm) and cheliceral muscles (cm). (b) The central venom duct with associated epithelium (e) and surrounding connective tissue (ct). (c) Cross-section S3 illustrates the course of the central venom duct towards the emerging venom gland (vg), which is linked to pharynx (phm) and cheliceral muscles (cm). (d) The anterior part of the venom gland with radially proceeding muscle fibers (ml) enclosing glandular cells (arrow). Nerves (n) are closely connected to the muscle fibers. (e) In cross-section S3/S4 the proximal venom duct (pvd) can be differentiated from the central duct. (f) Magnification of the proximal venom duct (S4). A layer of connective tissue with flattened nuclei (nu) surrounds the epithelial cells. The apical parts of the epithelium form thin projections (arrowheads) that enclose large granules (arrows) towards the lumen.

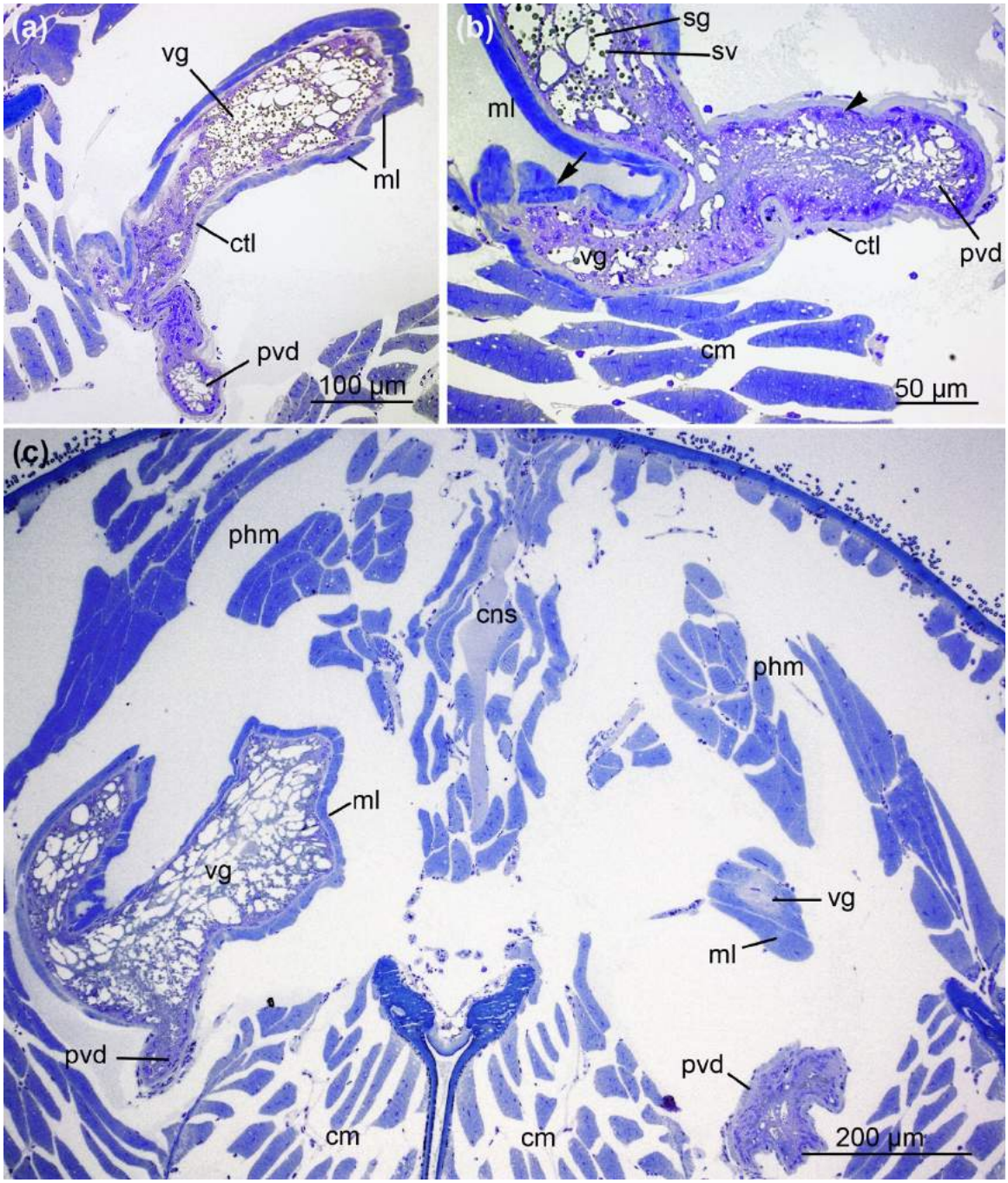


Figure 6. Convergence of venom duct and venom gland. ((a) The venom gland (vg) with secretory granules is surrounded by muscle fibers that are connected via a connective tissue layer (ctl) to the glandular cells. The proximal venom duct (pvd) reaches the venom gland. (c) The proximal venom duct enters the venom gland. The epithelium of the duct shows basally located nuclei (arrowheads) and apically projections, whereas the venom glands epithelium contains secretory granules (sg) and small vesicles (sv). Because of overlying fibers, the single layered muscle layer occasionally forms a double layer (arrow) (c) Cross section S4/S5 showing the proximal venom duct reaching the anterior part of the venom glands (vg) in the prosoma. cm = chelicera muscles, cns = central nervous system, ml = muscle layer, phm = pharynx muscles, pvd = proximal venom duct.

3.3. The architecture of venom glands

The anterior part of the venom gland is characterized by radially proceeding muscle fibers enclosing the first glandular cells. They are obliquely arranged in single layers alongside the gland and small nerves are closely connected to them (Figures 6d, 7). In sections, they are longitudinally or transversally aligned. Occasionally overlapping muscle fibers pretend to

form a double layer. When the gland enlarges, it abounds with a spongy network built by the glandular epithelium that is incorporated to a thick layer of densely packed muscles (Figure 7). The epithelial cells lie on a basal membrane which is again attached to a thick connective tissue with basally located nuclei. A large amount of secretory granules is apically located. The secretory epithelium of the gland forms an extended network of interdigitating cytoplasmic processes which enclose secretory granules with distinct patterns and include several vesicles of different sizes. These occur either isolated or in agglomerates inside the large secretory granules. In this area, the lumen of the gland is generally filled with secretory granules lacking smaller vesicles (Figure 7a). Nerve fibers are located along the venom gland and reach to the muscle layer (Figure 7b). The central part of the venom gland exhibits the largest diameter (Figure 7c). Here, the highest structural differences of secretory granules are found presuming that the content of the secretory granules of the glandular epithelium is highly diverse. Additionally, the amount and density of small coarse vesicles inside the secretory granules differs. In some parts of the gland, the whole apical cellular membranes of the epithelial cells dissolve and numerous small secretory vesicles discharge by holocrine secretion and distribute into the lumen. A magnification of the area elucidates a high diversity of secretory granules of the gland epithelial cells (Figure 7d). Most of the secretory granules are densely filled with small opaque vesicles whereas others seem to be rather empty as they are translucent. As the secretory granules increase in size, the epithelium is subdivided into units formed by an agglomeration of several cells with cytoplasmic projections reaching into the lumen. Alongside these projections, the large secretory granules proceed in the apical parts of the cells (Figure 7e). Towards the posterior part of the prosoma, the central glandular lumen of the venom gland subsequently becomes free of granules and vesicles (Figure 7e). This results in a canal-like empty luminal cavity. Here, distinct cellular membranes of the glandular epithelium enclose large densely packed secretory granules that are located around the lumen. Thus, the sponge-like arrangement of the epithelial apical cell region with its secretory granules drains into a central compartmentalized lumen.

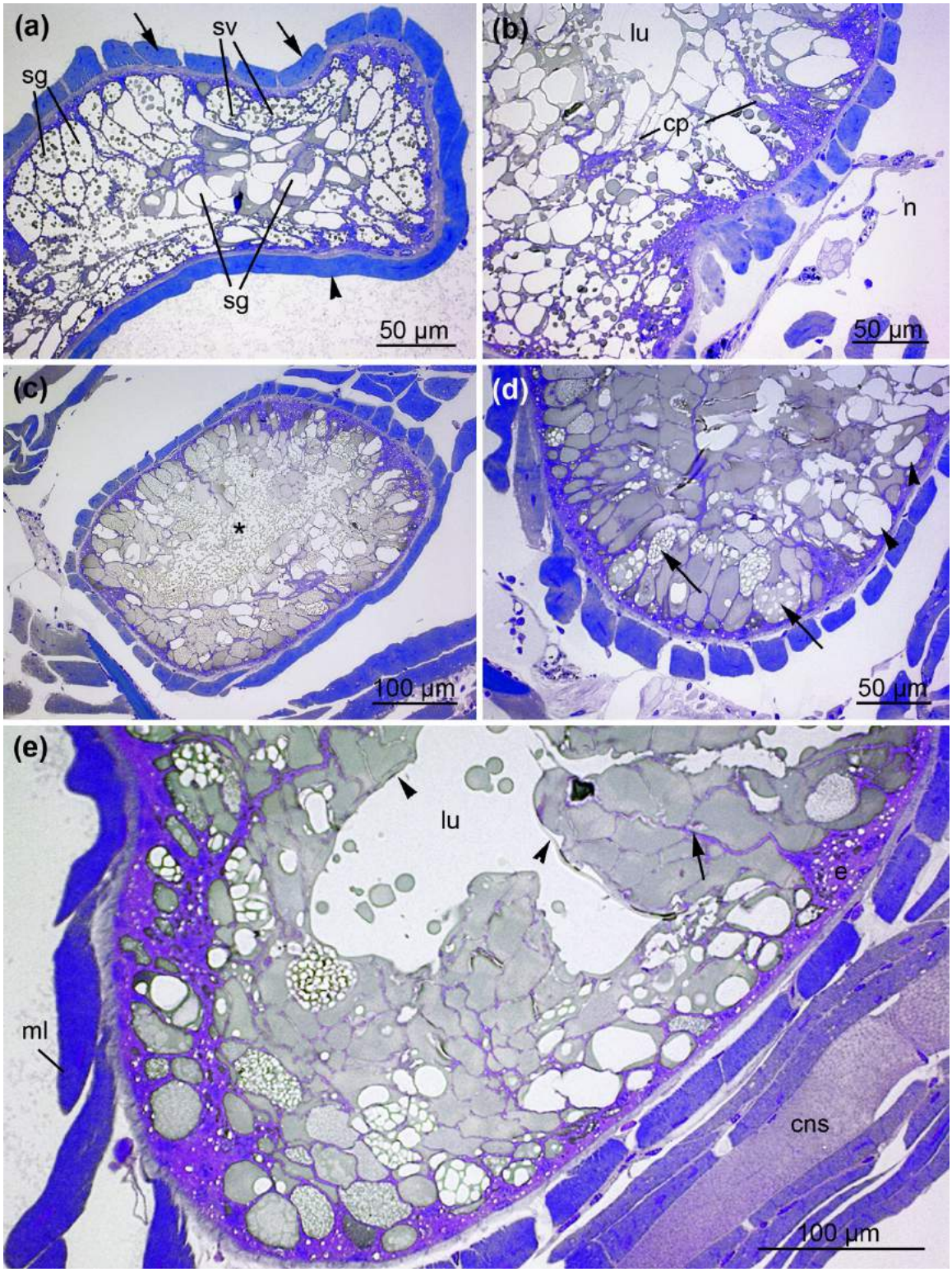


Figure 7. Anterior and central part of the venom gland. (a) Cross-section S5 illustrates the anterior part of the venom gland with longitudinally (arrowhead) and transversally (arrow) running muscle fibers. The gland harbors secretory granules (sg) and are either filled with or lack vesicles (sv. (b) Magnification reveals nervous fibers (n) linked to muscle layers (ml). Epithelial cells develop cytoplasmic projections (cp). (c) Cross-section S6 illustrates the central part of the venom gland in its biggest expansion. The asterisk indicates holocrine discharge of vesicles into the lumen. (d) Magnification elucidates the high diversity of secretory granules within central venom gland epithelial cells. Arrowheads indicate translucent granules, whereas arrows indicate those filled with opaque vesicles. (e) The central part of the venom gland with empty luminal cavity (lu). An agglomeration of cells forms cytoplasmic projections that reach into the lumen (arrow). A distinct membrane (arrowhead) covers the epithelium. cns = central nervous system

The posterior part of the venom gland is located at the level of the pharynx (Figure 8). The diameter of the venom gland decreases towards the posterior end of the prosoma. The central lumen is filled with large secretory granules which contain dense material whereas the number of secretory granules with small vesicles inside diminish.

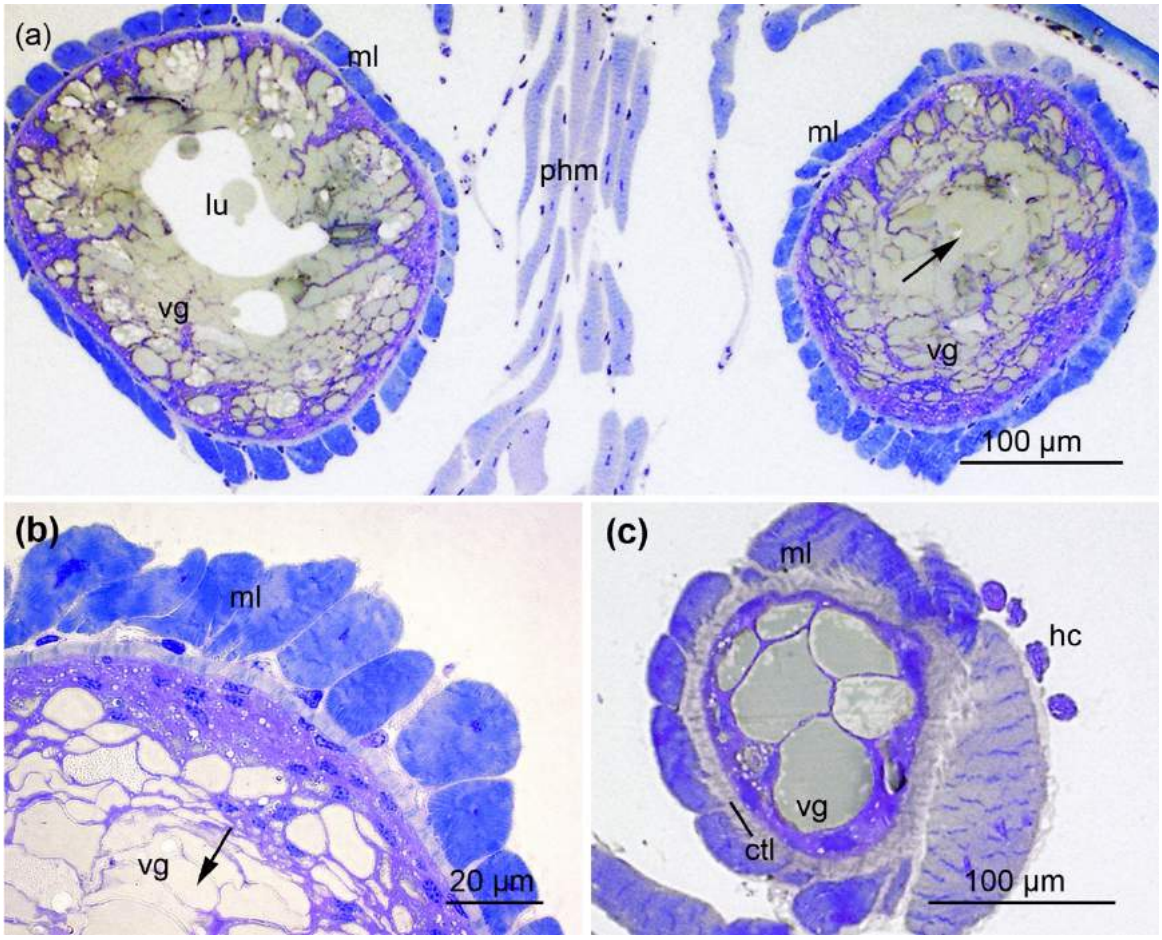


Figure 8. Posterior part of the venom gland. (a) Cross-section S7 gives an overview of the posterior part of the venom gland (vg). phm = pharynx muscles. Towards the posterior part of the gland the lumen is completely filled with secretory granules (arrow). (b) The magnification reveals secretory granules with differing dense material (arrow). (c) Almost posterior end of venom gland. ctl = connective tissue layer, hc = hemocyte, lu = lumen, ml = muscle layer.

4. Discussion

4.1. An overview on the venom system in wasp spiders

The venom gland of wasp spiders represents a relatively large organ that reaches deep into the prosoma. It is, like all so far studied spider venom glands, deeply covered by a variety of muscles that are the main tools enabling gland contraction and venom release. The gland represents a complex network of secretory cells that release vesicles which discharge their content into the lumen. These vesicles are likely filled with the venom components of wasp spiders that are synthesized in the secretory cells. Within these vesicles, the venom components migrate towards the lumen where the venom is stored until usage.

The venom system of spiders is constructed around chelicerae. These effectively work as hypodermic needles and enable the injection of venom into prey or predators [44]. In

chelicerae bearing arthropods, a variety of different types evolved that differ in their biomechanics [45]: The chelicerae of spiders follow a jack-knife mechanic. They are composed of two mobile units, a sharp fang on the tip followed by a larger basal segment that connects the fang with the prosoma. In resting position, the fang is folded against the basal segment. An unfolding of the fang enables the spider to deliver the venomous bite and venom is released from a small opening close to the fangs tip. In spiders two different types of chelicerae are differentiated by their orientation along the prosoma [45]: In orthognathous chelicerae, both fangs are facing downwards and work in parallel, whereas in labidognathous chelicerae the fangs face each other and work in a tweezer-like fashion. Moreover, orthognathous chelicerae are characterized by a large basal segment that connects the fang with the prosoma [44]. The venom gland is located within this basal segment. In labidognathous systems, the venom gland is instead positioned within the prosoma and the basal segment is reduced in comparison to orthognaths [44]. While orthognathous systems are found in the two ancestral infraorders Mesothelae and Mygalomorphae, labidognathous systems are present in Araneomorphae [44]. It has been proposed, that the migration of the venom gland from the basal segment of orthognaths into the prosoma of labidognaths enabled the reduction in body size that is observed in Araneomorphae versus other infraorders, without imposing spatial constraints on the venom system [6,46]. This reduction in body size seemingly enabled the evolution of a web-based lifestyle that is found in many of the araneomorph taxa and thus likely contributed to their evolutionary success [46].

Wasp spiders belong to the large family of Araneidae in which one of the most diverse radiations occurred within Araneomorphae [4]. It is thus rather unsurprising that we confirm the labidognathous character of *A. bruennichi* chelicerae. The wasp spider carries cheliceral teeth ventrally to its chelicerae. In these, the fang is placed when the venom delivery system is in resting position (Figures 3, 5). Cheliceral teeth are found in several distinct spider lineages and they may enable a secure grip on prey items [47]. Moreover, in spiders that evolved cheliceral teeth, these structures are used in extraintestinal digestion by mechanically breaking up tissue [44]. In addition to the release of digestive fluids on the prey, the spider uses the cheliceral teeth as a support to masticate and liquify its victim. Albeit the wasp spider follows a largely silk-based hunting behavior in which the venom apparatus is mostly omitted [32], the presence of cheliceral teeth indicates that the chelicerae represent valuable tools for the species. In particular, their versatility in prey handling and consumption instead a role in prey subjugation seems to be of importance for wasp spiders.

The fang is characterized by a thick cuticle (Figure 3) built by the underlying hypodermis and, together with the chitin of the exoskeleton, likely contributes to the stabilization of the fang during bite and venom injection. The venom duct, from the fangs orifice to the basal segment of the chelicerae, consists only of a flat epithelium forming a thin cuticle layer and is surrounded by a hemolymph rich in hemocytes (Figure 3). When the fang proceeds into the basal segment of the chelicerae, a variety of morphological alterations occur (Figures 5, 6). Primarily, the diameter of the venom system increases and the basal segment is filled with muscle fibers, probably enabling movement of the chelicerae. Externally of the chelicerae, the prosoma is rich in pharynx muscles that seem to be connected to parts of the cheliceral muscle apparatus, suggesting a functional interplay

of both systems. This is in particular the case for the venom gland that begins at the proximal end of the chelicerae. It is embedded in both, the cheliceral muscles and the pharynx muscles and this assembly may, besides facilitating the movement of the chelicera, support the primary muscles along the venom gland during contraction for venom release. The functionality of this complex system is moreover supported by their connection to small nerves that may regulate and fine-tune its utilization (Figure 5). A comprehensive overview about the venom system of *A. bruennichi* is given in Figure 9.

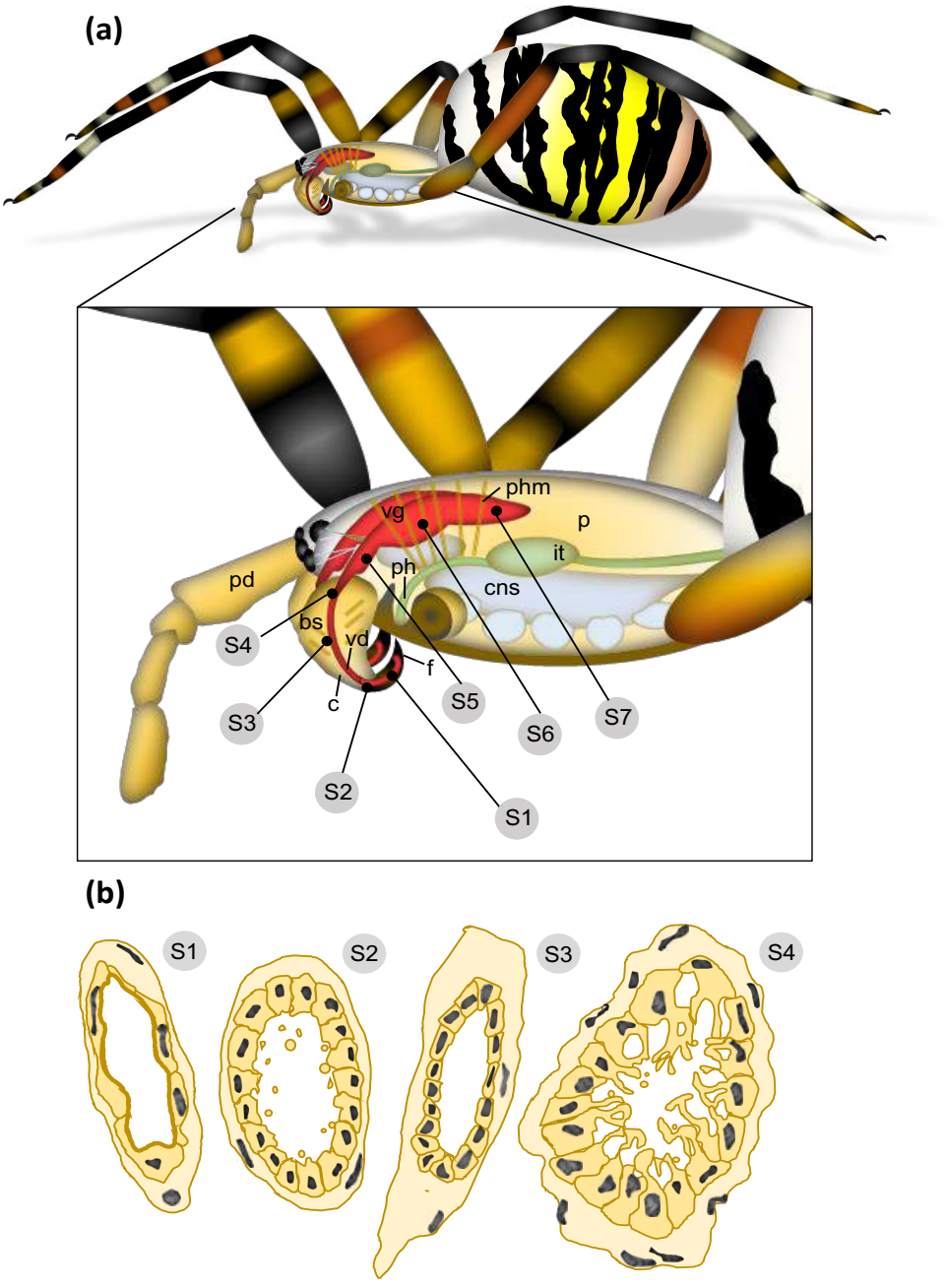


Figure 9. Internal anatomy of the venom system. (a) The labeling S1-S7 give information about the position of cross-sections shown in the following figures. Inside of the prosoma lies the central nervous system, an extensive musculature for the extremities and the pharynx, part of the intestinal tract, and a pair of venom glands. Each venom system consists of long, cylindrical part and an adjoining duct, which terminates at the tip of the cheliceral fang. (b) Schematic drawing of different subunits of the venom duct. S1: ovd = orificidal venom duct of the fang; S2: dvd = distal venom duct of the fang and the distal basal segment; S3: cvd = central venom duct of the basal segment; S4: pvd = proximal venom duct in the basal segment/prosoma bs = basal

segment of the chelicera, c = chelicera, cns = central nervous system, f = fang, it = intestinal tract, p = prosoma, pd = pedipalp , ph = pharynx, phm = pharynx muscle, vd = venom duct, vg = venom gland.

4.2. Hidden complexity of the venom duct

While the venom duct resembled a simple structure in the fang, it gains structural complexity as it proceeds through the chelicerae towards the venom gland (Figure 9b). In the chelicerae, it is composed by a flattened simple epithelium with slender nuclei and a layer of connective tissue (Figure 5). The venom duct alters proximal to the venom gland even more. Here, it comprises columnar epithelial cells that are surrounded by a thick layer of connective tissue, rich in flattened nuclei. Apical parts of the epithelium form projections enclosing large granules towards the ducts lumen. Given these obvious differences in cellular organization throughout the venom duct, our analyses recovered that it consists of four different subunits: A first orificidal region in the fang, a second distal and third central venom duct. Both being present throughout most parts of the basal segment within the chelicerae. A fourth proximal region is present close to the convergence of venom duct and venom gland in the prosoma.

According to our knowledge, this structural diversification throughout the venom duct has not yet been described for any spider. Generally, knowledge upon the importance of venom ducts for its linked venom system appears to be scarce and such ducts are mostly considered as the connection between the injector and the venom gland [44]. At least for cone snails and *Bothrops* pitvipers it has been demonstrated that toxins can be expressed by the venom duct and not exclusively by the venom gland [48, 49]. Similar to the wasp spider, venom ducts of cone snails were recently recovered as a functional trinity [50]. However, venom ducts may be much more important for the functionality of a given venom system. Firstly, they form a thin bottle-neck between the venom gland and the fang, thus they may represent major factors influencing injection pressure and thus envenomation efficiency. Secondly, venom ducts may have also a metabolic role for the venom system. Their venom duct is also subdivided into a distal, central and proximal duct and each of these were shown to express a different subset of toxins. It has been proposed that subdivisions within the cone snail venom duct were evolutionary specialized towards biosynthesis of specific conotoxins [50]. Our findings, that the venom duct of wasp spiders is structurally, and therefore likely functionally, subdivided into four discrete parts suggests a potential metabolic role of the venom duct. In particular, the proximal duct resembles parts of the venom gland as it contains a loose network of thin projections with large granules. This interesting subject is awaiting further investigation.

4.3. The arachno-atypical venom of wasp spiders is not reflected in the venom systems morphology

Parts of the herby conducted research tested if the unusual venom of wasp spiders, which is of astonishing simplicity in comparison to other spider venoms [41], is also reflected in the venom systems morphology. Venom is a costly resource and a loss of toxic components or even the whole venom cocktail can occur when a species adapts to a novel ecological niche where venom becomes obsolete. For example, sea snakes that switched from fish hunting to an egg-based diet completely lost their venom [51]. Scorpions that rely on their pedipalps for hunting usually have less complex venoms than their relatives that

favour their stinging apparatus and in centipedes it has recently been established, that their venom evolves under morphological constraints [43, 52].

The wasp spider evolved a rather simple venom composition, likely as a consequence of a mostly silk-based hunting strategy. However, our findings reject the hypothesis that the architecture of its venom apparatus differs from those described from other previously studied. Reflecting the typical morphology for araneomorphs, it is composed of labidognathous chelicerae that appear functionally and shape-wise comparable to spiders with a similar body size. Moreover, the associated venom gland is rather large and reaches deep into the prosoma, indicating that this system yields large quantities of venom for a small araneomorph spider. The muscular layers that cover the venom gland as well as the attached nerves mirror the structure present in spiders from other families. The overall structure of the venom gland equally resembles that of other spiders. In all species so far studied, the venom gland comprises a complex network of secretory cells, granules, vesicles and cytoplasmic projections albeit minor structural differences occur.

5. Conclusions

The spider kingdom represents one of the most diverse branches of the metazoan tree of life and virtually all (except one) of its included families carry functional venom systems. These venom systems are likely to yield a plethora of powerful bioressources and thus were investigated intensely in the past. Despite this, the vast majority of spiders remains unstudied for their venom so far and even less species have been studied for their venom systems architecture.

Here, we conducted a thorough morphological analysis on the wasp spider *A. bruennichi* a member of the understudied Araneidae family. By focusing on a venom system-wise neglected family, our study contributes to the understanding of spider venom system as it closes a prevailing taxonomic gap within the available literature. Among the array of insights regarding the architecture of the wasp spider venom system, we recovered a previously hidden structural complexity of the venom duct that is divided into four distinct regions. In particular, the proximal venom duct seems to be structurally reorganized towards a potential role in venom production. In subsequent studies this finding should be further pursued by applying more adequate technologies such as MALDI imaging and µ-CT 3D reconstructions. Our hypothesis, that the arachno-atypical venom of wasp spiders may be reflected in an unusual venom system morphology could be rejected. Contrary to our working hypothesis, we recovered a widespread agreement between structures of previously studied spider venom systems with *A. bruennichi*.

Author Contributions: Conceptualization, T.L., A.V. and S.L.; methodology, T.L., H.S.; software, H.S.; validation, H.S.; formal analysis, T.L., H.S.; investigation, T.L., H.S.; writing—original draft preparation, T.L.; writing—review and editing, B.M.vR, S.L., A.V.; visualization, H.S.; supervision, B.M.vR., S.L.; project administration, S.L.; funding acquisition, A.V. All authors have read and agreed to the published version of the manuscript.

Funding: Research was funded by the Hesse State Ministry of Science and Art (HMWK) in context of the LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG).

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] V. Schendel, L. D. Rash, R. A. Jenner, and E. A. B. Undheim, "The diversity of venom: The importance of behavior and venom system morphology in understanding its ecology and evolution," *Toxins*. 2019, doi: 10.3390/toxins11110666.

[2] B. G. Fry *et al.*, "The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms," *Annu. Rev. Genomics Hum. Genet.*, 2009, doi: 10.1146/annurev.genom.9.081307.164356.

[3] N. R. Casewell, W. Wüster, F. J. Vonk, R. A. Harrison, and B. G. Fry, "Complex cocktails: The evolutionary novelty of venoms," *Trends in Ecology and Evolution*. 2013, doi: 10.1016/j.tree.2012.10.020.

[4] World Spider Catalog, "World Spider Catalog Version 20.5.," *Nat. Hist. Museum Bern*, 2019, doi: 10.24436/2.

[5] N. J. Saez *et al.*, "Spider-venom peptides as therapeutics," *Toxins*. 2010, doi: 10.3390/toxins2122851.

[6] N. Langenegger, W. Nentwig, and L. Kuhn-Nentwig, "Spider venom: Components, modes of action, and novel strategies in transcriptomic and proteomic analyses," *Toxins*. 2019, doi: 10.3390/toxins11100611.

[7] N. J. Saez and V. Herzig, "Versatile spider venom peptides and their medical and agricultural applications," *Toxicon*, 2019, doi: 10.1016/j.toxicon.2018.11.298.

[8] K. L. Richards *et al.*, "Selective NaV1.1 activation rescues Dravet syndrome mice from seizures and premature death," *Proc. Natl. Acad. Sci. U. S. A.*, 2018, doi: 10.1073/pnas.1804764115.

[9] G. F. King and M. C. Hardy, "Spider-Venom Peptides: Structure, Pharmacology, and Potential for Control of Insect Pests," *Annu. Rev. Entomol.*, 2013, doi: 10.1146/annurev-ento-120811-153650.

[10] I. R. Chassagnon *et al.*, "Potent neuroprotection after stroke afforded by a double-knot spider-venom peptide that inhibits acid-sensing ion channel 1a," *Proc. Natl. Acad. Sci. U. S. A.*, 2017, doi: 10.1073/pnas.1614728114.

[11] V. Herzig, G. F. King, and E. A. B. Undheim, "Can we resolve the taxonomic bias in spider venom research?," *Toxicon X*, 2019, doi: 10.1016/j.toxcx.2018.100005.

[12] T. Lüddecke, A. Vilcinskas, and S. Lemke, "Phylogeny-guided selection of priority groups for venom bioprospecting: Harvesting toxin sequences in tarantulas as a case study," *Toxins (Basel).*, vol. 11, no. 9, 2019, doi: 10.3390/toxins11090488.

[13] V. L. P. Dos Santos *et al.*, "Structural and ultrastructural description of the venom gland of *Loxosceles intermedia* (brown spider)," *Toxicon*, 2000, doi: 10.1016/S0041-0101(99)00155-5.

[14] L. M. Silva *et al.*, "Structural analysis of the venom glands of the armed spider *Phoneutria nigriventer* (Keyserling, 1891): Microanatomy, fine structure and confocal observations," *Toxicon*, 2008, doi: 10.1016/j.toxicon.2007.12.009.

[15] E. V. Grishin, "Black widow spider toxins: The present and the future," in *Toxicon*, 1998, doi: 10.1016/S0041-0101(98)00162-7.

[16] J. E. Garb, "Extraction of venom and venom gland microdissections from spiders for proteomic and transcriptomic analyses," *J. Vis. Exp.*, 2014, doi: 10.3791/51618.

[17] U. Järlfors, D. S. Smith, and F. E. Russell, "Nerve endings in the venom gland of the spider *Latrodectus mactans*," *Toxicon*, 1969, doi: 10.1016/0041-0101(69)90025-7.

[18] D. S. Smith and F. E. Russel, "Structure of the venom gland of the black widow spider *Latrodectus mactans*. A preliminary light and electron microscopic study" in *Animal Toxins*, 1967.

[19] J. Kovoor and A. Muñoz-Cuevas, "Comparative histology of the venom glands in a lycosid and several oxyopid spiders (Araneae)," *Ekologia Bratislava*. 2000.

[20] N. Yigit, A. Bayram, T. Danisman, Z. Sancak, and M. G. Tel, "Morphological characterization of the venom apparatus in the wolf spider *Lycosa singoriensis* (Laxmann, 1770)," *J. Venom. Anim. Toxins Incl. Trop. Dis.*, 2009, doi: 10.1590/S1678-91992009000100013.

[21] K. Çavuşoğlu, A. Bayram, M. Maraş, T. Kirindi, and K. Çavuşoğlu, "A morphological study on the venom apparatus of spider *Larinioides cornustus* (Araneae, Araneidae)," *Turkish J. Zool.*, 2005.

[22] N. Yiğit, A. Bayram, T. Danişman, and Z. Sancak, "Functional morphology of the venom apparatus of *Larinioides ixobolus* (Araneae: Araneidae)," *Pakistan J. Biol. Sci.*, 2006, doi: 10.3923/pjbs.2006.1975.1978.

[23] T. A. A. Rocha-e-Silva, C. B. Collares-Buzato, M. A. da Cruz-Höfling, and S. Hyslop, "Venom apparatus of the brazilian tarantula *Vitalius dubius* Mello-Leitão 1923 (Theraphosidae)," *Cell Tissue Res.*, 2009, doi: 10.1007/s00441-008-0738-x.

[24] M. Benli, M. Karakas, N. Yigit, and S. Cebesoy, "Determining with SEM, structure of the venom apparatus in the tube web spider, *Segestria florentina* (Araneae: Segestriidae)," *J. Entomol. Zool. Stud.*, 2013.

[25] S. Malt, F. W. Sander, and G. Schaller, "Contribution to foraging ecology of selected Araneidae in xerophil grasslands with particular consideration of *Argiope brunnichii* Scop.," *Zool. Jahrbucher Abteilung fur Systenatik, Okol. und Geogr. der Tiere*, 1990.

[26] D. G. E. Gomes, "Orb-weaving spiders are fewer but larger and catch more prey in lit bridge panels from a natural artificial light experiment," *PeerJ*, 2020, doi: 10.7717/peerj.8808.

[27] R. E. Buskirk, "Coloniality, Activity Patterns and Feeding in a Tropical Orb-Weaving Spider," *Ecology*, 1975, doi: 10.2307/1934699.

[28] A. M. Heiling, "Why do nocturnal orb-web spiders (Araneidae) search for light?," *Behav. Ecol. Sociobiol.*, 1999, doi: 10.1007/s002650050590.

[29] J. M. Biere and G. W. Uetz, "Web Orientation in the Spider *Micrathena Gracilis* (Araneae: Araneidae)," *Ecology*, 1981, doi: 10.2307/1936708.

[30] K. W. Welke and J. M. Schneider, "Sexual cannibalism benefits offspring survival," *Anim. Behav.*, 2012, doi: 10.1016/j.anbehav.2011.10.027.

[31] M. Nyffeler and G. Benz, "Foraging ecology and predatory importance of a guild of orb-weaving spiders in a grassland habitat," *J. Appl. Entomol.*, 1989, doi: 10.1111/j.1439-0418.1989.tb00246.x.

[32] T. Eisner and J. Dean, "Ploy and counterploy in predator - prey interactions: orb weaving spiders versus bombardier beetles," *Proc. Natl. Acad. Sci. U. S. A.*, 1976, doi: 10.1073/pnas.73.4.1365.

[33] R. Václav and P. Prokop, "Does the appearance of orbweaving spiders attract prey?," *Ann. Zool. Fennici*, 2006.

[34] L. Fromhage, G. Uhl, and J. M. Schneider, "Fitness consequences of sexual cannibalism in female *Argiope bruennichi*," *Behav. Ecol. Sociobiol.*, 2003, doi: 10.1007/s00265-003-0656-6.

[35] S. P. Chinta, S. Goller, J. Lux, S. Funke, G. Uhl, and S. Schulz, "The sex pheromone of the wasp spider *Argiope bruennichi*," *Angew. Chemie - Int. Ed.*, 2010, doi: 10.1002/anie.200906311.

[36] A. Walter, P. Bliss, M. A. Elgar, and R. F. A. Moritz, "*Argiope bruennichi* shows a drinking-like behaviour in web hub decorations (Araneae, Araneidae)," *J. Ethol.*, 2009, doi: 10.1007/s10164-007-0077-5.

[37] H. Krehenwinkel, D. Rödder, and D. Tautz, "Eco-genomic analysis of the poleward range expansion of the wasp spider *Argiope bruennichi* shows rapid adaptation and genomic admixture," *Glob. Chang. Biol.*, 2015, doi: 10.1111/gcb.13042.

[38] H. Krehenwinkel and D. Tautz, "Northern range expansion of European populations of the wasp spider *Argiope bruennichi* is associated with global warming-correlated genetic admixture and population-specific temperature adaptations," *Mol. Ecol.*, 2013, doi: 10.1111/mec.12223.

[39] S. M. Zimmer, H. Krehenwinkel, and J. M. Schneider, "Rapid range expansion is not restricted by inbreeding in a sexually cannibalistic spider," *PLoS One*, 2014, doi: 10.1371/journal.pone.0095963.

[40] W. Wawer, R. Rutkowski, H. Krehenwinkel, D. Lutyk, K. Pusz-Bocheńska, and W. Bogdanowicz, " Population structure of the expansive wasp spider (*Argiope bruennichi*) at the edge of its range ," *J. Arachnol.*, 2017, doi: 10.1636/joa-s-16-056.1.

[41] T. Lüddecke *et al.*, "An Economic Dilemma Between Molecular Weapon Systems May Explain an Arachnological-atypical Venom in Wasp Spiders (*Argiope bruennichi*)," *Biomolecules*, vol. 10, no. 7, p. 978, 2020.

[42] Z. Duan, R. Cao, L. Jiang, and S. Liang, "A combined de novo protein sequencing and cDNA library approach to the venomic analysis of Chinese spider *Araneus ventricosus*," *J. Proteomics*, 2013, doi: 10.1016/j.jprot.2012.10.011.

[43] E. A. B. Undheim *et al.*, "Production and packaging of a biological arsenal: Evolution of centipede venoms under morphological constraint," *Proc. Natl. Acad. Sci. U. S. A.*, 2015, doi: 10.1073/pnas.1424068112.

[44] R. F. Foelix, "Biology of Spiders," *Insect Syst. Evol.*, 1983, doi: 10.1163/187631283X00371.

[45] S. L. Zonstein, "The spider chelicerae: some problems of origin evolution," *Proc. 21st Eur. Colloq. Arachnol.*, 2003.

[46] L. Kuhn-Nentwig, R. Stöcklin, and W. Nentwig, "Venom composition and strategies in spiders. is everything possible?," in *Advances in Insect Physiology*, 2011.

[47] M. J. Moon and M. H. Yu, "Fine structure of the chelicera in the spider *Nephila clavata*," *Entomol. Res.*, 2007, doi: 10.1111/j.1748-5967.2007.00108.x.

[48] H. Hu, P. K. Bandyopadhyay, B. M. Olivera, and M. Yandell, "Characterization of the *Conus bullatus* genome and its venom-duct transcriptome," *BMC Genomics*, 2011, doi: 10.1186/1471-2164-12-60.

[49] R. H. Valente *et al.*, "The primary duct of *Bothrops jararaca* glandular apparatus secretes toxins," *Toxins (Basel).*, 2018, doi: 10.3390/toxins10030121.

[50] L. L. Tayo, B. Lu, L. J. Cruz, and J. R. Yates, "Proteomic analysis provides insights on venom processing in *Conus textile*," *J. Proteome Res.*, 2010, doi: 10.1021/pr901032r.

[51] M. Li, B.G. Fry, R.M. Kini, "Eggs-only diet: its implications for the toxin rofile changes and ecology of the marbled sea snake (*Aipysurus eydouxii)*. J Mol Evol, 2005, 60(1): 81-89.

[52] G.S. Casper, "Prey capture and stinging behaviour in the Emporer Scorpion, *Pandinus imperator* (Koch) (Scorpiones, Scorpionidae). J Arachnol 1985, 13(3): 277-283.

Chapter V

An Economic Dilemma Between Molecular Weapon Systems May Explain an Arachno-atypical Venom in Wasp Spiders (*Argiope bruennichi*)

Tim Lüddecke, Björn M. von Reumont, Frank Förster, André Billion, Thomas Timm, Günther Lochnit, Andreas Vilcinskas & Sarah Lemke
Biomolecules (2020) 10(7): 978.

biomolecules

MDPI

Article

An Economic Dilemma Between Molecular Weapon Systems May Explain an Arachno-atypical Venom in Wasp Spiders (*Argiope bruennichi*)

Tim Lüddecke [1,2,*], Björn M. von Reumont [2,3], Frank Förster [4], André Billion [1], Thomas Timm [5], Günter Lochnit [5], Andreas Vilcinskas [1,2,3] and Sarah Lemke [2,3]

1 Department of Bioresources, Fraunhofer Institute for Molecular Biology and Applied Ecology, Ohlebergsweg 12, 35392 Gießen, Germany; andre.billion@ime.fraunhofer.de (A.B.); andreas.vilcinskas@ime.fraunhofer.de (A.V.)
2 LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Senckenberganlage 25, 60325 Frankfurt, Germany; bmvr@arcor.de (B.M.v.R.); sarah.lemke@ime.fraunhofer.de (S.L.)
3 Institute for Insect Biotechnology, Justus-Liebig-University of Gießen, Heinrich-Buff-Ring 26-32, 35392 Gießen, Germany
4 Institute for Bioinformatics and Systems Biology, Justus-Liebig-University of Gießen, Heinrich-Buff-Ring 58, 35392 Gießen, Germany; frank.foerster@computational.bio.uni-giessen.de
5 Institute of Biochemistry, Justus-Liebig-University of Gießen, Friedrichstr. 24, 35392 Gießen, Germany; thomas.timm@biochemie.med.uni-giessen.de (T.T.); guenter.lochnit@biochemie.med.uni-giessen.de (G.L.)
* Correspondence: tim.lueddecke@outlook.com

Received: 10 June 2020; Accepted: 26 June 2020; Published: 30 June 2020

Abstract: Spiders use venom to subdue their prey, but little is known about the diversity of venoms in different spider families. Given the limited data available for orb-weaver spiders (Araneidae), we selected the wasp spider *Argiope bruennichi* for detailed analysis. Our strategy combined a transcriptomics pipeline based on multiple assemblies with a dual proteomics workflow involving parallel mass spectrometry techniques and electrophoretic profiling. We found that the remarkably simple venom of *A. bruennichi* has an atypical composition compared to other spider venoms, prominently featuring members of the cysteine-rich secretory protein, antigen 5 and pathogenesis-related protein 1 (CAP) superfamily and other, mostly high-molecular-weight proteins. We also detected a subset of potentially novel toxins similar to neuropeptides. We discuss the potential function of these proteins in the context of the unique hunting behavior of wasp spiders, which rely mostly on silk to trap their prey. We propose that the simplicity of the venom evolved to solve an economic dilemma between two competing yet metabolically expensive weapon systems. This study emphasizes the importance of cutting-edge methods to encompass the lineages of smaller venomous species that have yet to be characterized in detail, allowing us to understand the biology of their venom systems and to mine this prolific resource for translational research.

Keywords: venomics; *Argiope bruennichi*; CAP superfamily; ICK; neuropeptides; hunting behavior; spider venom; proteotranscriptomics; bioresources

1. Introduction

Spiders are diverse and species-rich arthropods that have conquered most terrestrial habitats. The 48,463 extant spider species [1] share a common body plan that has changed little during ~380 million years of evolution. Spiders also possess a unique biochemical toolbox that uses a combination of venom and silk to subdue prey, contributing to their evolutionary success [2]. Moreover, they represent one of the few orders of terrestrial animals in which almost all extant species feature a

functional venom system and are thus considered as the most successful group of venomous animals [3]. Accordingly, spiders play a pivotal ecological role as venomous predators by maintaining the equilibrium of insect populations [4].

Venoms are complex mixtures of low-molecular-weight compounds, peptides and proteins, which act as toxins by disrupting important physiological processes when injected into prey [5,6]. They are used for defense, predation or competitor deterrence, but in all cases, they are physiologically expensive traits that have been optimized by strong selective pressure for specific functions. This evolutionary streamlining often results in high selectivity and target-specific bioactivity, meaning that animal venoms are now considered valuable bioresources in the field of drug discovery [7]. Several blockbuster drugs have been derived from venom components [7], but they were also investigated as research tools, cosmetics, industrial enzymes or bioinsecticides [8–11].

Spider venoms tend to be chemically more complex than other animal venoms, and up to 3000 different venom components can be present in a single species [12,13]. It has been estimated that the sum of all spider venoms could ultimately yield 10 million bioactive molecules, but only 0.02% of this diversity has been discovered thus far [14,15]. Several promising drug candidates for stroke, pain, cancer and neural disorders have been identified in spider venoms [16–19]. The major components of these spider venoms are low-molecular-weight inhibitor cysteine knot (ICK) peptides with robust tertiary structures conferred by the presence of a pseudoknot motif of interweaved disulfide bonds [13,20,21]. Additionally described as knottins, such peptides are often neurotoxic and remarkably resistant to heat, osmotic stress and enzymatic digestion, making them ideal drug candidates [13]. Although ICK peptides are found in other arthropod venoms [22–25], the diversity of these peptides in spider venoms is unprecedented. Approximately 60% of all spider venom components accessible in UniProt [26] are ICK peptides; hence, these peptides are commonly perceived as the principal component of spider venoms [13].

The analysis of venoms formerly relied on fractionation methods that require large amounts of starting material [27,28]. Therefore, previous studies of spider venoms have focused on species with strong anthropocentric connections, such as those posing direct medical threats or those of extraordinary size, making the venom easier to access in sufficient quantities. This restricted analyses to members of the families Atracidae, Ctenidae, Theraphosidae, Sicariidae and Theridiidae, which represent a narrow sample of spider diversity [1,15]. More recently, the advent of high-throughput methods compatible with miniscule samples has provided the means to expand the scope of such studies from less-accessible species [12,29]. Combinations of genomics, transcriptomics, proteomics and microbiomics [30,31] now allow the analyses of venoms from previously neglected taxa [32] in an emerging field known as modern evolutionary venomics [29].

Several of the most species-rich spider lineages have not been studied at all in the context of venom systems [15,33]. In this study, we therefore considered the family Araneidae (orb weavers), which is the third-largest spider family, comprising 3078 extant species [1]. Orb weavers construct conspicuous and often large orb-like foraging webs, attracting the interest of evolutionary biologists [34,35]. Little is known about their venoms, and only one species (the Chinese orb-weaver *Araneus ventricosus*) has been investigated using a venomics approach [36].

We selected the wasp spider *Argiope bruennichi*, which features a wasp-like banding pattern that may have evolved as a tool to lure prey [37]. This species is used as a model organism for the investigation of sexual dimorphism, chemical ecology, reproductive behavior, microbiome analysis and range expansion linked to climate change [38–47]. Its venom has been extracted for bioactivity assays but has not been analyzed in detail [27]. We applied a cutting-edge proteotranscriptomics workflow, in which an automated multiple-assembly strategy was used as a first step to identify and annotate venom gland-specific transcripts. These were matched to proteome components detected directly using two parallel mass spectrometry (MS) platforms to achieve exhaustive and sensitive protein detection and identification. We discuss the functional components of the venom in the context of wasp spider hunting behavior, in which silk rather than a venomous bite is the primary weapon used to overpower prey [48].

2. Materials and Methods

2.1. Collection of Specimens and Sample Preparation for Transcriptomics and Proteomics

Fourteen *A. bruennichi* adult females were collected in September 2018 in Gießen, Germany (N 50.5729555°, E 8.7280508°). Initially, we tried to milk venom from the collected spiders by applying electrostimulation. However, this approach failed due to the small size of the venom apparatus and the low venom yield. Therefore, the strategy was changed, and four days after electrostimulation, whole venom glands were dissected from CO_2-anesthetized specimens under a stereomicroscope, washed in distilled water and submerged in phosphate-buffered saline (PBS). Venom was released by gentle compression with forceps, and the extracts were centrifuged (10,000 × g, 10 min, room temperature) to pellet cell debris, before pooling the supernatants for lyophilization. The remaining venom gland tissue was transferred into 1-mL RNAlater solution, pooled and stored at –80 °C. Remaining body tissue was processed in the same manner.

2.2. Proteotranscriptomics Overall Workflow

RNA-Seq was used to identify transcripts from venom glands and remaining body tissues, followed by assembly and automated annotation. Crude venom was analyzed by one-dimensional (1D) and two-dimensional (2D) polyacrylamide electrophoresis (PAGE) before two parallel bottom-up MS methods were used to identify the venom proteins. The first was matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF)-MS to characterize peptides derived from the 2D gels, and the second was liquid chromatography electrospray ionization (nanoLC-ESI)-MS to characterize peptides directly from the crude venom samples. Transcripts matching the proteins identified by MS were then analyzed in more detail by examining differential expressions and annotations. The whole workflow in the present analysis was designed to minimize the effects that were described in recent studies, which revealed that transcriptome-based approaches led to the generation of large quantities of false-positive data points and, thus, caused an overestimation of toxic diversity [49]. Strict filter steps were applied—first, only transcripts were included that were two-fold enriched compared to the body tissue. Secondly, transcripts encoding for the vast majority of nonvenom transcripts were excluded from the analysis. On another level, the applied different proteomics platforms were combined to identify the presence of venom transcripts on the protein level. Therefore, our subsequent analysis only considered transcripts encoding for proteomically detectable venom proteins as the baseline in a rather conservative analysis approach that prevented an overinterpretation of the transcriptome data (Figure 1).

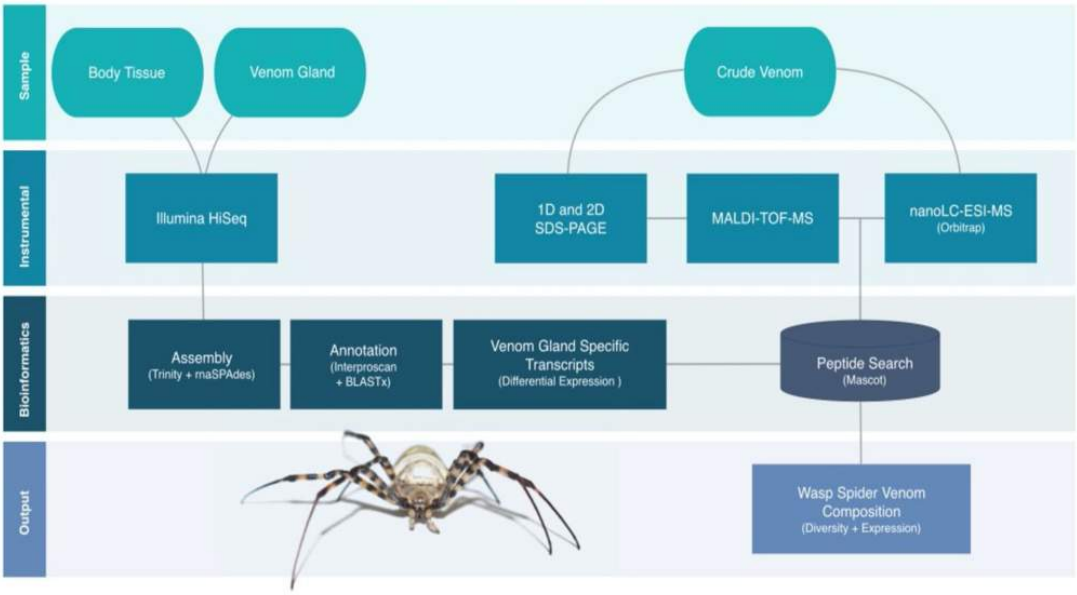


Figure 1. Proteotranscriptomics workflow to characterize the venom of *Argiope bruennichi*. Transcriptomes of venom glands and body tissue were sequenced and assembled. Crude venom was

analyzed by 1D/2D-polyacrylamide electrophoresis (PAGE) before combinatorial matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) and liquid chromatography electrospray ionization (nanoLC-ESI)-MS. The final transcriptome assembly was used for the MS peptide search. Venom-specific transcripts matching detected proteins were then investigated in terms of expression levels and annotations.

2.3. Transcriptomics of Venom Gland and Body Tissue

2.3.1. RNA Extraction and Sequencing

RNA extraction and sequencing were outsourced to Macrogen (Seoul, Korea). Following RNA extraction, libraries were constructed using the TruSeq RNA Sample Prep Kit v2 (paired-end, 151-bp read length). Quality was controlled by the verification of PCR-enriched fragment sizes on the Agilent Technologies 2100 Bioanalyzer with the DNA 1000 chip. The library quantity was determined by qPCR using the rapid library standard quantification solution and calculator (Roche). The libraries were sequenced on the Illumina Hi-Seq platform.

2.3.2. Transcriptome Assembly, Annotation and Quantification

Transcriptome data were processed using a modified version of our in-house assembly and annotation pipeline featuring different docker containers for enhanced reproducibility [25]. All containers (Table S1) were established using Nextflow v19.01.0 (https://www.nextflow.io/). Briefly, all input sequences were inspected using FastQC v0.11.7 before trimming in Trimmomatic v0.38 [50,51] using the settings 2:30:10, LEADING:5, TRAILING:5, SLIDINGWINDOW:4:15 and MINLEN:75. The trimmed reads were corrected using Rcorrector v1.0.3.1 and assembled de novo using a pipeline incorporating Trinity v2.8.4 and rnaSPAdes v3.12 with and without error correction [52–55]. All contigs were combined into a single assembly, in which transcripts from all assemblers were merged if they were identical. The reads were remapped to the assembly using Hisat2 v2.1.0, and expression values (transcripts per million, TPM) were calculated using stringtie v1.3.5 [56–58]. SAM and BAM files were converted using Samtools v1.9 [59]. Open reading frames were then predicted with Transdecoder v5.0.2 [55] and annotated at the amino acid level using Interproscan v5.35-74 and BLASTX v2.6.0+ [60,61] searches against the Swissprot, Toxprot and Arachnoserver databases [14,26]. The resulting assembly was used as a species-specific database for the identification of proteins detected by MS. Sequencing raw data are available at the SRA database (PRJNA634567).

To avoid the overinterpretation of our data, a differential expression analysis was applied to the two samples (venom gland versus remaining body tissue), and only putative venom components derived from transcripts with a logFC >2 within the venom gland dataset were considered further. Filtering steps were performed within the TBro v1.1.1 framework [62].

2.4. Venom Proteomics

2.4.1. Fractionation of Venom Proteins by PAGE

For 1D-PAGE, venom was mixed with tricine sample buffer (Bio-Rad) to make a total volume of 12 µL and incubated for 5 min at 95 °C. The sample was then loaded onto a 16.5% Mini-PROTEAN Tris-Tricine gel (Bio-Rad) in a Mini-PROTEAN Tetra System chamber (Bio-Rad) using 10x Tris-Tricine/SDS running buffer (Bio-Rad). Electrophoresis was carried out at 100 V for 100 min, and protein bands were detected with Flamingo stain (Bio-Rad).

For 2D-PAGE, contaminants were removed from the venom extract by the precipitation of 200-µg protein with 1:4 (v/v) chloroform/methanol [63]. The protein pellet was redissolved in 260-µL lysis buffer (6-M urea, 2-M thiourea, 4% CHAPS, 30-mM DTT and GE Healthcare 2% IPG buffer pH 3–10). GE Healthcare IEF strips (pH 3–10NL, 13 cm) were loaded with the sample by rehydration for 22 h, and isoelectric focusing was carried out at gradients of 0–100 V/1 mA/2 W for 5 h, 100–3500 V/ 2 mA/5 W for 6 h and 3500 V/2 mA/5 W for 6 h using a Multiphor II system (GE Healthcare). The IEF strip was then equilibrated for 15 min in 5-mL equilibration stock solution (6-M urea, 4% SDS, 0.01%

bromophenol blue, 50-mM Tris-HCl pH 8.8 and 30% (v/v) glycerol) containing 65-mM DTT and, then, for 15 min in the same solution containing 200-mM iodacetamide. Proteins were separated in the second dimension on a 14% SDS polyacrylamide gel [64] in a Hoefer600 cell (GE Healthcare) for 15 min at 15 mA (100 V/15 W limits) and 4 h at 150 mA (400 V/60 W limits). The proteins were detected with Flamingo stain (Bio-Rad).

2.4.2. MALDI-TOF-MS

The 2D gel was analyzed using PDQuest (Bio-Rad, CA, USA), and 152 spots (Figure S1) were excised using the ExQuest Spot Cutter (Bio-Rad) and transferred into 96-well plates (Greiner Bio-One). The samples were digested simultaneously by using a MicroStarlet pipetting robot (Hamilton Robotics, NV, USA) to execute the following steps: the excised gel plugs were destained with 25-mM ammonium hydrogen carbonate containing 50% (v/v) acetonitrile, dehydrated with 100% acetonitrile, rehydrated in 50-mM ammonium hydrogen carbonate, dehydrated with 100% acetonitrile, dried at 56 °C, rehydrated with 17-µL 25-mM ammonium hydrogen carbonate containing 4.5-ng/µL sequencing grade trypsin (Promega) and 0.025% Proteasemax (Promega) and incubated at 45 °C for 2 h. Peptides were recovered by extraction with 15-µL 1% trifluoroacetic acid (Applied Biosystems) and stored at 4 °C.

MALDI-TOF-MS was performed on an Ultraflex I TOF/TOF mass spectrometer (Bruker Daltonics) equipped with a nitrogen laser and a LIFT-MS/MS facility. Summed spectra consisting of 200–400 individual spectra were acquired in positive ion reflectron mode using 5-mg/mL 2,5-dihydroxybenzoic acid (Sigma-Aldrich) and 5-mg/mL methylendiphosphonic acid (Fluka) in 0.1% trifluoroacetic acid as the matrix. For data processing and instrument control, we used the Compass v1.4 software package consisting of FlexControl v3.4, FlexAnalysis v3.4 and BioTools v3.2. Data storage and database searches were carried out using ProteinScape v3.1 (Bruker Daltonics, MA, USA). Proteins were identified by Mascot v2.6.2 (Matrix Science, United Kingdom) peptide mass fingerprinting using the venom gland transcriptome as a database. The search was restricted to peptides larger than 10 amino acids with a mass tolerance of 75 ppm. Carbamidomethylation of cysteine was considered as a global modification, the oxidation of methionine was considered as a variable modification and one missed cleavage site was allowed. Only peptides with a Mascot Score > 80 were considered for further analysis (Table S2). The proteomic raw data are available at PRIDE (PXD018693).

2.4.3. NanoLC-ESI-MS

We dissolved 10 µg of protein in 25-mM ammonium bicarbonate containing 0.6-nM ProteasMax™. Cysteines were reduced with 5-mM DTT for 30 min at 50 °C and modified with 10-mM iodacetamide for 30 min at 24 °C. The reaction was quenched with an excess of cysteine, and the protein was digested with trypsin at a 50:1 ratio for 16 h at 37 °C. The reaction was stopped by addition trifluoroacetic acid to a final concentration of 1%. The sample was then purified using a C18-ZipTip (Millipore), dried under vacuum and redissolved in 10-µL 0.1% trifluoroacetic acid.

For analysis, 1 µg of the sample was loaded onto a 50-cm µPAC C18 column (Pharma Fluidics) in 0.1% formic acid at 35 °C. Peptides were eluted with a 3–44% linear gradient of acetonitrile over 240 min, followed by washing with 72% acetonitrile at a constant flow rate of 300 nl/min using a Thermo Fisher Scientific UltiMate 3000RSLCnano device (MA, USA). Eluted samples were injected into an Orbitrap Eclipse Tribrid MS (Thermo Fisher Scientific, MA, USA) in positive ionization mode via an Advion TriVersa NanoMate (Advion BioSciences, NY, USA) with a spray voltage of 1.5 kV and a source temperature at 250 °C. Using the data-independent acquisition mode, full MS scans were acquired every 3 s over a mass range of *m*/*z* 375–1500 with a resolution of 120,000 and auto-gain control (AGC) set to standard with a maximum injection time of 50 ms. In each cycle, the most intense ions (charge states 2–7) above a threshold ion count of 50,000 were selected with an isolation window of 1.6 *m*/*z* for higher-energy collisional (HCD) dissociation at a normalized collision energy of 30%. Fragment ion spectra were acquired in the linear ion trap with the scan rate set to rapid, a normal

mass range and a maximum injection time of 100 ms. Following fragmentation, selected precursor ions were excluded for 15 s.

Data were acquired with Xcalibur v4.3.73.11. (Thermo Fisher Scientific, MA, USA) and analyzed using Proteome Discoverer v2.4.0.305 (Thermo Fisher Scientific, MA, USA). Mascot v2.6.2 was used to search against the transcriptome database. A precursor ion mass tolerance of 10 ppm was applied. Carbamidomethylation of cysteine was considered as a global modification, the oxidation of methionine was considered as a variable modification and one missed cleavage site was allowed. Fragment ion mass tolerance was set to 0.8 Da for the linear ion trap MS^2 detection. The false discovery rate (FDR) for peptide identification was limited to 0.01 using a decoy database. For subsequent analysis, we only considered proteins with a Mascot Score > 30 and at least two verified peptides (Table S3 and Table S4). The raw proteomic raw data are available at PRIDE (PXD018693).

2.5. Reanalysis of Araneus Ventricosus Venom

In order to increase the comparability of our results regarding the general conclusions on venoms of Araneidae, we reanalyzed the available proteomic data for only the in-detail studied araneid taxon *Araneus ventricosus*. Since the original work did not perform any assignments to protein classes, all known proteins of *A. ventricosus* were retrieved from the Arachnoserver database [14] and assigned to protein classes by Interproscan [61] (Table S5). The resulting sequences and protein diversity were assessed manually; however, expression levels could not be included, because the original work did not include quantitative data.

3. Results

3.1. The A. bruennichi Venom Gland and Body Tissue Yield High-quality Transcriptome Libraries

Venom glands were dissected from 14 female *A. bruennichi* specimens, and the venom was extracted and set aside for proteomic analysis. The venom glands and remaining body tissues were separately pooled, and RNA was extracted for RNA-Seq analysis. The resulting paired-end libraries were checked for DNA quantity and quality. The concentration of the venom gland transcriptome library was 116.26 ng/μL (fragment size = 387 bp), and the concentration of the remaining body tissue library was 91.47 ng/μL (fragment size = 363 bp). The venom gland transcriptome contained a total of 133,263,138 paired-end reads, with a GC content of 42.2%, a Q20 of 98.2% and a Q30 of 94.5%. The remaining body tissue transcriptome contained a total of 145,808,360 paired-end reads, with a GC content of 41.6%, a Q20 of 98.3% and a Q30 of 94.8%. The libraries were sequenced and annotated using our automated pipeline.

3.2. Only Large A. bruennichi Venom Proteins are Detected by SDS-PAGE and MALDI-TOF-MS

The crude venom set aside prior to RNA extraction was first fractionated by 1D-PAGE. The lanes representing both concentrations showed identical banding patterns after staining, and the vast majority of the protein bands were in the 25–100 kDa range, with a few weaker bands of 15–25 kDa and no prominent bands below 15 kDa (Figure 2). To characterize the properties of these proteins in more detail, the venom sample was fractionated by 2D-PAGE. The isoelectric focusing steps (pH 3–10) revealed proteins with a range of pI values, although predominantly focused around pH 7 (Figure 2). In agreement with the 1D-PAGE results, the orthogonal SDS-PAGE step indicated that most spots represented proteins of 25 kDa or more, with only a few in the size range 10–25 kDa and hardly any below 10 kDa (Figure 2).

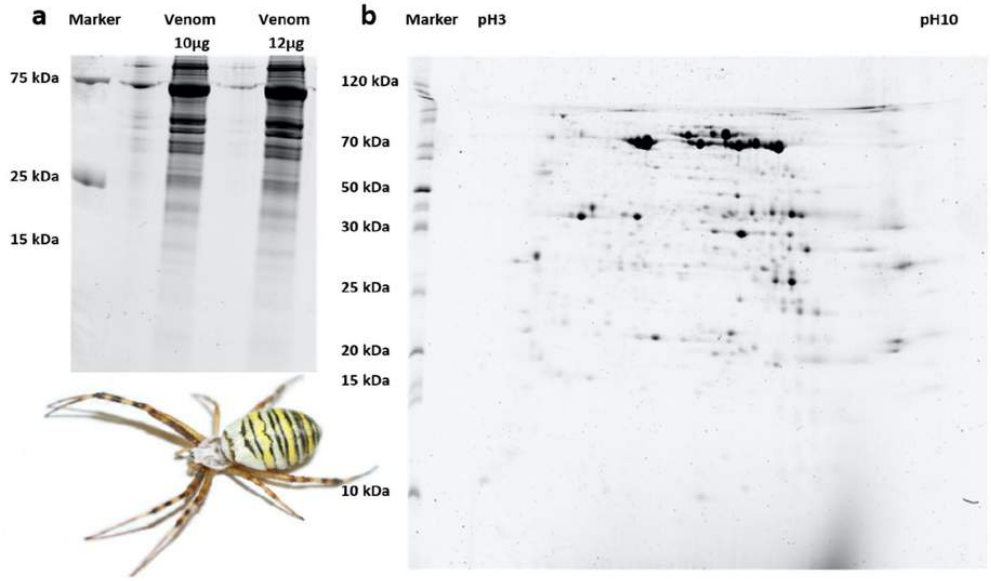


Figure 2. Analysis of *A. bruennichi* venom proteins by PAGE. (**a**) 1D-PAGE of venom proteins at two concentrations, showing identical banding patterns, with most proteins larger than 25 kDa. (**b**) 2D-PAGE, showing that the proteins cover a range of pI values but cluster around pH 7 and confirming that most proteins are larger than 25 kDa.

We excised 152 spots from the 2D gels for MALDI-TOF-MS analysis, 41 of which matched to significantly enriched predicted coding regions from our venom gland transcriptome, while the remainder matched nonvenom proteins. Among the 41 venom-related spots, only six were ultimately assigned to protein classes known to be present in other animal venoms (Table 1). The sequences of the venom proteins identified in *A. bruennichi* were highly similar to the toxins previously identified in its close relative, the Chinese orb-weaver *A. ventricosus* [36]. Their molecular masses fell with the range 28.3–50.5 kDa, and functional annotation revealed that they all belong to the cysteine-rich secretory protein, antigen 5 and pathogenesis-related protein 1 (CAP) superfamily.

Table 1. Identification of *Argiope bruennichi* venom proteins by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Among 152 spots excised from 2D gels, 41 represented proteins that were enriched in the venom glands, and six of these were similar to previously identified venom components, all putative members of the CAP superfamily.

Spot ID	Class	Score	kDa	Peptides	Coverage (%)	ppm
8501	CAP	136.0	48.9	14	21.1	13.81
7309	CAP	70.7	28.3	6	18.3	7.42
7501	CAP	182.0	50.0	18	35.3	23.55
6502	CAP	85.9	50.5	10	24.8	22.95
6502	CAP	124.0	45.7	15	33.7	23.73
6502	CAP	89.2	48.1	11	25.3	23.57

CAP = cysteine-rich secretory protein, antigen 5 and pathogenesis-related protein 1.

3.3. Further A. bruennichi Venom Proteins are Revealed by High-Resolution NanoLC-ESI-MS

In a parallel proteomics workflow, the crude venom was analyzed by high-resolution nanoLC-ESI-MS (Orbitrap), revealing a total of 1806 protein groups, including 415 predicted coding regions matching significantly enriched predicted coding regions from our venom gland transcriptome. In size, protein groups ranged from 2 kDa to 950 kDa (Figure 3). The majority of the identified protein groups within the wasp spider venom are composed of isoforms between 10 kDa and 100 kDa; only a marginal fraction of the protein groups is sized below 10 kDa, over 100 kDa or even over 200 kDa.

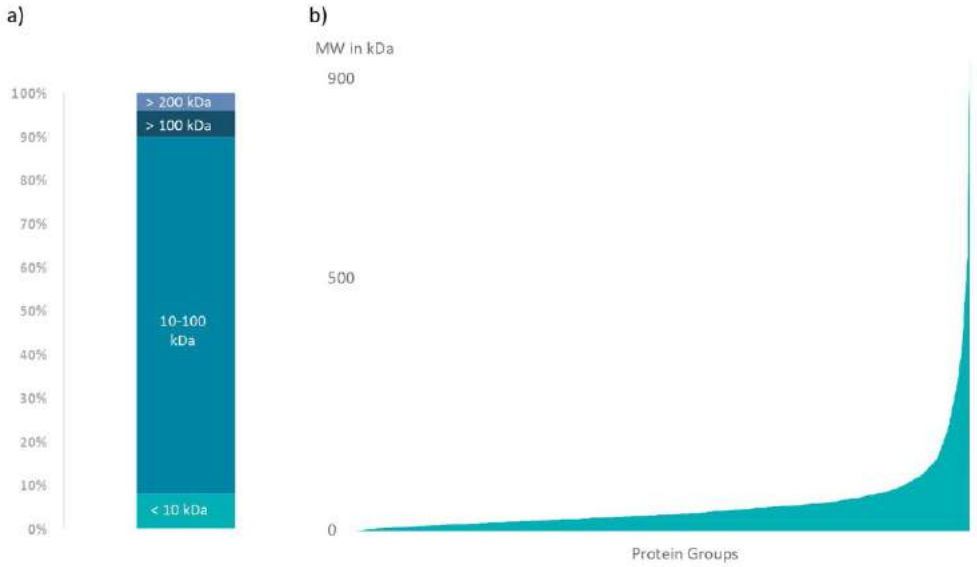


Figure 3. Size distribution within the *A. bruennichi* venom proteome. (**a**) Relative distribution of the identified protein groups within the proteome. Small protein groups of <10 kDa account for 8%, and protein groups between 10 kDa and 100 kDa account for 82%, while protein groups larger than 100 kDa and larger than 200 kDa contribute 6% and 4%, respectively, to the proteomic dataset. (**b**) Absolute size distribution of identified protein groups in kDa.

From identified protein groups, we retrieved 54 protein groups with putative venom functions, representing 20 different protein families. Many of these protein families have previously been identified in spider venoms, including Kunitz-type serine protease inhibitors, prokineticin, EF-hand proteins, MIT-atracotoxins, astacin-like metalloproteases and ICK peptides. Three others showed similarities to hormones and neuropeptides (insulin-like growth factor binding protein (IGFBP), diuretic hormone (DH) and ITG-like peptides). Another class of proteins showed a high sequence similarity to uncharacterized toxins previously isolated from *A. ventricosus* [36]. BLAST searches did not recover any further similar sequences, so the remaining proteins were defined as "unidentified aranetoxins". The nanoLC-ESI-MS experiment also confirmed our MALDI-TOF-MS data by showing that *A. bruennichi* venom contains multiple CAP proteins that are also the most abundant proteins among the venom components (Table 2).

Table 2. Identification of *A. bruennichi* venom proteins by liquid chromatography electrospray ionization (nanoLC-ESI)-MS. The analysis of peptide fragments allowed us to identify protein groups with putative venom functions, representing 20 different protein families. Confirming the parallel MALDI-TOF-MS analysis, most proteins could be assigned to the CAP superfamily. ICK = inhibitor cysteine knot and IGFBP = insulin-like growth factor binding protein.

Protein Class	Matched Peptides	MW (kDa)	Calc. pI	Mascot Score	Coverage (%)
5' Nucleotidase	4	24.3	4.54	228	17
Astacin-like metalloprotease	2	10.9	4.79	149	15
Astacin-like metalloprotease	2	14.4	5.87	76	13
Astacin-like metalloprotease	3	17.0	7.93	50	24
CAP	23	51.2	7.77	3697	57
CAP	18	50.8	8.19	2333	52
CAP	14	50.0	7.65	2164	40
CAP	9	28.3	7.66	1918	59
CAP	16	48.9	8.44	1904	46
CAP	8	50.5	7.97	1661	26
CAP	15	51.0	8.50	1127	40
CAP	14	48.0	8.07	1020	46
CAP	11	28.9	9.09	944	44

CAP	5	17.0	5.14	604	32
CAP	6	13.2	8.79	311	49
CAP	3	25.1	6.84	173	27
Cystatin	2	15.6	7.33	48	22
Diuretic hormone-like	2	14.6	9.55	267	23
EF-hand	7	22.3	5.25	112	33
ICK	8	15.0	6.15	1076	40
ICK	2	14.7	4.68	209	14
ICK	4	15.7	6.76	59	24
IGFBP	2	18.8	4.92	63	16
ITG-like peptide	10	26.6	4.78	1666	51
ITG-like peptide	8	24.7	4.96	673	38
Kunitz	2	25.1	7.46	38	8
Leucine-rich-repeat domain	20	39.7	4.93	4279	71
Leucine-rich-repeat domain	10	41.0	5.29	771	34
Leucine-rich-repeat domain	9	36.9	5.74	428	37
Leucine-rich-repeat domain	7	39.2	5.11	215	27
Leucine-rich-repeat domain	5	41.2	5.50	117	23
MIT-atracotoxin	4	10.7	4,94	448	56
MIT-atracotoxin	5	9.8	5.50	140	66
Prokineticin	4	13.8	7.97	636	33
Putative chitinase	12	35.5	7.30	1159	45
Putative chitinase	8	30.1	6.49	219	39
Putative chitinase	2	18.0	5.19	112	11
Putative chitinase	3	47.4	11.15	53	13
S10 peptidase	4	51.4	8.07	118	15
Serine protease	8	53.2	6.40	608	24
Serine protease	8	53.2	6.64	469	23
Serine protease	5	86.2	6.54	135	9
Serine protease	3	99.1	6.27	100	4
Serine protease	3	55.2	6.13	54	7
Techylectin	3	40.3	7.17	88	7
Thyroglobulin-like	3	10.9	7.56	134	31
Unclassified aranetoxins	3	8.3	8.07	630	55
Unclassified aranetoxins	5	12.9	9.57	220	32
Unclassified aranetoxins	4	8.2	8.18	206	39
Unclassified aranetoxins	4	8.3	8.18	202	39
Unclassified aranetoxins	2	8.4	7.71	56	39
Venom protein 11	2	9.5	8.00	234	35

3.4. Data Integration Reveals that A. bruennichi Venom is Atypical for Spiders

The transcriptomic and proteomic data were integrated for the comprehensive analysis of venom composition in terms of the diversity and abundance (TPM) of venom proteins (Table S4). In terms of overall diversity, the CAP superfamily was the most represented, with 15 different CAP proteins accounting for more than 27% of all the identified protein components. Leucine-rich repeat proteins and unclassified aranetoxins were also well represented, each with five members, accounting for ~9% of the total diversity. We identified four putative chitinases and serine proteases, each accounting for ~7.5% of the total diversity, and three ICKs and astacin-like metalloproteases, each accounting for ~5.5% of the total diversity. There were two members of the MIT-atracotoxin and ITG-like peptide families, each contributing ~3.5% of the total diversity. Finally, the Kunitz serine protease inhibitor, cystatin, diuretic hormone-like peptide, techylectin, IGFBP, venom protein 11, 5' nucleotidase, prokineticin, thyroglobulin, S10 peptidase and EF-hand families were represented by one member, each contributing <2% to the total diversity (Figure 4). In terms of abundance, CAPs represented 64.3% of the total protein content of *A. bruennichi* venom and were by far the most dominant component, followed by ITG-like peptides (9.5%), unclassified aranetoxins (7.7%) and leucine-rich repeat proteins (7.7%). The other components were expressed at much lower levels, with ICKs contributing only 3.3% of the total protein content, followed by putative chitinases (2.6%) and serine proteases (2.7%) and the others each contributing < 1% (Figure 4).

In agreement with the 1D/2D-PAGE experiments, higher-molecular-weight proteins accounted for most of the diversity of the *A. bruennichi* venom proteome and were also the most abundant components. We identified 23 proteins/peptides with molecular masses < 20 kDa, but these accounted for only ~42% of the proteome diversity and only ~13% of the total protein content (Figure 4).

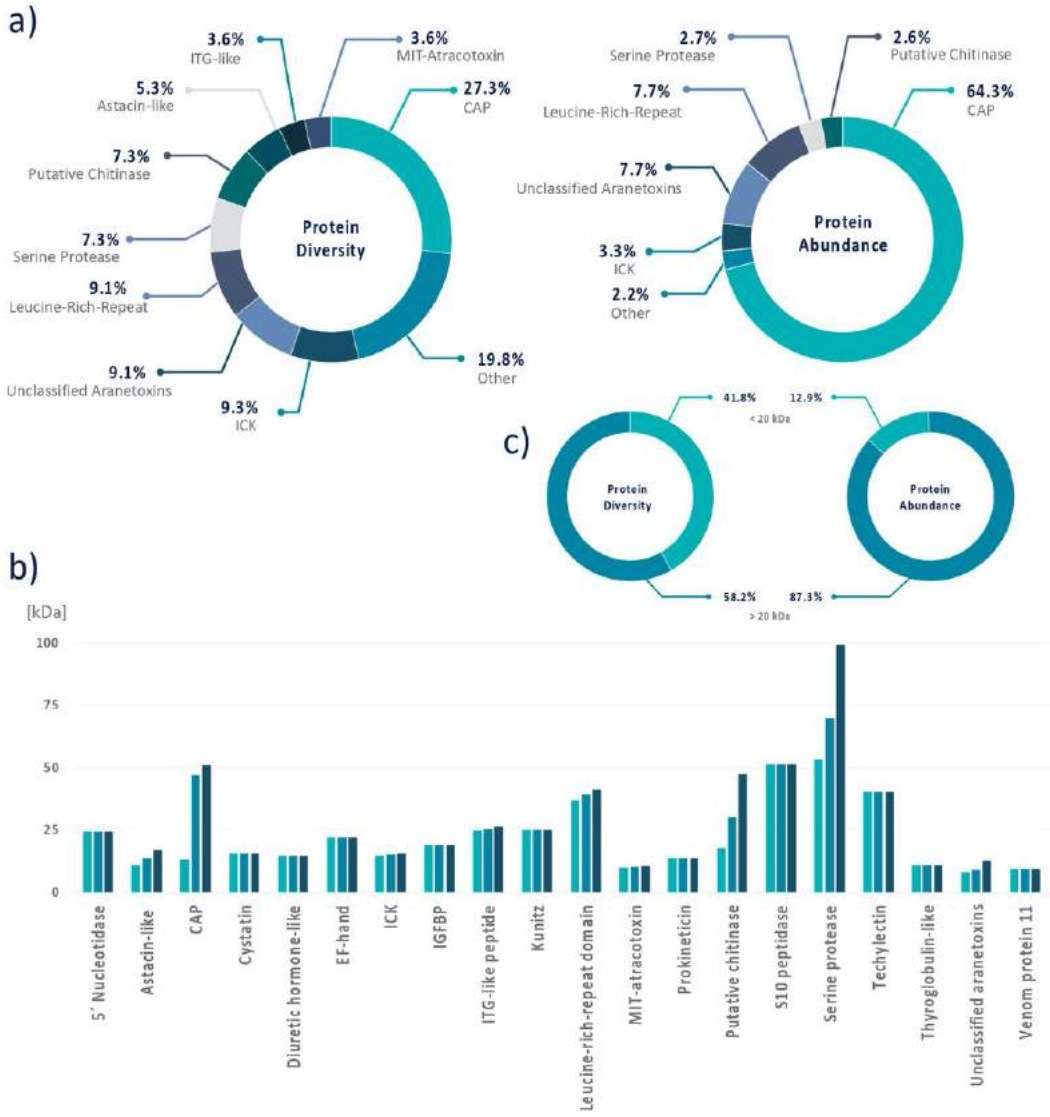


Figure 4. The venom protein profile of *A. bruennichi*. (**a**) Pie charts depict the venom composition in terms of protein diversity based on the number of distinct predicted coding sequences compared to protein abundance based on the transcripts per million reads for each coding sequence. By both measures, CAP family proteins are the dominant venom component, with 15 different members, many expressed at high levels. (**b**) The molecular weight (kDa) of identified venom proteins, with the lowest, average and highest molecular weights per group from left to right. (**c**) The distribution of small (< 20 kDa) and large (> 20 kDa) proteins in terms of protein diversity and protein abundance (TPM).

3.5. The Venom of Araneus Ventricosus Has Similarity to Argiope bruennichi Venom

The reanalysis of the *A. ventricosus* venom recovered 61 proteins that could be assigned to known protein families, which resembled only a third of its venom components. The remaining

uncharacterized proteins mostly comprised short peptides that were currently assigned as diverse but unknown U-aranetoxins, some of these having counterparts in *A. bruennichi*. However, without any available functional information, they cannot be further discussed.

Proteins of known families include Kunitz serine protease inhibitors (1.6%); MIT-atracotoxins (3.2%); serine proteases (3.2%); ICKs (CSTX, conotoxin and huwentoxin-1 families) (14.5%); thyroglobulins (46.7%) and CAPs (22.6%). Albeit thyroglobulins represent the most diverse protein class in *A. ventricosus* venom and several of the components identified in *A. bruennichi* venom are absent, both venoms share some similarities. Most identified proteins reflect a rather similar venom diversity (Figure 5). This includes CAP proteins identified in *A. ventricosus*, although they are the second-most diverse proteins in its venom. Similar to *A. bruennichi*, many of the identified proteins from *A. ventricosus* represent rather large classes, exceeding 20 kDa.

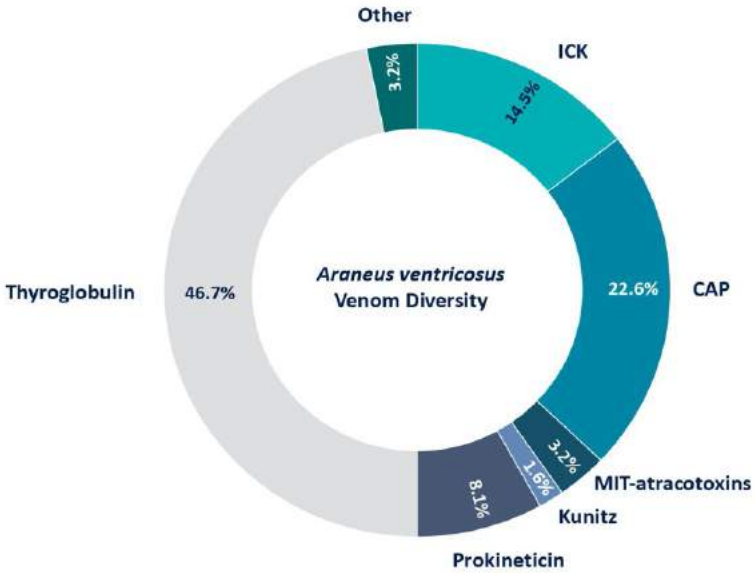


Figure 5. The venom protein profile of *Araneus ventricosus* derived from the reanalysis of the original study [36]. Given are identified proteins that could be annotated via Interproscan in percentages of the total diversity. From the 62 proteins within the dataset, most were assigned as thyroglobulins and, similar to *A. bruennichi*, CAPs.

4. Discussion

Spider venoms typically consist of mostly low-molecular-weight peptides, with ICK peptides as the predominant neurotoxic components. For example, ICK peptides represent 93% of the diversity in *Phoneutria nigriventer* venom [65] and 42 of 46 identified venom components in the barychelid *Trittame loki* [66]. In *Cupiennius salei*, short cationic peptides and ICK peptides together comprise 39% of the venom components, whereas larger proteins only contribute 15% to its diversity [67]. The predominance of ICK peptides has also been reported in venom isolated from *Cyriopagopus hainanus* (formerly *Haplopelma hainanum*), *Selenocosmia jiafu*, *Lycosa singoriensis* and *Pamphobeteus verdolaga* [68–71]. The general assumption is that ICK peptides are highly diverse components of spider venom, and dozens of different peptides may be present per species [13]. In contrast, we found that ICK peptides were only a minor component of the *A. bruennichi* venom, with only three different peptides identified (~5.5% of the overall diversity) and a low abundance (only 3.3% of the total content based on TPM counts), suggesting a less important role in wasp spider venom compared to other spiders. Instead, we found that CAP superfamily proteins were both the most diverse (15 different members, >27% of the overall diversity) and the most abundant (>64% of the total content based on TPM

counts), suggesting these proteins are particularly important for the function of *A. bruennichi* venom. Given that *A. ventricosus* venom also contains several CAP proteins (accounting for 22.6% of the venom) (Figure 5) [36], we speculate that CAP proteins may be generally important for venom functions in orb-weaving spiders. Unfortunately, we cannot compare our results directly to this previous study, because it was based solely on a proteomic analysis and lacked quantitative data. This is, in particular, of importance, as the simplicity of *A. bruennichi* venom has been recovered by our proteotranscriptomics approach on the level of protein abundance (quantified in TPM) (Figure 4). The extent to which our results regarding the wasp spider venom are to generalize for Araneidae should be a subject for future investigations. However, the atypical nature of *A. bruennichi* venom allows us to develop a functional hypothesis based on the ecology of wasp spiders, focusing on the most dominant venom components.

4.1. The Importance of CAP Superfamily Proteins in Wasp Spider Venom

The CAP superfamily is one of several protein groups that have undergone convergent recruitment and neofunctionalization in venom systems, and CAP proteins have therefore been isolated from the venoms of snakes, spiders, cone snails, scorpions, fish, cephalopods and a variety of insects [5]. This taxonomic ubiquity reflects the ability of CAP proteins to adopt diverse functions. For example, CAP proteins in snake venoms act as neurotoxins by interacting with ion channels [5,72,73], whereas CAP proteins in the venoms of lampreys, hematophagous insects and ticks are thought to facilitate feeding [5,74,75]. In bees, wasps and ants, CAP proteins are major allergenic components of venom and are therefore associated with inflammation and potentially fatal anaphylaxis [76,77]. CAP proteins have been detected in several spider venoms but, generally, as minor components, and their function is unknown [66,67]. Thus far, the only known spider with a venom dominated by CAP proteins is *A. bruennichi*.

The CAP proteins in *A. bruennichi* venom, as in other arthropods, are unlikely to act as neurotoxins, because they lack the C-terminal cysteine-rich domain that confers the neurotoxicity of CAP proteins in snake venom [5]. Phylogenetic reconstructions indicate that spider CAP proteins are similar to Tex31, a well-characterized CAP protein from the venom of the cone snail *Conus textile* [5] that has proteolytic activity [78]. This suggests that CAP proteins in *A. bruennichi* venom (and other spider venoms) may support extra-oral digestion, toxin maturation or act as spreading factors to promote the uptake of other venom components. The lack of CAP superfamily neurotoxins in spiders would not be a disadvantage, because the venom contains other neurotoxic components, including ICK peptides, prokineticins and MIT-atracotoxins. Assuming that the newly identified aranetoxins and neuropeptides also act as neurotoxins, wasp spider venom clearly contains an impressive arsenal of bioactive components that may facilitate hunting. Interestingly, an early study on the effects of spider venom against cockroaches and meal beetles demonstrated that wasp spider venom can paralyze but not kill both these prey [27]. Therefore, despite the atypical composition of *A. bruennichi* venom, this species is nevertheless capable of neurotoxic envenomation.

4.2. Wasp spider Venom Contains Potential New Toxin Classes Similar to Arthropod Neuropeptides

Our proteomic analysis of *A. bruennichi* venom identified five polypeptides that we grouped as unclassified aranetoxins, showing a high sequence similarity to the recently discovered U_6 and U_8 aranetoxins in *A. ventricosus* [36]. They are not yet formally assigned to any known class of toxins, and their molecular and biological functions remain to be determined. However, the five unclassified aranetoxins were expressed at high levels in our venom gland transcriptome dataset and are therefore likely to fulfill important functions in the *A. bruennichi* venom system. Their presence in two orb weavers but no other spider families suggests their role may be specific to the unique ecological niche of orb weavers.

We also identified one diuretic hormone-like peptide, one IGFBP and two ITG-like peptides. Diuretic hormone-like peptides contain a DH31-like domain and are related to a diuretic hormone from the Florida carpenter ant (*Camponotus floridanus*) and, to a lesser extent, U-scoloptoxin-Sm2a from the centipede *Scolopendra morsitans* [79]. This class of protein has not been detected in other

spider venoms. The *A. bruennichi* IGFBP is closely related to a protein found in the venom of the tiger wandering spider *Cupiennius salei* [80]. Such proteins are commonly found in arachnid venoms but also in scorpions (*Superstitia donensis*, *Hadrurus spadix* and *Centruroides hentzi*) and ticks of the genus *Amblyomma* [81–84]. A related protein is encoded in the genome of *A. ventricosus*, but its function has not yet been determined [85]. Whereas the *A. bruennichi* diuretic hormone-like peptide and insulin-like growth factor-binding protein are relatively minor venom components, the two ITG-like peptides were expressed at high levels in the venom gland transcriptome and may therefore fulfil more important functions. They are closely related to peptides found in the black cutworm moth (*Agrotis ipsilon*) and *C. floridanus* but have not been identified in other spider venoms.

The role of all three classes of proteins described above is unclear, but hormones in other venomous animals have been weaponized as toxins to subdue prey. Such neofunctionalization might occur when hormones recruited to the venom gland for normal physiological activity undergo mutations that affect their surface chemistry and potential for functional interactions. For example, a neuropeptide that regulates physiological processes in the predator could become a toxin if a mutation causes it to interact with a different receptor in a prey species following envenomation. If this process occurs in the context of gene duplication and divergence, the new role in envenomation could be unlinked from the original physiological role, allowing evolutionary forces to fix the neuropeptide as a venom toxin. The neofunctionalization of hormones and neuropeptides in venom systems is further highlighted by the recent discovery of the convergent recruitment of hyperglycemic hormones in the venom of spiders and centipedes [79]. This study demonstrated that helical arthropod-neuropeptide-derived (HAND) toxins are derived from hormones of the ion transport peptide/crustacean hyperglycemic hormone (ITP/CHH) family, which are ubiquitous and functionally diverse neuropeptides in arthropods. ITP/CHH peptides have also been recruited into the venom systems of ticks and wasps and are not restricted to the HAND family. For example, emerald jewel wasp (*Ampulex compressa*) venom contains tachykinin and corazonin neuropeptides that induce hypokinesia in cockroaches [86], whereas exendin from the venom of helodermatid lizards is a modified glucagon-like peptide that interferes with pancreatic insulin release [87]. Amphibians have also recruited a variety of hormone peptides as skin toxins [88–91]. The novel neuropeptides in *A. bruennichi* might fulfill other functions, and their potential role as toxins must be tested, but the strong expression of the ITG-like peptides indicates an important function in the venom system.

4.3. The Potential Ecological Role of Atypical Wasp Spider Venom

The atypical venom composition of *A. bruennichi* could be explained by trophic specialization, which would select for simple venoms prioritizing specific components needed to subdue selected prey species. This would contrast with generalist predators, where diverse venom components would confer a selective advantage [92–96]. However, *A. bruennichi* is not regarded as a specialist feeder, and an alternative explanation must be sought [97].

In a pioneering study, the hunting behaviors of three orb weavers (*Nephila claviceps*, *Argiope aurantia* and *Argiope argentata*) were compared using bombardier beetles (*Brachinus* spp.) as prey [48]. These beetles have evolved a unique chemical defense involving the stress-triggered release of phenolic compounds from abdominal glands under high pressure. The phenolic compounds undergo rapid exothermal oxidation to benzoquinones, thus spraying predators with a pressurized discharge at temperatures up to 100 °C, allowing the beetle to escape from most situations [98]. When such beetles were offered to the three spiders, the interactions were distinct: *N. claviceps* always tried to inject venom as a first-attack strategy, and this resulted in a successful defense and escape by the beetles, whereas both *Argiope* species deployed silk as a first-attack strategy, and envenomation would follow only when the beetle was fully covered and unable to move [48]. In another study using *A. bruennichi* as the model predator, silk was deployed to overpower most prey insects, including those with other robust defense systems such as wasps, and only lepidopteran prey were attacked with venom first [99]. Similar findings have been reported for eight other *Argiope* species, suggesting that this specialized hunting behavior is highly conserved within the genus [99–101]. The prevalence

of this silk-based hunting strategy may help to explain the simplicity of its venom, which would be under selection solely for its ability to subdue lepidopteran prey.

In venomous animals, each toxin is a valuable resource that contributes to its fitness by facilitating predation, but this advantage must be balanced against the metabolic costs of replenishing venom stocks [102–104]. Many venomous animals have evolved as trophic specialists to reduce these costs, and some even produce different venoms for different purposes [105]. Spiders face a similar dilemma, because they possess two potentially competing systems to subdue prey, namely their venom and silk glands. In both cases, protein resources are deployed as a means to facilitate predation, and in both cases, the glands must be replenished at a significant metabolic cost [106]. We propose that the simplicity of *A. bruennichi* venom may reflect the evolutionary consequences of the competition for resources between the venom and silk systems, which have driven its behavioral specialization to use different strategies against different prey species. Intriguingly, the "silk-first" strategy provides the wasp spider with the competitive advantage of a high success rate against even well-defended prey [48,99], potentially contributing to its unprecedented success during its recent range expansion [42,43].

5. Conclusions

Our detailed analysis of *A. bruennichi* venom identified potential new classes of toxins and potential new roles for known protein families, including the predominant CAP superfamily. A comparison to the venom of *A. ventricosus* revealed also that other araneid spiders harbor many CAP proteins in their venom, and thus, these proteins may represent group-specific key components for orb-weaver venoms. The molecular functions and biological roles of these proteins should be investigated in detail to disentangle the venom biology of wasp spiders and their relatives and to identify new drug leads. The sequences we identified can be used to produce recombinant *A. bruennichi* venom proteins in larger quantities for detailed functional analyses, and that work is already underway in our laboratory. However, wasp spiders are small animals with limited venom yields and are therefore unsuitable for traditional fractionation workflows [29,32]. Our venomics workflow overcomes this issue by combining the data-driven selection of interesting candidates based on venom gland transcriptome analysis with a dual proteomics strategy for the comprehensive identification of venom proteins directly. Importantly, our highly sensitive venomics workflow means that a comprehensive venom composition is possible starting with only 14 spiders. A similar approach was recently used to analyze the venom proteome of the spider family Pholcidae [107].

Such novel workflows and technical platforms will help to extend our knowledge of venom compositions beyond the small collection of amenable organisms with readily accessible venom systems. This has already shown that many commonly held assumptions about venom (based on the limited number of investigated species) are not supported when more diverse species are included. We describe the venom of *A. bruennichi* as atypical for spiders, because its composition, dominated by CAP superfamily proteins and with ICK peptides fulfilling a minor role, differs from the restricted range of spider venoms that have been investigated thus far. Similarly, the recent proteomic analysis of pholcid venom also revealed a distinct composition dominated by neprilysin metalloproteases [107]. Both studies highlight the importance of filling the taxonomic gaps in venom research in order to fully understand the hidden molecular diversity. This is likely to reveal there is no "typical" spider venom but, rather, a spectrum of compositions reflecting different ecological niches. Such diversity will not only illuminate the field of arachnid evolutionary biology but will also provide many more promising candidates for translational research.

Supplementary Materials: The following are available online at https://zenodo.org/record/3922250#.XvoO0C1XY0o: Figure S1: 2D-SDS PAGE of wasp spider venom with excised spots (numbered). Table S1: Docker images applied for venom gland transcriptomics. Table S2: Proteomic results for MALDI-TOF-MS. Table S3: Proteomic results for nanoLC-ESI-MS. Table S4: Annotation of venom components identified in proteomic experiments. Table S5: Annotation of *Araneus ventricosus* venom components.

Author Contributions: T.L., S.L., A.B., B.M.v.R. and A.V. designed the study. T.L. performed the fieldwork and generated the sample material for transcriptome sequencing and venom proteomics. T.L., S.L. and T.T. performed the laboratory proteomics work. T.T. and G.L. generated and analyzed the mass spectra. T.L., F.F., B.M.v.R. and A.B. generated and analyzed the transcriptome datasets. A.V. attracted the funding for the study. T.L., B.M.v.R. and S.L. wrote the manuscript with substantial input from A.B., F.F., G.L., T.T. and A.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Hessen State Ministry of Higher Education, Research and the Arts (HMWK) for the project "Animal Venomics" via the LOEWE Center "Translational Biodiversity Genomics".

Acknowledgments: The authors are grateful to Philipp Heise, Lea Talmann and Anne Paas for their support in the field and laboratory work. The work has been conducted in the Animal Venomics Research Group at the JLU Gießen and the Fraunhofer IME-BR, the authors express their gratitude to all members of the who supported the study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Spider Catalog, Version 21. Available online: http://wsc.nmbe.ch, (accessed on 16 February 2020).
2. Garrison, N.L.; Rodriguez, J.; Agnarsson, I.; Coddington, J.A.; Griswold, C.E.; Hamilton, C.A.; Hedin, M.; Kocot, K.M.; Ledford, J.M.; Bond, J.E. Spider phylogenomics: Untangling the spider tree of life. *PeerJ.*; **2016**, *doi: 10.7717/peerj.1719*.
3. Saez, N.J.; Senff, S.; Jensen, J.E.; Er, S.Y.; Herzig, V.; Rash, L.D.; King, G.F. Spider-venom peptides as therapeutics. *Toxins*, **2010**, *2, 2851-2871*.
4. Riechert, S.E. Thoughts on the Ecological Significance of Spiders. *BioScience*, **1974**, *24*, 352–356.
5. Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.A.; King, G.F.; Nevaleinen, T.J.; Norman, J.A.; Lewis, R.J.; Norton, R.S.; et. al. The Toxicogenomic Multiverse: Convergent Recruitment of Proteins into Animal Venoms. *Annu. Rev. Genomics Hum. Genet.*, **2009**, *10*, 483–511.
6. Nelsen, D.R.; Nisani, Z.; Cooper, A.M.; Fox, G.A.; Gren, E.C.K.; Corbit, A.G.; Hayes, W.K. Poisons, toxungens, and venoms: Redefining and classifying toxic biological secretions and the organisms that employ them. *Biol Rev Camb Philos Soc*, **2014**, *89*, 450–465.
7. Holford, M.; Daly, M.; King, G.F.; Norton, R.S. Venoms to the rescue. *Science*, **2018**, *361, 842–844*.
8. Duterte, S.; Lewis, R.J. Use of Venom Peptides to Probe Ion Channel Structure and Function. *J. Bio. Chem.*, **2010**, *285, 13315–13320*.
9. Herzig, V.; Cristofori-Armstrong, B.; Israel, M.R.; Nixon, S.A.; Vetter, I.; King, G.F. Animal toxins – Nature's evolutionary refined toolkit for basic research and drug discovery. *Biochem. Pharmacol.* ,**2020**, *114096*, doi: 10.1016/j.bcp.2020.114096.
10. Dongol, Y.; Cardoso, F.C.; Lewis, R.J. Spider Knottin Pharmacology at Voltage-Gated Sodium Channels and Their Potential to Modulate Pain Pathways. *Toxins*, **2019**, *11, 626*.
11. Saez, N.J.; Herzig, V. Versatile spider venom peptides and their medical and agricultural applications. *Toxicon*, **2019**, *158*, 109–126.
12. Pineda, S.; Chin, Y.K.Y.; Undheim, E.A.B.; Senff, S.; Mobli, M.; Dauly, C.; Nicholson, G.; Kaas, Q.; Guo, S.; Herzig, V.; et, al. Structural Venomics Reveals Evolution of a Complex Venom by Duplication and Diversification of an Ancient Peptide-Encoding Gene. *Proc Natl Sci USA*, **2020**, *117, 11399–11408*.
13. Langenegger, N.; Nentwig, W.; Kuhn-Nentwig, L. Spider Venom: Components, Modes of Action, and Novel Strategies in Transcriptomic and Proteomic Analyses. *Toxins*, **2019**, *11, 611*.
14. Pineda, S.S.; Chaumeil, P.A.; Kunert, A.; Kaas, Q.; Thang, M.W.C.; Le, L.; Nuhn, M.; Herzig, V.; Saez, N.J.; Cristofori-Armstrong, B.; et, al. ArachnoServer 3.0: An online resource for automated discovery, analysis and annotation of spider toxins. *Bioinformatics*, **2018**, *34*, 1074–1076.
15. Lüddecke, T.; Vilcinskas, A.; Lemke, S. Phylogeny-Guided Selection of Priority Groups for Venom Bioprospecting: Harvesting Toxin Sequences in Tarantulas as a case Study. *Toxins*, **2019**, *11*, 488.
16. Osteen, J.D.; Herzig, V.; Gilchrist, J.; Emrick, J.J.; Zhang, C.; Wang, X.; Castro, J.; Garcia-Caraballo, S.; Grundy, L.; Rychkov, G.Y.; et, al. Selective spider toxins reveal a role for Nav1.1 in mechanical pain. *Nature*, **2016**, *534, 494–499*.
17. Chasagnon, I.R.; McCarthy, C.A.; Chin, Y.K.Y.; Pineda, S.S.; Keramidas, A.; Mobli, M.; Pham, V.; De Silva, T.M.; Lynch, J.W.; Widdop, R.E.; et, al. Potent neuroprotection after stroke afforded by a double-knot

spider-venom peptide that inhibits acid-sensing ion channel 1a. *Proc. Natl. Acad. Sci. USA*, **2017**, *114*, 3750–3755.

18. Richards, K.L.; Milligan, C.J.; Richardson, R.J.; Jancovski, N.; Grunnet, M.; Jacobsen, L.H.; Undheim, E.A.B.; Mobli, M.; Chow, C.Y.; Herzig, V.; et, al. Selective Nav1.1 activation rescues Dravet syndrome mice from seizures and premature death. *Proc. Natl. Acad. Sci. USA*, **2018**, *115*, 8077–8085.
19. Wu, T.; Wang, M.; Wu, W.; Luo, Q.; Jiang, L.; Tao, H.; Deng, M. Spider venom peptides as potential drug candidates due to their anticancer and antinociceptive sctivities. *J Venom Anim Toxins Incl Trop Dis*, **2019**, *25*, *doi: 10.1590/1678–9199-JVATITD-14–63–18*.
20. Pallaghy, P.K.; Nielsen, K.J.; Craik, D.J.; Norton, R.S. A common structural motif incorporating a cysteine knot and a triple-stranded beta-sheet in toxic and inhibitory polypeptides. *Protein Sci*, **1994**, *3*, *1833–1839*.
21. Norton, R.S.; Pallaghy, P.K. The cysteine knot structure of ion channel toxins and related polypeptides. *Toxicon*, **1998**, *36*, *1573–1583*.
22. Von Reumont, B.M.; Undheim, E.A.B.; Jauss, R.T.; Jenner, R.T. Venomics of Remipede Crustaceans Reveals Novel peptide Diversity and Illuminates the Venoms Biological Role. *Toxins*, **2017**, *9*, doi: *10.3390/toxins9080234*.
23. Drukewitz, S.H.; Fuhrmann, N.; Undheim, E.A.B.; Blanke, A.; Giribaldi, J.; Mary, R.; Laconde, G.; Duterte, S.; von Reumont, B.M. A Dipteran´s Novel Sucker Punch: Evolution of Arthropod Atypical Venom with a neurotoxic Component in Robber Flies (Asilidae, Diptera). *Toxins*, **2018**, *10*, *29*, doi: 10.3390/toxins10010029.
24. Drukewitz, S.H.; Bokelmann, L.; Undheim, E.A.B.; von Reumont, B.M. Toxins from scratch? Diverse, multimodal gene origins in the predatory robber fly *Dasypogon diadema* indicate a dynamic venom evolution in dipteran insects. *Gigascience*, **2019**, *8*, *giz081*.
25. Özbek, R.; Wielsch, N.; Vogel, H.; Lochnit, G.; Förster, F.; Vilcinskas, A.; von Reumont, B.M. Proteo-Transcriptomic Characterization of the venom from the Endoparasitoid Wasp *Pimpla turionellae* with Aspects on its Biology and Evolution. *Toxins*, **2019**, *11*, 721.
26. The UniProt Consortium. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res*, **2019**, *47*, *D506–D515*.
27. Friedel, T.; Nentwig, W. Immobilizing and lethal effects of spider venoms on the cockroach and and the common mealbeetle. *Toxicon*, **1989**, *27*, *305–3016*.
28. Hardy, M.C.; Daly, N.L.; Mobli, M.; Morales, R.A.; King, G.F. Isolation of an orally active insecticidal toxin from the venom of an Australian tarantula. *PLoS One*, **2013**, *doi: 10.1371/journal.pone.0073136*.
29. Von Reumont, B.M. Studying Smaller and Neglected Organisms in Modern Evolutionary Venomics Implementing RNAseq (Transcriptomics) – A Critical Guide. *Toxins*, **2018**, *10*, *doi: 10.3390/toxins10010029*.
30. Oldrati, V.; Arrell, M.; Violette, A.; Perret, F.; Sprüngli, X.; Wolfender, J.L.; Stöcklin, R. Advances in Venomics. *Mol Biosyst*, **2016**, *12*, *3530–3543*.
31. Ul-Saban, S.; Rodriguez-Roman, E.; Reitzel, A.M.; Adams, R.M.M.; Herzig, V.; Nobile, C.J.; Saviola, A.J.; Trim, S.A.; Stiers, E.E.; Moschos, S.A.; et, al. The emerging field of venom-microbiomics for exploring venom as a microenvironment, and the corresponding Initiative for Venom Associated Microbes and Parasites (iVAMP). *Toxicon:X*, **2019**, *4*, 100016.
32. Von Reumont, B.M.; Campbell, L.I.; Jenner, R.A. Quo vadis venomics? A roadmap to neglected venomous invertebrates. *Toxins*, **2014**, *6*, 3488–3551.
33. Herzig, V.; King, G.F.; Undheim, E.A.B. Can we resolve the taxonomic bias in spider venom research? *Toxicon:X*, **2019**, *1*, 100005.
34. Bond, J.E.; Garrison, N.L.; Hamilton, C.; Godwin, R.L.; Hedin, M.; Agnarsson, I. Phylogenomics resolves a spider backbone phylogeny and rejects a prevailing paradigm for orb web evolution. *Curr Biol*, **2014**, *24*, *1765–1771*.
35. Fernandez, R.; Hormiga, G.; Giribet, G. Phylogenomic analysis of spiders reveals nonmonophyly of orb weavers. *Curr Biol*, **2014**, *24*, *1772–1777*.
36. Duan, Z.; Cao, R.; Jiang, L.; Liang, S. A combined de novo protein sequencing and cDNA approach to the venomic analysis of Chinese spider *Araneus ventricosus*. *J Proteom.* **2013**, *78*, 416–427.
37. Bush, A.A.; Yu, D.W.; Herberstein, M.E. Function of bright coloration in the wasp spider *Argiope bruennichi* (Araneae: Araneidae). *Proc Biol Sci*, **2008**, *275*, *1337–1342*.
38. Zimmer, S.M.; Welke, K.W.; Schneider, J.M. Determinants of natural mating success in the cannibalistic orb-web spider *Argiope bruennichi*. *PLoS One*, **2012**, *7*, *doi: 10.1371/journal.pone.0031389*.

39. Chinta, S.P.; Goller, S.; Lux, S.; Uhl, G.; Schulz, S. The sex pheromone of the wasp spider *Argiope bruennichi*. *Angew Chem Int Ed*, **2010**, *49*, *2033–2036*.
40. Cory, A.L.; Schneider, J.M. Effects of social information on life history and mating tactics in the orb-web spider *Argiope bruennichi*. *Ecol Evol*, **2017**, *8*, *344–355*.
41. Cory, A.L.; Schneider, J.M. Mate availability does not influence mating strategies in males of the sexually cannibalistic spider *Argiope bruennichi*. *PeerJ*, **2018**, *doi: 10.7717/peerj.5360*.
42. Krehenwinkel, H.; Tautz, D. Northern range expansion of European populations of the wasp spider *Argiope bruennichi* is associated with global warming-correlated genetic admixture and population-specific temperature adaptions. *Mol. Ecol.*, **2013**, 2, 2232–2248.
43. Krehenwinkel, H.; Rödder, D.; Tautz, D. Eco-genomic analysis of the poleward range expansion of the wasp spider *Argiope bruennichi* shows rapid adaption and genomic admixture. *Global Change Biol*, **2015**, *21*, 4320–4332.
44. Krehenwinkel, H.; Pekar, S. An Analysis of factors Affecting Genotyping Success from Museum Specimens reveals an Increase of Genetic and Morphological variation during a Historical Range Expansion of a European Spider. *PLoS One*, **2015**, *10*, *doi: 10.1371/journal.pone.0136337*.
45. Sheffer, M.M.; Uhl, G.; Prost, S.; Lueders, T.; Ulrich, T.; Bengtsson, M.M. Tissue and Population-Level Microbiome Analysis of the Wasp Spider *Argiope bruennichi* identified a Novel bacterial Symbiont. *Microorganisms*, **2019**, *8*, 1. *doi: 10.3390/microorganisms8010008*.
46. Welke, K.W.; Schneider, J.W. Males of the orb-web spider *Argiope bruennichi* sacrifice themselves to unrelated females. *Biol Lett*, **2010**, *6*, *585–588*.
47. Zhang, P.; Fang, H.Y.; Pan, W.J.; Pan, H.C. The complete mitochondrial genome of the wasp spider (*Argiope bruennichi*). *Mitochondrial DNA A DNA Mapp Seq Anal*, **2016**, *27*, *996–997*.
48. Eisner, T.; Dean, J. Ploy and counterploy in predator-prey interactions: Orb-weaving spiders versus bombardier beetles. *Proc. Nat. Acad. Sci. USA*, **1976**, *73*, *1365–1367*.
49. Smith, J.; Undheim, E.A.B. True Lies: Using Proteomics to Assess the Accuracy of Transcriptome-Based Venomics in Centipedes Uncovers False Positives and Reveals Startling Intraspecific Variation in *Scolopendra subspinipes*. Toxins, 2018, *10*, doi:10.3390/toxins10030096.
50. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. *Babraham Institute: Babraham, UK*, **2019**, Available online: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. (accessed on 16 February 2020)
51. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, **2014**, *30*, 2114–2120.
52. Song, L.; Florea, L. Rcorrector: Efficient and accurate error correction for Illumina-seq reads. *Gigascience*, **2015**, *https://doi.org/10.1186/s13742–015–0089-y*.
53. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zheng, Q.; et, al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnol*, **2011**, *29*, 644–652.
54. Bushmanova, E.; Antipov, D.; Lapidus, A.; Prjibelski, A.D. rnaSPAdes: A de novo transcriptome assembler and its application to RNA-Seq data. *Gigascience*, **2019**, *8*, giz100, https://doi.org/10.1093/gigascience/giz100.
55. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Lieber, M.; MacManes, M.D.; et, al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protoc*, **2013**, *8*, 1494–1512.
56. Kim, D.; Paggi, J.M.; Park, C.; Bennett, C.; Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnol*, **2019**, *37*, 907–915.
57. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnol*, **2015**, *33*, 290–295.
58. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protoc*, **2016**, *11*, 1650–1667.
59. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The sequence Alignment/Map format and SAMtools. *Bioinformatics*, **2009**, *25: 2078–2079*.
60. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulus, J.; Bealer, K.; Madden, T.L. Blast+: Architecture and applications. *BMC Bioinformatics*, **2009**, *10*, doi: 10.1186/1471-2105-10-421.

61. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et, al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, **2014**, *30*, 1236–1240.
62. Ankenbrand, M.J.; Weber, L.; Becker, D.; Förster, F.; Bemm, F. TBro: Visualization and management of *de novo* transcriptomes. *Database*, **2016**, *baw146. doi: 10.1093/database/baw146*.
63. Wessel, D.; Flugge, U.I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem*, **1984**, *138, 141–143*.
64. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T*Nature*, **1970**, *227*, 680–685.
65. Diniz, M.R.V.; Paiva, A.L.B.; Guerra-Duarte, C.; Nishiyama, M.Y.; Mudadu, M.A.; Oliveira, U.; Borges, M.H.; Yates, J.R.; Junqueira-de-Azevedo, I.L. An overview of *Phoneutria nigriventer* spider venom using combined transcriptomic and proteomic approaches. *PLoS One*, **2018**, *13, doi: 10.1371/journal.pone.0200628*.
66. Undheim, E.A.B.; Sunagar, K.; Herzig, V.; Kely, L.; Low, D.H.; Jackson, T.N.; Jones, A.; Kurniawan, N.; King, G.F.; Ali, S.A.; et, al. A proteomics and transcriptomics investigation of the venom from the barychelid spider *Trittame loki*. *Toxins*, **2013**, *5, 2488–2503*.
67. Kuhn-Nentwig, L.; Langenegger, N.; Heller, M.; Koua, D.; Nentwig, W. The Dual Prey-Inactivation Strategy of Spiders – In-Depth Venomic Analysis of *Cupiennius salei*. *Toxins*, **2019**, *11, 167*.
68. Zhang, Y.; Tang, C.J.; Wang, F.; Jiang, L.; Xiong, X.; Wang, M.; Rong, M.; Liu, Z.; Liang, S. (2010). Transcriptome analysis of the venom glands of the Chinese wolf spider *Lycosa singoriensis*. *Zoology*, **2010**, *113, 10–18*.
69. Hu, Z.; Chen, B.; Xiao, Z.; Zhou, X.; Liu, Z. Transcriptomic Analysis of the Spider Venom Gland Reveals Venom Diversity and Species Consanguinity. *Toxins*, **2019**, *11, 68*.
70. Cheng, T.C.; Long, R.W.; Wu, Y.Q.; Guo, Y.B.; Liu, D.L.; Peng, L.; Li, D.Q.; Yang, D.W.; Liu, F.X.; Xia, Q.Y. Identification and characterization of toxins in the venom gland of the Chinese bird spider, *Haplopelma hainanum*, by transcriptomic analysis. *Insect Sci*, **2016**, *23, 487–499*.
71. Estrada-Gomez, S.; Cardoso, F.C.; Vargas-Munoz, L.J.; Quintana-Castillo, J.C.; Arenas Gomez, C.M.; Pineda, S.S.; Saldarriaga-Cordoba, M.M.; Venomic, Transcriptomic, and Bioactivity Analyses of *Pamphobeteus verdolaga* venom reveal Complex Disuldide-Rich Peptides That Modulate Calcium Channels. *Toxins*, **2019**, *11, 496*.
72. Brown, R.L.; Haley, T.L.; West, K.A.; Crabb, J.W. Pseudechetoxin: A peptide blocker of cyclic nucleotide-gated ion channels. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 754–759.
73. Yamazaki, Y.; Brown, R.L.; Morita, T. Purification and cloning of toxins from elapid venoms that target cyclic nucleotide-gated ion channels. *Biochemistry*, **2002**, *41*, 11331–11337.
74. Ito, N.; Mita, M.; Takahashi, Y.; Matsushima, A.; Watanabe, Y.G.; Hirano, S.; Odani, S. Novel cysteine-rich secretory protein in the buccal gland secretion of the parasitic lamprey *Lethenteron japonicum*. *Biochem Biophys Res Commun*, **2007**, *358, 35–40*.
75. Ribeiro, J.M.C.; Andersen, J.; Silva-Neto, M.A.C.; Pham, V.M.; Garfield, M.K.; Valenzuela, J.G. Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. *Insect Biochem Mol Biol*, **2004**, *34*, 61–79.
76. Caruso, B.; Bonadonna, P.; Bovo, C.; Melloni, N.; Lombardo, C.; Senna, G.; Lippi, G. Wasp venom allergy screening with recombinant allergen testing. Diagnostic performance of rPol d 5 and and rVes v 5 for differentiating sensitization to *Vespula* and *Polistes* subspecies. *Clinica Chimica Acta*, **2016**, *453*, 170–173.
77. Fang, K.S.Y.; Vitale, M.; Fehlner, P.; King, T.P. cDNA cloning and primary structure of a white-face hornet venom allergen, antigen *Proc. Natl. Acad. Sci. USA*, **1988**, *85*, 895–899.
78. Milne, T.J.; Abbenante, G.; Tyndall, J.D.A.; Halliday, J.; Lewis, R.J. Isolation and characterization of of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *J. Biol. Chem.*, **2003**, *278*, 31105–31110.
79. Undheim, E.A.B.; Grimm, L.L.; Low, C.F.; Morgenstern, D.; Herzig, V.; Zobel-Thropp, P.A.; Pineda, S.S.; Habib, R.; Dziemborowicz, S.; Fry, B.G.; et, al. Weaponization of a Hormone: Convergent recruitment of Hyperglycemic Hormone into the Venom of Arthropod Predators. *Structure*, **2015**, *23, 1283–1292*.
80. Kuhn-Nentwig, L.; Largiader, C.R.; Streitberger, K.; Chandru, S.; Baumann, T.; Kampfer, U.; Schaller, J.; Schurch, S.; Nentwig, W. Purification, cDNA structure and biological significance of a single insulin-like growth factor-binding domain protein (SIBD-1) identified in the hemocytes of the spider *Cupiennius salei*. *Insect Biochem. Mol. Biol.*, **2011**, *41*, 891–901.

81. Esteves, E.; Maruyama, S.R.; Kawahara, R.; Fujita, A.; Martins, L.A.; Righi, A.A.; Costa, F.B.; Palmisano, G.; Labruna, M.B.; Sa-Nunes, A.; et, al. Analysis of the Salivary Gland Transcriptome of Unfed and Partially Fed *Amblyomma sculptum* Ticks and descriptive Proteome of the Sliva. *Front. Cell. Infect. Microbiol.*, **2017**, *7*, 476–476.
82. Rokyta, D.R.; Ward, M. Venom-gland transcriptomics and venom proteomics of the black-back scorpion (*Hadrurus spadix*) reveal detectability challenges and an unexplored realm of animal toxin diversity. *Toxicon*, **2017**, *128*, 23–37.
83. Santibanez-Lopez, C.E.; Cid-Uribe, J.I.; Batista, C.V.; Ortiz, E.; Possani, L.D. Venom Gland Transcriptomic and Proteomic Analyses of the Enigmatic Scorpion *Superstitionia donensis* (Scorpiones: Superstitioniidae), with Insights on the Evolution of its Venom Components. *Toxins*, **2016**, *8*, *E367–E367*.
84. Ward, M.J.; Ellsworth, S.A.; Rokyta, D.R. Venom-gland transcriptomics and proteomics of the Hentz striped scorpion (*Centruroides hentzi*; Buthidae) reveal high toxin diversity in a harmless member of a lethal family. *Toxicon*, **2017**, *142: 14–29*.
85. Kono, N.; Nakamura, H.; Ohtoshi, R.; Moran, D.A.P.; Shinohara, A.; Yoshida, Y.; Fujiwara, M.; Mori, M.; Tomita, M.; Arakawa, K. Orb-weaving spider *Araneus ventricosus* genome elucidates the spidroin gene catalogue. *Sci. Rep.*, **2019**, *9*, *8380*.
86. Arvidson, R.; Kaiser, M.; Lee, S.S.; Urenda, J.P.; Dail, C.; Mohammed, H.; Nolan, C.; Pan, S.; Stajich, J.E.; Libersat, F.; et, al. Parasitoid Jewel Wasp Mounts Multipronged Neurochemical Attack to Hijack a Host brain. *Mol Cell Proteomics*, **2019**, *18*, *99–114*.
87. Yap, M.K.K.; Misuan, N. Exendin-4 from *Heloderma suspectum* venom: From discovery to its latest application as type II diabetes combatant. *Basic Clin Pharmacol Toxicol*, **2019**, *124*, *513–527*.
88. Gaudino, G.; Fasolo, A.; Merlo, G.; Lazarus, L.H.; Renda, T.; D.′este, L.; Vandesande, F. Active peptides from amphibian skin are also amphibian neuropeptides. *Peptides*, **1985**, *6*, 209–213.
89. Roelants, K.; Fry, B.G.; Norman, J.A.; Clynen, E.; Schoofs, L.; Bossuyt, F. Identical skin toxins by convergent molecular adaptions in frogs. *Curr Biol*, **2010**, *20*, *125–130*.
90. Roelants, K.; Fry, B.G.; Lumen, Y.; Stijlemans, B.; Brys, L.; Kok, P.; Clynen, E.; Schoofs, L.; Cornelis, P.; Bossuyt, F. Origin and Functional Diversification of an Amphibian Defense Peptide Arsenal. *PLoS genet*, **2013**, *9*, e1003662, doi: 10.1371/journal.pgen.1003662.
91. Lüddecke, T.; Schulz, S.; Steinfartz, S.; Vences, M. A salamander's toxic arsenal: Review of skin poison diversity and function in true salamanders, genus *Salamandra*. *The Science of Nature*, **2018**, *105*, *56*.
92. Daltry, J.C.; Wüster, W. Diet and snake venom evolution. *Nature*, **1996**, *379*, 537–540.
93. Li, M.; Fry, B.G.; Kini, M. Eggs-Only Diet: Its Implications for the Toxin Profile Changes and Ecology of the Marbled Sea Snake (*Aipysurus eydouxii*). *J. Mol. Evol.*, **2005**, *60*, *81–89*.
94. Fry, B.G.; Wüster, W.; Kini, M.; Brusic, V.; Khan, A.; Venkataraman, D.; Rooney, A.P. Molecular evolution and phylogeny of elapid snake venom three-finger toxins. *J. Mol. Evol.*, **2006**, *57*, 110–129.
95. Lyons, K.; Dugon, M.M.; Healy, K. Diet Breadth Mediates the Prey Specificity of Venom Potency in Snakes. *Toxins*, **2020**, *12*, *74*.
96. Phuong, M.A.; Mahardika, G.N.; Alfaro, M.E. Dietary breadth is positively correlated with venom complexity in cone snails. *BMC Genomics*, **2016**, *17*, 401.
97. Szymkowiak, P.; Tryjanowski, P.; Winiecki, A.; Grobelny, S.; Konwerski, S. Habitat differences in the food composition of the wasp-like spider *Argiope bruennichi* (Scop.) (Aranei: Araneidae) in Poland. *Belg. J. Zool.*, **2005**, *135*, *33–37*.
98. Dean, J.; Aneshansley, D.J.; Edgerton, H.E.; Eisner, T. Defensive spray of the bombardier beetle: A biological pulse jet. *Science*, **1990**, *248*, *1219–1221*.
99. Nyffeler, M.; Benz, G. Eine Notiz zum Beutefangverhalten der Radnetzspinne *Argiope bruennichi* (Scopoli) (Araneae, Araneidae). *Revue suisse Zool.*, **1982**, *89*, *23–25*.
100. Harwood, R.H. Predatory behaviour of *Argiope aurantia* (Lucas). *Am. Midl. Nat.*, **1974**, *91*, 130–139.
101. Robinson, M.H. Predatory behaviour of *Argiope argentata* (Fabricius). *Am. Zool.*, **1969**, *9*, 161–174.
102. Nisani, Z.; Boskovic, D.S.; Dunbar, S.G.; Kelln, W.; Hayes, W.K. Investigating the chemical profile of regenerated scorpion (*Parabuthus transvaalicus*) venom in relation to metabolic cost and toxicity. *Toxicon*, **2012**, *60*, 315–323.
103. Blennerhassett, R.A.; Bell-Anderson, K.; Shine, R.; Brown, G.P. The cost of chemical defence: The impact of toxin depletion on growth and behaviour of cane toads (*Rhinella marina*). *Proc. Biol. Sci.*, **2019**, *286*, *20190867*.

104. Saggiomo, S.L.; Zelenka, C.; Seymor, J. Relationship between food and venom production in the estuarine stonefish *Synanceia horrida*. *Toxicon,* **2017**, *125*, 19–23.
105. Walker, A.A.; Mayhew, M.L.; Jin, J.; Herzig, V.; Undheim, E.A.B.; Sombke, A.; Fry, B.G.; Merrit, D.J.; King, G.F. The assassin bug *Pristesancus plagipennis* produces two distinct venoms in separate gland lumens. *Nat Comm,* **2018**, *9*, *755*.
106. Guehrs, K.H.; Schlott, B.; Grosse, F.; Weisshart, K. Environmental conditions impinge on dragline silk protein composition. *Insect Mol Biol,* **2008**, *17*, *553–564*.
107. Zobel-Thropp, P.A.; Mullins, J.; Kristensen, C.; Kronmiller, B.A.; David, C.L.; Breci, L.A.; Binford, G.J. Not so dangerous After All? Venom Composition and Potency of the Pholcid (Daddy Long-leg) Spider *Physocyclus mexicanus*. Front. Ecol. Evol., **2019**, *7*, 256, doi: 10.3389/fevo.2019.00256.

VI Appendix

Further Publications

In order to broaden the scope of this research, several studies on metazoan toxin systems, their evolution and natural history were conducted in parallel to the herein depicted works on spiders. These additional works were conducted mostly on amphibian systems, in particular the fire salamander *Salamandra salamandra*, but include also ribbon worms (Nemertea).
The following studies have been conducted and published:

Jung, L., J. Dusek, **T. Lüddecke**, V. Schulz, K. Maier-Sam, L. Habich, A. Mosebach, M. Lienz, H.P. Ziemek (2020): Epidemiological screening of captive salamanders reveals current absence of *Batrachochytrium salamandrivorans* in private collections throughout the federal state of Hesse (Germany). Salamandra 56(3): 233-238.

von Reumont, B.M., T. **Lüddecke, T**. Timm, G. Lochnit, A. Vilcinskas, J. von Döhren, M. Nilsson (2020): Proteo-Transcriptomic Analysis Identifies Potential Novel Toxins Secreted by the Predatory, Prey-Piercing Ribbon Worm *Amphiporus lactifloreus*. Marine Drugs 18: 407.

Preißler K., Gippner S., **Lüddecke T**., Krause E.T., Schulz S., Vences M., Steinfartz S. (2019): More yellow more toxic? Sex rather than alkaloid content is correlated with yellow coloration in the fire salamander. Journal of Zoology, doi: 10.1111/jzo.12676.

Knepper, J., **T. Lüddecke**, K. Preißler, M. Vences, S. Schulz (2019): Isolation and Identification of Alkaloids from Poisons of Fire Salamanders (*Salamandra salamandra*). Journal of Natural Products 82(5).

Sanchez, E., H. Pröhl, **T. Lüddecke**, S. Schulz, S. Steinfartz, M. Vences (2019): The conspicuous postmetamorphic coloration of fire salamanders, but not their toxicity, is affected by larval background albedo. Journal of Experimental Zoology B – Molecular and Developmental Evolution doi: 10.1002/jez.b.22845.

Stawikowski R., **Lüddecke T**. (2019): Description of defensive postures of the natterjack toad *Epidalea calamita* (Laurenti 1768) and notes on the release of toxic secretions. Herpetology Notes 12: 443-445.

Conference Contributions
20th World Congress of the International Society on Toxinology (2019)
Buenos Aires, Argentina
Oral presentation:" Recent advances in *Salamandra* skin poison research"

20th World Congress of the International Society on Toxinology (2019)
Buenos Aires, Argentina
Oral presentation:" Translational tarantula phylogenomics: Evolution of theraphosid spiders and their defensive arsenal with implications for venom bioprospecting"

20th World Congress of the International Society on Toxinology (2019)
Buenos Aires, Argentina
Poster presentation:" The venom gland transcriptome of the wasp spider *Argiope bruennichi*"

2nd Congresso Latinoamericano de Biogeografia & Humboldt 250 Conference of the International Biogeography Society (2019)
Quito, Ecuador

Oral presentation: "On the Evolution of Theraphosidae: Phylogenomics of tarantulas unravels major genetic lineages and suggests a rapid diversification in the Americas"

Recombinant Protein Production 10 Conference of the European Federation of Biotechnology (2019)
Hersonissos, Greece (Crete)
Poster presentation: "Animal toxins for application in medicine and plant protection"

Net@Phys Workshop of the German Zoological Society – Section for Physiology (2019)
Ebsdorfergrund, Germany
Oral presentation: "Genetic and phenotypic changes during metamorphosis and determination of the developmental origin of toxicity in the common fire salamander (*Salamandra salamandra*)"

Deutscher Herpetologentag of the German Society for Herpetology and Herpetoculture (2018)
Magdeburg, Germany
Oral presentation: "The past, present and future of *Salamandra* skin poison research"

Grants raised
German Academic Exchange Service (2019): Travel Grant for 20th World Congress of the International Society on Toxinology (2019); Buenos Aires, Argentina

German Embassy of Ecuador: Travel Grant for 2nd Congresso Latinoamericano de Biogeografia & Humboldt 250 Conference of the International Biogeography Society (2019); Quito, Ecuador

Contributions per Chapter
This thesis is the outcome of my work on spider evolutionary systematics and toxinology over the last years. This research has been conducted in a highly prolific framework of me and my colleagues of different working groups and institutions dispersed around the world. I will here provide an overview of the contributions of each researcher associated to this research per chapter.

Chapter I: I designed the study together with HK and MV. I and HK, FG, GC, IW, SJL and RT collected sample material. I performed lab work, bioinformatic data analyses and submission to repositories. I produced picture material of all samples included and deposited voucher specimen to the zoological state collection. MV and I wrote the manuscript.

Chapter II: I and SF designed the study and acquired sample materials. I, SF and SK performed laboratory work. I, MV and WHP attracted funding for the study. SF, WHP, MV and DQC conducted bioinformatic analyses. HK, SJL, IW, VvW and RT contributed to experimental design and assured valid taxonomical plus morphological assignments. TL and SF wrote the manuscript with support of all other authors.

Chapter III: I designed the work together with SL and AV. I analysed the data and wrote the manuscript with support of the other colleagues.

Chapter IV: I designed the study with SL and AV. AV acquired funding for the study. I collected sample material from the wild. I and HS performed laboratory work. Data was analysed and interpreted by me, HS and BMvR. I wrote the manuscript with input of HS, BMvR, SL and AV.

Chapter V: I designed the study with SL, AV and BMvR. AV acquired funding for the project. I collected fieldwork and obtained sample material for transcriptomics and proteomics. I, SL, TT and GL performed the laboratory work. I analysed the transcriptomic data together with BMvR, AB and FF. Proteome data was analysed by me, TT, GL and SL. I wrote the manuscript with input from all other authors.

List of abbreviations

AA	=	Aminoacids
AS	=	Arachnoserver database
BI	=	Bayesian inference
CAP	=	cysteine-rich secretory protein/antigen 5 / pathogenesis-related 1 protein
DDH	=	disulfide-directed beta-hairpin fold
dICK	=	Double inhibitory cysteine knot peptide
ESI	=	Electrospray ionisation
HAND	=	Helical arthropod derived neuro peptide
ICK	=	Inhibitory cysteine knot peptide, knottin
ITP/CHH	=	ion transport peptide/ crustacean hyperglycemic hormone neuropeptides
LC	=	Liquid chromatography
LCA	=	Last common ancestor
LTX	=	Latrotoxin (alpha-Latrotoxin)
MALDI	=	Matrix assisted laser desorption ionisation
MIT-1	=	Venom peptides with colipase fold (atracotoxins)
ML	=	Maximum likelihood
MP	=	Maximum parsimony
MS	=	Mass spectrometry
OG	=	Orthologous genes
PLD	=	Phospholipase D, Sphingomyelinase D
SDS-PAGE	=	Sodium-dodecylsulfate electrophoresis
SRA	=	Sequence read archive
Tex31	=	A CAP protein from venom of *Conus textile*, proteolytic
TOF	=	Time of flight
VZ	=	Venomzone database
WSC	=	World spider catalog database

Der Lebenslauf wurde aus der elektronischen Fassung der Dissertation entfernt

The curriculum vitae was removed from the electronical version of the thesis