Phylogenomic analysis and revised classification of atypoid mygalomorph spiders (Araneae, Mygalomorphae), with notes on arachnid ultraconserved element loci

Marshal Hedin [Corresp., 1], **Shahan Derkarabetian** [1, 2, 3], **Adan Alfaro** [1], **Martín J Ramírez** [4], **Jason E Bond** [5]

[1] Department of Biology, San Diego State University, San Diego, California, United States

[2] Department of Biology, University of California, Riverside, Riverside, California, United States

[3] Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts, United States

[4] Division of Arachnology, Museo Argentino de Ciencias Naturales “Bernardino Rivadavia”, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina

[5] Department of Entomology and Nematology, University of California, Davis, Davis, California, United States

Corresponding Author: Marshal Hedin
Email address: mhedin@sdsu.edu

The atypoid mygalomorphs include spiders from three described families that build a diverse array of entrance web constructs, including funnel-and-sheet webs, purse webs, trapdoors, turrets and silken collars. Molecular phylogenetic analyses have generally supported the monophyly of Atypoidea, but prior studies have not sampled all relevant taxa. Here we generated a dataset of ultraconserved element loci for all described atypoid genera, including taxa (*Mecicobothrium* and *Hexurella)* key to understanding familial monophyly, divergence times, and patterns of entrance web evolution. We show that the conserved regions of the arachnid UCE probe set target exons, such that it should be possible to combine UCE and transcriptome datasets in arachnids. We also show that different UCE probes sometimes target the same protein, and under the matching parameters used here show that UCE alignments sometimes include non-orthologs. Using multiple curated phylogenomic matrices we recover a monophyletic Atypoidea, and reveal that the family Mecicobothriidae comprises four separate and divergent lineages. Fossil-calibrated divergence time analyses suggest ancient Triassic (or older) origins for several relictual atypoid lineages, with late Cretaceous / early Tertiary divergences within some genera indicating a high potential for cryptic species diversity. The ancestral entrance web construct for atypoids, and all mygalomorphs, is reconstructed as a funnel-and-sheet web.

Phylogenomic analysis and revised classification of atypoid mygalomorph spiders (Araneae, Mygalomorphae), with notes on arachnid ultraconserved element loci

Marshal Hedin[1*], Shahan Derkarabetian[1,2,3], Adan Alfaro[1], Martín J. Ramírez[4], Jason E. Bond[5]

[1] Department of Biology
San Diego State University
San Diego, California 92182-4614, United States

[2] Department of Biology
University of California, Riverside
Riverside, California 92521, United States

[3] Department of Organismic and Evolutionary Biology
Museum of Comparative Zoology
Harvard University
Cambridge, Massachusetts 02138, United States

[4] Division of Arachnology
Museo Argentino de Ciencias Naturales “Bernardino Rivadavia”
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

[5] Department of Entomology and Nematology
University of California, Davis
Davis, California 95616-5270, United States

* Correspondence: mhedin@sdsu.edu

Abstract

The atypoid mygalomorphs include spiders from three described families that build a diverse array of entrance web constructs, including funnel-and-sheet webs, purse webs, trapdoors, turrets and silken collars. Molecular phylogenetic analyses have generally supported the monophyly of Atypoidea, but prior studies have not sampled all relevant taxa. Here we generated a dataset of ultraconserved element loci for all described atypoid genera, including taxa (*Mecicobothrium* and *Hexurella)* key to understanding familial monophyly, divergence times, and patterns of entrance web evolution. We show that the conserved regions of the arachnid UCE probe set target exons, such that it should be possible to combine UCE and transcriptome datasets in arachnids. We also show that different UCE probes sometimes target the same protein, and under the matching parameters used here show that UCE alignments sometimes include non-orthologs. Using multiple curated phylogenomic matrices we recover a monophyletic Atypoidea, and reveal that the family Mecicobothriidae comprises four separate and divergent lineages. Fossil-calibrated divergence time analyses suggest ancient Triassic (or older) origins for several relictual atypoid lineages, with late Cretaceous / early Tertiary divergences within some genera indicating a high potential for cryptic species diversity. The ancestral entrance web construct for atypoids, and all mygalomorphs, is reconstructed as a funnel-and-sheet web.

Introduction

Phylogenetic evidence now overwhelmingly indicates that the mygalomorph spiders, including trapdoor spiders and their kin, are divided into the primary clades Avicularioidea and Atypoidea (Hedin & Bond, 2006; Bond et al. 2012; Hamilton et al. 2016; Garrison et al. 2016, Wheeler et al. 2017; Hedin et al. 2018a; Fernández et al. 2018, Opatova et al. 2019). Avicularioidea includes the most familiar mygalomorphs (e.g., tarantulas), and the bulk of known taxonomic diversity (World Spider Catalog, 2019). Phylogenomic analyses based on sequence-capture data have now dramatically changed our understanding of family-level diversity and interrelationships within the avicularioids (Hamilton et al. 2016; Hedin et al. 2018a; Opatova et al. 2019), with many families previously suspected of non-monophyly now known to constitute multiple independent lineages (Hedin et al. 2018a; Opatova et al. 2019).

Avicularioids are sister to Atypoidea, the latter group representing an old taxonomic hypothesis (Simon, 1892). Atypoidea was first suggested then refuted by morphology, then supported by few-gene molecular studies, and is now seemingly confirmed by phylogenomic approaches. This clade, sometimes referred to as the “atypical tarantulas” (Gertsch, 1949), includes three described families (Antrodiaetidae, Atypidae, Mecicobothriidae) whose members possess dorsal abdominal tergites (**Fig. 1B, E, G**). These tergites are believed to represent the vestiges of abdominal segmentation, as found in spider relatives and early-diverging spiders. Adult male atypoids possess a palpus with a conductor, females have bipartite spermathecal organs, and members of both sexes typically possess six spinnerets (Eskov & Zonstein, 1990). This clade is relatively ancient, as multiple fossil genera placed within the three described families are known from the Lower Cretaceous (100-112 MYA) of Mongolia (Eskov & Zonstein, 1990)**.** Dalla Vecchia and Selden (2013) placed the Triassic (210-215 MYA) *Friularachne* into Atypoidea, but left the family-level placement unspecified.

Atypoids utilize silk to build many different types of burrow entrance constructs (Coyle, 1986). The mecicobothriid genera *Mecicobothrium, Megahexura, Hexura,* and *Hexurella* are all ground-dwelling spiders found living under objects or in earthen crevices, using elongate spinnerets to build silken funnel-and-sheet webs (Gertsch & Platnick, 1979; Costa & Pérez-Miles, 1998; pers. obs.; **Fig. 1A, B, D, F**). The atypid genera either live in subterranean burrows with open silk-lined entrances (*Calommata),* or build cryptic silken capture tubes extending horizontally or vertically from burrow entrances (all atypid genera, Schwendinger, 1990; Fourie, Haddad & Jocqué, 2011; **Fig. 1C**). Finally, the antrodiaetids live in subterranean burrows with silken turret or collapsible collar entrances (*Antrodiaetus*), or build trapdoors to cover their burrows (*Aliatypus,* Coyle, 1971; **Fig. 1E, G, H**). Most atypoid taxa are distributed on northern continents, although *Mecicobothrium* occurs in southern South America, and *Calommata* species are found in east Asia and throughout sub-Saharan Africa (World Spider Catalog, 2019).

Faircloth et al. (2012) first used the sequence capture of ultraconserved elements (UCEs) in phylogenomic analyses of various amniote lineages. In vertebrates more

generally, core UCE regions show extreme sequence conservation, making design of broad-utility nucleotide baits possible (e.g., for all fishes, all amniotes, etc.). The function and genomic position of vertebrate UCEs has remained somewhat elusive, although most are believed to have regulatory functions and lie outside of exons (e.g., Bejerano et al. 2004; Polychronopoulos et al. 2017; McCole et al. 2018). More recently, UCE baits have been designed for megadiverse arthropod lineages, including arachnids and multiple insect orders (Faircloth et al. 2015; Faircloth, 2017). Bossert and Danforth (2018) showed that a universal set of 100 UCEs are shared across all arthropods, and that these "core" UCEs are entirely or partially exonic in origin, thus differing from vertebrate UCEs. In this paper we further explore the function of genomic regions captured by the arachnid UCE bait set. This set was tested *in situ* by Starrett et al. (2017), and has been used in multiple phylogenomic studies (Derkarabetian et al. 2018a; Hedin et al. 2018a, 2018b; Wood et al. 2018). Knowing the functional role of arachnid UCEs has clear importance in phylogenomic analyses, potentially impacting sequence alignment, model selection, data partitioning, detection of paralogy, and so on. This is particularly true in a lineage such as spiders, where an ancient whole-genome duplication event has occurred (Clarke et al. 2015; Schwager et al. 2017), perhaps complicating orthology assignment.

Interrelationships within Atypoidea have varied considerably in past molecular phylogenetic studies (**Fig. 2**), and no prior studies have simultaneously sampled all known (described) atypoid genera. Here we present such an analysis with all genera, including key taxa such as *Mecicobothrium* and the diminutive *Hexurella,* neither included in prior molecular phylogenetic analyses. Using an annotated UCE locus set with BLAST evidence for gene function and orthology, we demonstrate that Atypoidea is monophyletic, while revealing multiple cases of non-monophyly within described families. Early-diverging atypoid lineages are often species-poor (approximating monotypic) and use silk to build funnel-and-sheet webs, while more diverse silken constructs have evolved in derived atypoid lineages. Similar patterns of species and web diversification occur in parallel in the avicularioids (Opatova et al. 2019).

Materials & Methods

Taxon sampling. Representatives of all nine described atypoid genera (World Spider Catalog, 2019) were sampled. Within genera, the sample included all three known species in the synonymized genus *Atypoides* (species now included in *Antrodiaetus,* Hendrixson & Bond, 2007), two species of originally-described *Antrodiaetus* which span the hypothesized root node of this taxon (Hendrixson & Bond, 2007, 2009), two species of *Aliatypus* which span the hypothesized root node of this genus (Satler et al. 2011), both described *Hexura* species, two geographically separated species of *Hexurella*, two geographically distant populations of the monotypic *Megahexura fulva*, and two species of the genus *Sphodros*. Only *Mecicobothrium, Calommata* and *Atypus* were represented by single specimens (**Table S1**). To confirm atypoid monophyly we sampled a handful of representative avicularioid taxa, including genera representing multiple early-diverging avicularioid lineages (Bond et al. 2012; Hedin et al. 2018a; Opatova et al. 2019). Mygalomorphs are sister to araneomorph spiders – we used an

early-diverging araneomorph lineage (*Hypochilus*) to root trees. In total, we gathered original UCE data for 15 specimens; data for 12 specimens were taken from previous studies (Starrett et al. 2017; Hedin et al. 2018a; **Table S1**). Permits for the collection of Australian specimens were granted by the Queensland Environmental Protection Agency (permit #WISP01242003).

DNA extraction. Most specimens were preserved for DNA studies (preserved in high percentage ethyl alcohol at -80C), and genomic DNA was extracted from leg tissue using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA). For a handful of tissues preserved in 70–80% we used either standard phenol/chloroform extractions with 24-hour incubation for lysis, or used a modification of the Tin, Economo & Mikheyev (2014) protocol (**Table S1**). All extractions were quantified using a Qubit Fluorometer (Life Technologies, Inc.) and quality was assessed on agarose gels. Between 22–500 ng total DNA was used for UCE library preparation (**Table S1**).

UCE data collection & matrix assembly. UCE data were collected in multiple library preparation and sequencing experiments. Up to 500 ng of genomic DNA was used in sonication, using a Covaris M220 Focused-ultrasonicator. Library preparation followed methods previously used for arachnids as in Starrett et al. (2017), Derkarabetian et al. (2018a, 2018b), and Hedin et al. (2018a, 2018b). Target enrichment was performed using the MYbaits Arachnida 1.1K version 1 kit (Arbor Biosciences; Faircloth, 2017) following the Target Enrichment of Illumina Libraries v. 1.5 protocol (http://ultraconserved.org/#protocols). Libraries were sequenced on an Illumina HiSeq 2500 (Brigham Young University DNA Sequencing Center).

Raw demultiplexed reads were processed with the PHYLUCE pipeline (Faircloth, 2016). Quality control and adapter removal were conducted with the ILLUMIPROCESSOR wrapper (Faircloth, 2013). Assemblies were created with VELVET (Zerbino et al. 2008) and/or TRINITY (Grabherr et al. 2011), both at default settings. When contigs from both assemblies were available, these were combined for probe matching, retrieving assembly-specific UCEs and overall increasing the number of UCEs per sample relative to using only a single assembly method. Contigs were matched to probes using minimum coverage and minimum identity values at liberal values of 65. UCE loci were aligned with MAFFT (Katoh & Standley, 2013) and trimmed with GBLOCKS (Castresana, 2000; Talavera & Castresana, 2007), using --b1 0.5 --b2 0.5 --b3 6 --b4 6 settings in the PHYLUCE pipeline.

UCE locus annotation, Matrix filtering, and Phylogenomic analyses. 698 loci were found in a PHYLUCE 70% occupancy matrix. For the consensus sequence from each locus alignment, BLAST X searches in Geneious 10.1 (Biomatters Ltd.) were conducted against a local database (max e value of $1x10^{-10}$) comprising protein sequences for four arachnid taxa: *Limulus polyphemus* (https://www.ncbi.nlm.nih.gov/genome/787), *Ixodes scapularis* (https://www.ncbi.nlm.nih.gov/genome/?term=523), *Stegodyphus mimosarum* (https://www.ncbi.nlm.nih.gov/genome/?term=12925) and *Parasteatoda tepidariorum* (https://www.ncbi.nlm.nih.gov/genome/?term=13270).

BLAST annotation indicated that essentially all spider UCE loci are either entirely exonic, or exons with flanking introns (see *Results*). This annotation information allowed us to further curate PHYLUCE alignments in several ways. First, we discovered that some individual loci were part of the same protein, likely exons (or parts thereof) separated by introns (see *Results*). Second, annotation indicated that some UCE loci could potentially include paralogs of the same protein, or orthologs of two or more different proteins. We thus visually inspected all UCE locus alignments and excluded loci with non-orthology as evidenced by congeneric taxa with divergent sequences, using RAxML gene trees (see below) to confirm this non-orthology. Finally, annotation allowed us to define exon/intron boundaries, and exclude a majority of intron sequence for some analyses.

Three matrices were assembled for phylogenomic analyses, including 1) 70% occupancy PHYLUCE unfiltered (including same protein duplicates, some loci with non-orthologs), 2) 70% exon + intron, no “paralogs”, retaining one UCE locus from a set including duplicates (alignment with most sequences, or longest alignment if approximately same number of taxa), 3) 70% filtered as #2 above, plus using stricter GBLOCKS settings (--b1 0.5 --b2 0.85 --b3 4 --b4 8) to further trim alignments. We visually checked to confirm that these trimmed alignments comprised mostly exon data. Unpartitioned and partitioned concatenated maximum likelihood analyses were run for each dataset above. Unpartitioned analyses were conducted with RAxML version 8.2 (Stamatakis, 2014) using a complex GTRGAMMA model and 200 rapid bootstrap replicates. Partitioned maximum likelihood analyses were conducted using IQ-TREE (Nguyen et al. 2015; Chernomor, von Haeseler & Minh, 2016) with partitions and models determined using ModelFinder (Kalyaanamoorthy et al. 2017), and support estimated via 1000 ultrafast bootstrap replicates (Hoang et al. 2018). Finally, we used SVDquartets (Chifman & Kubatko 2014; Chifman & Kubatko 2015) with n = 500 bootstraps, as implemented in PAUP* 4.0a163 (Swofford, 2003).

Web Evolution and Divergence Time Analysis. Mesquite version 3.51 (Maddison & Maddison, 2018) was used to reconstruct ancestral states for entrance web constructs, with tip values scored as seven different discrete states. Tip scorings were derived from published literature (references in *Introduction*), supplemented with original observation. Maximum likelihood reconstructions were produced using the one-parameter Markov k-state model (Lewis, 2001), using the RAxML exon only topology as input.

We estimated absolute divergence times using the lognormal relaxed clock model (Thorne, Kishino & Painter, 1998) implemented in Phylobayes 4.1c (Lartillot & Philippe, 2004). We used the exon only matrix, with the RAxML topology as a constraint tree. Four MCMC chains were run in parallel, stopping after 30,000 points. Analyses were checked for convergence, and considered converged when the largest discrepancy observed across bipartitions (maxdiff) was equal to 0. Posterior estimates of ages and highest posterior density (HPD) values were summarized on a single target tree from all input trees using TreeAnnotator (Bouckaert et al. 2014). Three fossil calibrations were used, with a soft bounds model (Yang & Rannala, 2006) and a birth death prior on divergence times, as follows: 1) minimum age for the root node of mygalomorphs = 240 MYA, based on *Rosamygale*, the oldest known mygalomorph fossil (Selden & Gall,

1992). This taxon was placed by original authors as an avicularioid, but is treated more conservatively here. 2) minimum age for the root node of Atypoidea = 210 MYA, based on *Friularachne* (Dalla Vecchia & Selden, 2013). We also used an alternative second calibration, using the Eskov and Zonstein (1990) fossils *Ambiortiphagus* and *Cretacattyma* to set the minimum age for the most recent common ancestor of Atypidae <> Antrodiaetidae at 100 MYA. 3) minimum age for the root node of Avicularioidea = 216 MYA, based on *Edwa* (Raven, Jell & Knezour, 2015), a likely early-diverging avicularioid. For all three calibrations we used a maximum age of 390 MYA, corresponding to the age of fossil Uraraneida, the putative sister group of spiders (Selden, Shear & Sutton, 2008). This approximate age is in accord with maximum dates derived from other molecular clock analyses of spiders (Ayoub et al. 2007; Wood et al. 2012; Starrett et al. 2013; Fernández et al. 2018; Opatova et al. 2019).

Results

Voucher data, input DNA values, assembled contig numbers, and UCE locus numbers are found in **Table S1**. Except for museum samples of *Mecicobothrium*, all samples returned multiple 100s of loci for all matrices. We highlight *Mecicobothrium* – although we are confident in the results presented here (based on identical placement across all analyses), future studies with fresh specimens should verify the phylogenetic placement discussed below. Raw reads from fifteen original samples have been submitted to the SRA (SAMN10839235 - 10839249).

Annotation of the ~ 700 loci derived from the PHYLUCE pipeline indicates that spider UCEs are primary exonic in origin, as essentially all (> 98%) alignments BLAST to proteins found in *Stegodyphus* and *Parasteatoda* spiders, with relatively high BIT scores (**Tables S2, S3**). We note that *Stegodyphus* and *Parasteatoda* are true spiders in the clade sister to mygalomorphs; we did not conduct custom BLAST searches against mygalomorphs, as the only sequenced genome (*Acanthoscurria*) is low coverage and incomplete (Sanggaard et al. 2014). Even the handful of UCE loci without BLAST hits contained open reading frames of variable length, and these could represent proteins that are particularly divergent from araneomorphs, or restricted to mygalomorphs.

We found that 112 total alignments mapped to the same 74 proteins (i.e., different alignments hit same protein; **Tables S2, S3**). We confirmed that the conserved regions of these separate alignments represented different exons of typically large proteins, and that these exons are likely separated by very long introns (using the known short exon- long intron structure of spiders as models, see Sanggaard et al. 2014). BLAST and visual assessment of the 70% PHYLUCE matrices indicated that 106 alignments included non-orthologous sequences, and this was confirmed via RAxML analysis of these individual alignments (.tre files in **Data S1**). Non-orthology was also indicated by annotation, as most alignments including “paralogs” hit two or more different proteins at similar BIT score values (**Table S2**). The issue of non-orthology is further discussed below. The final matrices were populated as follows: 1) PHYLUCE unfiltered 70% occupancy (698 loci, 191855 basepairs), 2) 70% filtered exon

+ intron (480 loci, 137170 basepairs), 3) 70% filtered exon only (480 loci, 71483 basepairs). All aligned matrices and .tre files are available in **Data S1.**

Except for one node, all nine phylogenomic analyses recover an identical branching topology within Atypoidea, albeit with variation in branch lengths and node support (**Fig. 3**). The single node in question involves the interrelationships of *Antrodiaetus riversi*, *A. gertschi*, and *A. hadros*, all previously in the synonymized genus *Atypoides*. Overall, the following pertinent clades were recovered with high support (bootstrap > 95 and posterior probability > 0.95) in all analyses: Avicularioidea, Atypoidea, Atypidae, and all genera with multiple sampled species. The fragmentation of mecicobothriids into four separate lineages is strongly supported, with the genus *Hexura* nested within Antrodiaetidae. The three known species in the synonymized genus *Atypoides* form a clade sister to "traditional" *Antrodiaetus* species (**Fig. 3**), consistent with the well-supported 4-gene results of Hendrixson & Bond (2009, figs. 1, 2). Results of character evolution and divergence time analyses are presented and discussed below.

Discussion

Arachnid UCEs. We discovered that the arachnid bait set targets and recovers mostly exons, as suggested by Bossert and Danforth (2018) for arthropod UCE baits in general (see also Branstetter et al. 2017 and Bossert et al. 2018 for hymenopterans). As such, arachnid UCE work is essentially exon capture, with flanking introns also captured for some loci. This of course has important implications for data analysis, because as we have shown here, this functional information can be used to refine analyses is various ways. Our finding also means that it might be possible to extract UCE loci from large spider / arachnid transcriptome datasets (e.g., Sharma et al. 2015; Garrison et al. 2016; Fernández et al. 2018), particularly at deeper phylogenetic levels where exon-only data would provide sufficient signal. Such a combined strategy was recently used in bee phylogenomics (Bossert et al. 2018).

Obviously, orthology is a fundamental premise in phylogenetic analyses. We found that the PHYLUCE unfiltered matrix included alignments with non-orthologs, confirmed via RAxML analysis. This "paralogy" persisted despite bioinformatic filters in place at both probe design (Faircloth, 2017) and PHYLUCE pipeline (Faircloth, 2016) stages. Our findings should not be taken as a criticism of these filters, because initial probe-design does not guarantee perfect orthology (Faircloth, 2017), and because we matched contigs to probes at liberal values (minimum coverage and minimum identity values of 65). Here we anticipated a tradeoff, as increasing this value would likely decrease non-orthology, but at the same time reduce the number of returned loci. Part of the issue is that the arachnid bait set was designed for sequence capture across all arachnids (Faircloth, 2017; Starrett et al. 2017), with a common ancestor that likely lived over 500 MYA (e.g., Rota-Stabelli, Daley & Pisani, 2013). Of all available UCE bait sets (e.g., amniotes, fish, various insects), this represents the greatest phylogenetic depth – the design of more taxon-specific bait sets within

Arachnida, in combination with more stringent probe matching values is expected to largely (but probably not entirely) alleviate issues with non-orthology.

Empirical studies have shown that large phylogenomic datasets can be misled even when a minute fraction of loci include non-orthologs (e.g., Brown & Thomson, 2017; Gatesy et al. 2018). Here analysis of the PHYLUCE unfiltered matrix (with most characters but also non-orthologs) returned trees with the same branching topology within Atypoidea as for filtered matrices (**Fig. 3**). However, these trees vary somewhat in branch support (**Fig. 3**), but importantly produce maximum likelihood topologies that differ conspicuously in estimated branch lengths (measured in nucleotide substitutions per nucleotide site). For example, estimated IQ-TREE branch lengths derived from the PHYLUCE unfiltered matrix are 1.5-3X longer than those estimated from the 70% filtered exon + intron matrix (**Fig. S1**), with both matrices produced using the same GBLOCKS settings. Exon-only trees have even shorter branch lengths (.tre files in **Data S1**), but this comparison is confounded by removal of a different class of data (faster-evolving intron sites). To the extent that branch lengths influence downstream inferences (e.g., estimates of divergence times, lineage-through-time analyses, etc.), these differences in matrix filtering could have potential analytical impacts.

We discovered that some UCE loci treated as separate alignments actually represent exons of the same protein. Via annotation, we confirmed that the conserved regions of these separate alignments represented different exons of typically very large proteins. Although unknown for the taxa studied here, these exons are likely separated by very long introns (using the known short exon- long intron structure of spiders as models, Sanggaard et al. 2014). Inclusion of “duplicate” loci should not negatively impact concatenated phylogenomic analyses. But if the exons represent a single recombinational unit, then treating duplicate alignments as independent would violate analytical assumptions of coalescent-based analyses. Also, for population-level analyses relying upon SNPs from UCE loci (e.g., Derkarabetian et al. 2018b), many commonly-used downstream analyses assume no linkage and inclusion of duplicate loci would not be justified.

To summarize, we used custom annotation and manual checking of alignments to show that 1) core regions of arachnid UCEs represent exons, 2) non-orthology sometimes persists in UCE alignments, despite upstream bioinformatic filters, 3) some “separate” loci in the arachnid bait set represent different exons of the same protein (although separated by introns of unknown length). We argue that manual checking of alignments derived from an analytical pipeline remains important (see also Bossert et al. 2018 for another UCE example). **Table S3** summarizes which UCE loci have been recovered in arachnid studies to date, and whether these loci are duplicates or potentially non-orthologous. This summary information could be used to further refine UCE analyses in arachnids, e.g., to manually adjust the published bait set to remove duplicates and paralogous loci, where non-orthology is unlikely to be rectified with more stringent probe match values. As has happened for almost all

other UCE bait sets, the refinement of the arachnid set is an expected and natural outcome of knowledge gained through empirical study.

Atypoid Phylogeny. We found strong support for the monophyly of Atypoidea (following Simon, 1892), based on a molecular phylogenetic sample with all described living genera. Our sample included the key genera *Hexurella* and *Mecicobothrium*, never previously sampled in a molecular phylogenetic analysis, and also included multiple early-diverging lineages from Avicularioidea (Hedin et al. 2018a, Opatova et al. 2019). The Atypoidea hypothesis was championed early (Chamberlin & Ivie, 1945; Coyle 1971; Coyle 1974) but ultimately fell out of favor as putative synapomorphies for the group were interpreted as plesiomorphies (Platnick, 1977; Gertsch & Platnick, 1979), and the original cladistic morphological analyses for mygalomorphs failed to recover this clade (Raven, 1985; Goloboff, 1993). However, at approximately the same time, Eskov and Zonstein (1990) argued for atypoid monophyly, and these ideas were later supported by early Sanger-based research (Hedin & Bond, 2006; Bond et al. 2012), although these molecular studies never included all described genera.

The presumed monophyly and placement of mecicobothriids is key in arguments regarding atypoid monophyly. Similar to early-diverging “diplurid” mygalomorphs, living mecicobothriid genera use elongate lateral spinnerets to build silken funnel-and-sheet webs. Platnick (1977) considered mecicobothriids to be more closely related to “diplurids” than to atypids or antrodiaetids, although he only examined *Megahexura* and *Hexura*. Similarly, Goloboff (1993) recovered mecicobothriids (scored as a single terminal) in an early-diverging grade with “diplurids”, but moving the root placement in his preferred phylogeny by one branch recovers atypoid monophyly. In this sense, both the exposed polyphyly of mecicobothriids (see below), and the phylogenomic placement of *Hexurella* and *Mecicobothrium* as ancient, early-diverging atypoids that closely straddle the primary division in mygalomorphs (**Fig. 3**), become centrally important in helping to understand past arguments over morphological homology and polarity. Proposed morphological polarities and diagnostic characters for all primary atypoid lineages are discussed below in the *Taxonomy* section.

Our phylogenomic results for all described meciobothriid genera convincingly confirm the non-monophyly of this family (**Fig. 3**). This result is consistent with prior molecular phylogenetic analyses that included *Megahexura* and *Hexura*, never recovered as sister taxa (**Fig. 2**). Mecicobothriid genera are actually morphologically heterogeneous, with each living genus displaying morphological apomorphies in somatic and genital morphology, particularly in female spermathecal morphology (see Gertsch & Platnick, 1979, Eskov & Zonsthen1990, see below). Non-monophyly and ancient divergences also help to explain the vexing biogeographic disjunction (*Hexurella, Hexura, Megahexura* from the western US; *Mecicobothrium* from southern South America) observed for included genera. Both fossil-calibrated molecular clock estimates indicate that *Hexurella* and *Mecicobothrium* stem lineages were likely present during the Triassic, well before the fragmentation of Pangea (**Fig. 4, Fig. S2**).

Cryptic Species, Webs, Parallel Diversification – Many mygalomorph genera are relatively ancient, morphologically conserved, and dispersal-limited, traits which lead to cryptic speciation. Cryptic species are common in mygalomorphs (e.g., Bond et al. 2001; Castalanelli et al. 2014; Leavitt et al. 2015), and found in the atypoids that have been examined closely, antrodiaetids in particular (Hendrixson & Bond 2007; Satler et al. 2011; Starrett et al. 2018). For example, the single described species *Antrodiaetus riversi* from central California is actually a complex of multiple cryptic species (Hedin, Starrett & Hayashi, 2013). Based on relative branch lengths recovered in phylogenomic analyses (**Fig. 3**), and estimated Cretaceous / early Tertiary ages for genera (**Fig. 4, Fig. S2**) we predict that cryptic species also occur in the Californian *Megahexura*, in *Hexurella*, and in *Hexura* from Oregon. *Hexura* is interesting in that the two described parapatric species are apparently ancient, perhaps similar to patterns seen in *Ensatina oregonensis* / *picta* salamanders from the rich mesic forests of Oregon (e.g., Kuchta et al. 2009).

Character reconstructions indicate rather unambiguously that the ancestral entrance web construct for Atypoidea is a funnel-and-sheet web (**Fig. 5**), with multiple entrance types derived from this state. Trapdoors in the antrodiaetid genus *Aliatypus* may have evolved directly from funnel-and-sheet webs, rather than from collapsible collars (*contra* Coyle, 1971). The well-supported placement of *Hexura* inside Antrodiaetidae (**Fig. 3**), as also found in the phylogenomic results of Opatova et al. (2019), is key in this character evolution inference.

We also reconstructed a funnel-and-sheet web as the ancestral state for all mygalomorphs (**Fig. 5**). Our sample for avicularioids is small, but importantly, includes all key early-diverging lineages (Bond et al. 2012; Hedin et al. 2018a; Opatova et al. 2019). Using a much more comprehensive taxon sample, Opatova et al. (2019) also reconstruct the ancestral web for avicularioids as a funnel-and-sheet web. Many authors have discussed mecicobothriid and “diplurid” web similarities as an example of convergence, for example Gertsch (1949) stated that “*the hind spinnerets of these spiders are greatly elongated …. probably an adaptation for spinning the sheet web, … illustrates how in widely unrelated creatures similar activities often lead to the production of similar morphological features*”. Instead, our phylogenomic results indicate that the funnel-and-sheet, and elongate lateral spinnerets used to produce these webs, is likely the plesiomorphic condition in mygalomorphs. One caveat is that our funnel-and-sheet scoring may be an over-simplification of homology for these taxa. For example, many early-diverging “diplurids” build massive sheet-like space webs that serve to capture prey (Coyle, 1986), features not obviously present in early-diverging atypoid webs.

Atypoid Taxonomy – Here we summarize the revised taxonomy of Atypoidea and all included families, focusing on extant taxa (summarized in **Fig. 6**). The composition of the family Mecicobothriidae is revised. *Megahexura* and *Hexurella* are removed from Mecicobothriidae and each included in new families, while *Hexura* is transferred to the family Antrodiaetidae. Also within Antrodiaetidae, the genus *Atypoides* is formally

removed from synonymy with *Antrodiaetus*. All nomenclatural changes proposed are to be attributed to Hedin and Bond.

The electronic version of this article in Portable Document Format (PDF) will represent a published work according to the International Commission on Zoological Nomenclature (ICZN), and hence the new names contained in the electronic version are effectively published under that Code from the electronic edition alone. This published work and the nomenclatural acts it contains have been registered in ZooBank, the online registration system for the ICZN. The ZooBank LSIDs (Life Science Identifiers) can be resolved and the associated information viewed through any standard web browser by appending the LSID to the prefix http://zoobank.org/. The LSID for this publication is: urn:lsid:zoobank.org:pub:A7E6FD73-9D49-4B55-911F-5D105B09A52C. The online version of this work is archived and available from the following digital repositories: PeerJ, PubMed Central and CLOCKSS."

Family Hexurellidae (NEW FAMILY) (urn:lsid:zoobank.org:act:504C322E-8FAC-4E25-806C-DCE37372112E)

Type genus. *Hexurella* Gertsch & Platnick, 1979 (urn:lsid:nmbe.ch:spidergen:00010) (type species *H. pinea* Gertsch & Platnick, 1979)

Diagnosis. As a consequence of its monogeneric status, characters used to diagnose Hexurellidae are those characters also attributed to the type genus *Hexurella*, as follows: 1) males having a gently coiled embolus (not corkscrew shaped (illustrated by Gertsch & Platnick, 1979, figures 77, 84, 87, 90); 2) posterior lateral spinnerets with four segments; and 3) spermathecae composed of a single bursal opening branching into four short, and relatively thicker bulbs (Gertsch & Platnick, 1979, figure 79). Conversely, megahexurid taxa appear to have much thinner spermathecal bulbs in which pairs share a bursal opening. As is the case for other new taxa and ranks proposed below, a more thorough examination of this new family's morphology will be an important next step in diagnosing these groups.

Distribution. Distributed in upland habitats of southern California, northern Baja California, and central/southern Arizona (Gertsch & Platnick, 1979). Undescribed species likely occur in the mountains of northern Sonora, Mexico.

Family Mecicobothriidae Holmberg, 1882 (urn:lsid:nmbe.ch:spiderfam:0003) **(new circumscription)**

Type genus. *Mecicobothrium* Holmberg , 1882 (urn:lsid:nmbe.ch:spidergen:00011) (type species *Mecicobothrium thorelli* Holmberg, 1882)

Diagnosis. Characters used to diagnose the family are those characters attributed to the type genus. Adult males of described species have a long and distinctly coiled corkscrew-shaped palpal embolus (e.g., Gertsch & Platnick, 1979, figures 45, 48, 49; Lucas et al. 2006 figures 1-3) that distinguishes members of this family from all other

atypoid taxa. Males also have a unique anterior cheliceral apophysis (Gertsch & Platnick, 1979, figures 40-42; Lucas et al. 2006 figures 20-21). Females have distinct spermathecal bulbs comprising four receptacles with the outer pair much shorter and rounder than the inner two (Gertsch & Platnick, 1979, figure 38); we note that females of *M. baccai* are unknown.

Distribution. The two described species are known from Argentina, Uruguay, and Brazil.

Family Megahexuridae (NEW FAMILY) (urn:lsid:zoobank.org:act:0D009AAF-B71C-4FFA-A580-DCD67BAA48AB)

Type genus. *Megahexura* Kaston, 1972 (urn:lsid:nmbe.ch:spidergen:00012) (type species *Hexura fulva* Chamberlin, 1919)

Diagnosis. Characters used to diagnose the family Megahexuridae are those attributed to the type genus. Members of this family can be diagnosed from other atypoid taxa by having a carapace with expanded pleurites at the posterior lateral corners (Gertsch & Platnick, 1979, figures 51, 53). Megahexurid females have spermathecae with four thin elongate bulbs, with a single receptacle opening for each pair (Gertsch & Platnick, 1979, figure 57).

Distribution. The single described species (*M. fulva)* is known from upland habitats of southern and central California (Gertsch & Platnick, 1979), although populations likely occur in northern Baja California. *Megahexura fulva* likely includes cryptic species (**Fig. 4, Fig. S2**).

Family Antrodiaetidae Gertsch,1940 (urn:lsid:nmbe.ch:spiderfam:0002) **(new circumscription)**

Type genus. *Antrodiaetus* Ausserer, 1871 (urn:lsid:nmbe.ch:spidergen:00007) (type species *Antrodiaetus unicolor* (Hentz, 1842))

List of included genera.
Aliatypus Smith, 1908 (urn:lsid:nmbe.ch:spidergen:00006)
Antrodiaetus Ausserer, 1871 (urn:lsid:nmbe.ch:spidergen:00007)
Hexura Simon, 1884 (urn:lsid:nmbe.ch:spidergen:00009)

Atypoides O. Pickard-Cambridge, 1883. (type species *Atypoides riversi* O. Pickard-Cambridge, 1883 by monotypy). Here formally removed from synonymy of *Antrodiaetus* Ausserer, 1871 contra Hendrixson and Bond 2007: 752.

List of included species.
Atypoides riversi O. Pickard-Cambridge, 1883
Atypoides hadros Coyle 1968
Atypoides gertschi Coyle 1968

Diagnosis. Adult male antrodiaetids possess a palpal bulb with a branched conductor, with inner and outer conductor sclerites (following Coyle, 1971, figure 325). The possession of this character state in *Hexura* was noted in the addendum of Eskov and Zonstein (1990), based on observations of Dr. F. Coyle, and confirmed by our study of male *Hexura* specimens.

Following Coyle (1968), the genus *Atypoides* can be distinguished from *Antrodiaetus* in having three pairs of spinnerets (Coyle, 1968, figures 30-32), with adult males possessing cheliceral apophyses (Coyle, 1968, figures 46-52). Many features separate *Atypoides* and *Antrodiaetus* from *Hexura* and *Aliatypus.*

Distribution. *Aliatypus* and *Hexura* are known from the western United States (Coyle, 1974; Gertsch & Platnick, 1979), *Atypoides* is from the western US and the southern Ozarks (Coyle, 1968; Hedin, Starrett & Hayashi, 2013), while *Antrodiaetus* includes species in Japan and more broadly in North America (Coyle, 1971; Hendrixson & Bond, 2007). Cryptic species are likely in all four genera.

Comments. Although megahexurids are sister to antrodiaetids, we do not place them in the same family for three primary reasons. First, these families share a common ancestor that likely existed over 200 million years ago (**Fig. 4**). This level of divergence would exceed any intra-familial divergence in described mygalomorph families (see Opatova et al. 2019). Second, these families differ in important diagnostic characters, including female spermathecal morphology, but importantly megahexurid males lack the key antrodiaetid palpal bulb with diagnostic inner and outer conductor sclerites (**Fig. 6**).

Conversely, one could argue that *Aliatypus* and *Hexura* each deserve family-level status (the latter an available family rank name, Hexurinae Simon 1889), sister to other antrodiaetids. Again, although heterogenous from a web construct perspective (**Fig. 5**), antrodiaetids share morphological synapomorphies, with a level of inter-generic temporal divergence comparable to other described mygalomorph families (**Fig. 4,** Opatova et al. 2019).

Conclusions

Early-diverging atypoid lineages are ancient, often species-poor (approximating monotypic), and use silk to build funnel-and-sheet webs. The evolution of more diverse silken entrance constructs is found in more derived atypoid lineages. Similar patterns of species-poor early-diverging lineages, and diverse entrance constructs evolving in more derived lineages occurs in parallel in the avicularioid mygalomorphs (Opatova et al. 2019). In this sense, atypoids and avicularioids represent comparable evolutionary experiments, although the latter clade has clearly evolved a greater diversity of taxa, morphologies, and web constructs. How the competitive interplay of these parallel lineages has impacted diversification dynamics in deep time would be an interesting topic for further study.

Acknowledgements

We thank Nobuo Tsurusaki, Dean Leavitt, Jordan Satler and in particular Jim Starrett for help with specimen collection. Gabriel Pompozzi provided images of *Mecicobothrium*. Comments of Brent Hendrixson, D. Leavitt, J. Starrett and an anonymous reviewer helped to improve the manuscript.

References

Ayoub NA, Garb JE, Hedin M, Hayashi CY (2007) Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1γ), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). Molecular Phylogenetics and Evolution 42, 394–409.

Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Matttick JS, Haussler D (2004) Ultraconserved elements in the human genome. Science 304, 1321–1325.

Branstetter MG, Longino JT, Ward PS, Faircloth BC (2017) Enriching the ant tree of life: enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods in Ecology and Evolution 8, 768-776.

Bond JE, Hedin MC, Ramirez MG, Opell BD (2001) Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Molecular Ecology 10, 899-910.

Bond JE, Hendrixson BE, Hamilton CA, 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(6), e38753.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology, 10(4), e1003537.

Borowiec ML (2019) Convergent evolution of the army ant syndrome and congruence in big-data phylogenetics Systematic Biology, *In Press*.

Bossert S, Danforth BN (2018) On the universality of target-enrichment baits for phylogenomic research. Methods in Ecology and Evolution 00:1-8. doi: 10.1111/2041-210X.12988.

Bossert S, Murray EA, Almeida EAB, Brady SG, Blaimer BB, Danforth BN (2018) Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. Molecular Phylogenetics and Evolution https://doi.org/10.1016/j.ympev.2018.10.012.

Brown JM, Thomson RC (2017) Bayes Factors unmask highly variable information content, bias, and extreme influence in phylogenomic analyses. Systematic Biology 66, 517–530.

Castalanelli MA, Teale R, Rix MG, Kennington WJ, Harvey MS (2014) Barcoding of mygalomorph spiders (Araneae: Mygalomorphae) in the Pilbara bioregion of Western Australia reveals a highly diverse biota. Invertebrate Systematics 28, 375–385.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology Evolution 17, 540–552

Chamberlin RV, Ivie W (1945) On some Nearctic mygalomorph spiders. Annals of the Entomological Society of America 38, 549-558.

Chernomor O, von Haeseler A, Minh BQ (2016) Terrace aware data structure for phylogenomic inference from supermatrices. Systematic Biology 65, 997-1008.

Chifman J, Kubatko L (2014) Quartet inference from SNP data under the coalescent. Bioinformatics 30, 3317–3324.

Chifman J, Kubatko L (2015) Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. Journal of Theoretical Biology 374, 35–47.
Clarke TH, Garb JE, Hayashi CY, Arensburger P, Ayoub NA (2015) Spider transcriptomes identify ancient large-scale gene duplication event potentially important in silk gland evolution. Genome Biology and Evolution 7, 1856–1870.
Costa FG, Pérez-Miles F (1998) Behavior, life cycle and webs of *Mecicobothrium thorelli* (Araneae, Mygalomorphae, Mecicobothriidae). Journal of Arachnology 26, 317-329.
Coyle FA (1968) The mygalomorph spider genus *Atypoides* (Araneae: Antrodiaetidae). Psyche, Cambridge 75, 157-194.
Coyle FA (1971) Systematics and natural history of the mygalomorph spider genus *Antrodiaetus* and related genera (Araneae: Antrodiaetidae). Bulletin of the Museum of Comparative Zoology 141, 269-402.
Coyle FA (1974) Systematics of the trapdoor spider genus *Aliatypus* (Araneae: Antrodiaetidae). Psyche, Cambridge 81, 431-500.
Coyle FA (1986) The role of silk in prey capture by nonaraneomorph spiders. In: Shear WA, ed. *Spiders: Webs, Behavior, and Evolution*. Stanford University Press. 269-305.
Dalla Vecchia FM, Selden PA (2013) A Triassic spider from Italy. Acta Palaeontologica Polonica 58, 325–330.
Derkarabetian S, Starrett J, Tsurusaki N, Ubick D, Castillo S, Hedin M (2018a) A stable phylogenomic classification of Travunioidea (Arachnida, Opiliones, Laniatores) based on sequence capture of ultraconserved elements. ZooKeys (760), 1–36.
Derkarabetian S, Castillo S, Koo PK, Ovchinnikov S, Hedin M (2018b). An empirical demonstration of unsupervised machine learning in species delimitation. bioRxiv doi: https://doi.org/10.1101/429662.
Eskov KY, Zonstein S (1990) First Mesozoic mygalomorph spiders from the Lower Cretaceous of Siberia and Mongolia, with notes on the system and evolution of the infraorder Mygalomorphae (Chelicerata: Araneae). Neues Jahrbuch für Mineralogie, Geologie und Paläontologie, Abhandlungen 178, 325-368.
Faircloth BC (2013) illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. http://dx.doi.org/10.6079/J9ILL
Faircloth BC (2016) PHYLUCE is a software package for the analysis of conserved genomic loci**.** Bioinformatics 32, 786–788.
Faircloth BC (2017) Identifying conserved genomic elements and designing universal probe sets to enrich them**.** Methods in Ecology and Evolution 8, 1103–1112.
Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61, 717–726.
Faircloth BC, Branstetter MG, White ND, Brady SG (2015) Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Molecular Ecology Resources 15, 489–501.
Fernández R, Kallal RJ, Dimitrov D, Ballesteros JA, Arnedo MA, Giribet G, Hormiga G (2018) Phylogenomics, diversification dynamics, and comparative transcriptomics across the Spider Tree of Life. Current Biology 28, 1489-1497.e5

704 Fourie R, Haddad CR, Jocqué R (2011). A revision of the purse-web spider genus
705 *Calommata* Lucas, 1837 (Araneae, Atypidae) in the Afrotropical region. ZooKeys
706 95, 1-28.
707 Garrison NL, Rodriguez J, Agnarsson I, Coddington JA, Griswold CE, Hamilton CA,
708 Hedin M, Kocot KM, Ledford JM, Bond JE (2016) Spider Phylogenomics:
709 Untangling the Spider Tree of Life. PeerJ 4:e1719; DOI 10.7717/peerj.1719.
710 Gatesy J, Sloan DB, Warren JM, Baker RH, Simmons MP, Springer MS (2018)
711 Partitioned coalescence support reveals biases in species-tree methods and
712 detects gene trees that determine phylogenomic conflicts. bioRxiv doi:
713 http://dx.doi.org/10.1101/461699.
714 Gertsch WJ (1949) *American Spiders*. D. van Nostrand, New York.
715 Gertsch WJ, Platnick NI (1979) A revision of the spider family Mecicobothriidae
716 (Araneae, Mygalomorphae). American Museum Novitates 2687, 1-32.
717 Goloboff P (1993) A reanalysis of mygalomorph spider families (Araneae). American
718 Museum Novitates 3056, 1–32
719 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
720 Raychowdhury R, Zeng Q, Chen Z (2011) Full-length transcriptome assembly from
721 RNA-Seq data without a reference genome. Nature Biotechnology 29, 644.
722 Hamilton CA, Lemmon AR, Lemmon EM, Bond JE (2016) Expanding anchored hybrid
723 enrichment to resolve both deep and shallow relationships within the spider tree of
724 life. BMC Evolutionary Biology 16, 212.
725 Hedin M, Bond JE (2006) Molecular phylogenetics of the spider infraorder
726 Mygalomorphae using nuclear rRNA genes (18S and 28S): Conflict and agreement
727 with the current system of classification. Molecular Phylogenetics and Evolution 41,
728 454–471.
729 Hedin M, Starrett J, Hayashi C (2013) Crossing the uncrossable: novel trans-valley
730 biogeographic patterns revealed in the genetic history of low-dispersal
731 mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California. Molecular
732 Ecology 22, 508–526.
733 Hedin M, Derkarabetian S, Ramírez M, Vink C, Bond J (2018a) Phylogenomic
734 reclassification of the world's most venomous spiders (Mygalomorphae, Atracinae),
735 with implications for venom evolution. Scientific Reports 8, 1636 D.
736 Hedin M, Derkarabetian S, Blair J, Paquin P (2018b) Sequence capture phylogenomics
737 of eyeless *Cicurina* spiders from Texas caves, with emphasis on US federally-
738 endangered species from Bexar County (Araneae, Hahniidae). ZooKeys 769, 49–
739 76.
740 Hendrixson BE, Bond JE (2007) Molecular phylogeny and biogeography of an ancient
741 Holarctic lineage of mygalomorph spiders (Araneae: Antrodiaetidae: *Antrodiaetus*).
742 Molecular Phylogenetics and Evolution 42, 738-755.
743 Hendrixson BE, Bond JE (2009) Evaluating the efficacy of continuous quantitative
744 characters for reconstructing the phylogeny of a morphologically homogeneous
745 spider taxon (Araneae, Mygalomorphae, Antrodiaetidae, *Antrodiaetus*). Molecular
746 Phylogenetics and Evolution. 53, 300-13.
747 Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2:
748 Improving the ultrafast bootstrap approximation. Molecular Biology and Evolution
749 35, 518–522.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nature Methods 14, 587-589.
Katoh D, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30, 772–780.
Kieran TJ, Gordon ERL, Forthman M, Hoey-Chamberlain R, Kimball RT, Faircloth BC, Weirauch C, Glenn TC (2018) Insight from an Ultraconserved element bait set designed for Hemipteran phylogenetics integrated with genomic resources. Molecular Phylogenetics and Evolution https://doi.org/10.1016/j.ympev.2018.10.026.
Kuchta SR, Parks DS, Mueller RL, Wake DB (2009) Closing the ring: historical biogeography of the salamander ring species *Ensatina eschscholtzii*. Journal of Biogeography 36, 982-995.
Lartillot N, Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Molecular Biology and Evolution 21, 1095-1109.
Leavitt DH, Starrett J, Westphal MF, Hedin M (2015) Multilocus sequence data reveal dozens of putative cryptic species in a radiation of endemic Californian mygalomorph spiders (Araneae, Mygalomorphae, Nemesiidae). Molecular Phylogenetics and Evolution 91, 56-67.
Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic Biology 50, 913-925.
Lucas SM, Indicatti RP, Brescovit AD Francisco RC (2006) First record of the Mecicobothriidae Holmberg from Brazil, with a description of a new species of *Mecicobothrium* (Araneae, Mygalomorphae). Zootaxa 1326, 45-53.
Maddison WP, Maddison DR (2018) Mesquite: a modular system for evolutionary analysis. Version 3.51 http://www.mesquiteproject.org
McCole RB, Erceg J, Saylor W, Wu C-T (2018) Ultraconserved elements occupy specific arenas of three-dimensional mammalian genome organization. Cell Reports 24, 479–488.
Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Molecular Biology and Evolution 32, 268-274.
Opatova V, Hamilton CA, Hedin M, Montes de Oca L, Král J, Haddad CR, Bond JE (2019) Reevaluation of the systematics and classification of the spider infraorder Mygalomorphae using genomic scale data. bioRxiv doi: https://doi.org/10.1101/531756
Platnick NI (1977) The hypochiloid spiders: a cladistic analysis, with notes on the Atypoidea (Arachnida, Araneae). American Museum Novitates 2627, 1-23.
Polychronopoulos D, King JWD, Nash AJ, Tan G, Lenhard B (2017) Conserved non-coding elements: developmental gene regulation meets genome organization, Nucleic Acids Research 45, 12611–12624.
Raven RJ (1985) The spider infraorder Mygalomorphae (Araneae): Cladistics and systematics. Bulletin American Museum of Natural History 182, 1–180.

795 Raven RJ, Jell PA, Knezour RA (2015) *Edwa maryae* gen. et sp. nov. in the Norian
796 Blackstone Formation of the Ipswich Basin—the first Triassic spider
797 (Mygalomorphae) from Australia. Alcheringa 39, 259-263.
798 Rota-Stabelli O, Daley AC, Pisani D (2013) Molecular timetrees reveal a Cambrian
799 colonization of land and a new scenario for ecdysozoan evolution. Current Biology
800 23, 392-8.
801 Sanggaard KW, Bechsgaard JS, Fang X, Duan J, Dyrlund TF, Gupta V, Jiang X, Cheng
802 L, Fan D, Feng Y, Han L (2014) Spider genomes provide insight into composition
803 and evolution of venom and silk. Nature Communications 5, 3765.
804 Satler JD, Starrett J, Hayashi CY, Hedin M (2011) Inferring species trees from gene
805 trees in a radiation of California trapdoor spiders (Araneae, Antrodiaetidae,
806 *Aliatypus*). PLoS One, 6(9) doi:http://dx.doi.org/10.1371/journal.pone.0025355
807 Schwager EE, Sharma PP, Clarke T, Leite DJ, Wierschin T, Pechmann M, Akiyama-
808 Oda Y, Esposito L, Arensburger P, Bechsgaard J, Bilde T, Buffry A, Chao H, Dinh
809 H, Doddapaneni H, Dugan S, Eibner C, Extavour CG, Funch P, Garb J, Gonzalez
810 VL, Griffiths-Jones S, Han Y, Hayashi C, Hilbrant M, Hughes DST, Janssen R, Lee
811 SL, Maeso I, Murali SC, Muzny DM, da Fonseca RN, Qu J, Ronshaugen M,
812 Schomburg C, Schönauer A, Stollewerk A, Torres-Oliva M, Turetzek N,
813 Vanthournout B, Werren J, Wolff C, Worley KC, Gibbs RA, Coddington J, Oda H,
814 Stanke M, Ayoub NA, Damen WGM, Prpic N-M, Flot J-F, Posnien N, Richards S,
815 McGregor AP (2017) The house spider genome reveals a whole genome
816 duplication during arachnid evolution. BMC Biology 15, 62.
817 Schwendinger PJ (1990) A synopsis of the genus *Atypus* (Araneae, Atypidae).
818 Zoologica Scripta 19, 353-366.
819 Selden PA, Gall J-C (1992) A Triassic mygalomorph spider from the northern Vosges,
820 France. Palaeontology 35, 211–235.
821 Selden PA, Shear WA, Sutton MD (2008) Fossil evidence for the origin of spider
822 spinnerets, and a proposed arachnid order. Proceedings of the National Academy
823 of Sciences of the United States of America 105, 20781–20785.
824 Sharma PP, Fernández R, Esposito L, González-Santillán E, Monod L (2015)
825 Phylogenomic resolution of scorpions reveals multilevel discordance with
826 morphological phylogenetic signal. Proceedings of the Royal Society of London B:
827 Biological Sciences 282, 20142953.
828 Simon E (1892) *Histoire naturelle des araignées*. Paris 1, 1-256.
829 Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-
830 analysis of large phylogenies. Bioinformatics 30, 1312–1313.
831 Starrett J, Hedin M, Ayoub N, Hayashi CY (2013) Hemocyanin gene family evolution in
832 spiders (Araneae), with implications for phylogenetic relationships and divergence
833 times in the infraorder Mygalomorphae. Gene 524, 175-186.
834 Starrett J, Derkarabetian S, Hedin M, Bryson Jr. RW, McCormack JE, Faircloth BC (
835 2017) High phylogenetic utility of an Ultraconserved element probe set designed
836 for Arachnida. Molecular Ecology Resources 17, 812-823.
837 Starrett J, Hayashi CY, Derkarabetian S, Hedin M (2018) Cryptic elevational zonation in
838 trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus janus* complex) from the
839 California southern Sierra Nevada. Molecular Phylogenetics and Evolution 118,
840 403-413.

841 Swofford DL (2003) PAUP*: phylogenetic analysis using parsimony, version 4.0a163.
842 Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent
843 and ambiguously aligned blocks from protein sequence alignments. Systematic
844 Biology 56, 564–577.
845 Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of
846 molecular evolution. Molecular Biology and Evolution 15, 1647-1657.
847 Tin MM-Y, Economo EP, Mikheyev AS (2014) Sequencing degraded DNA from non-
848 destructively sampled museum specimens for RAD-Tagging and low-coverage
849 shotgun phylogenetics. PLoS ONE 9, e96793.
850 https://doi.org/10.1371/journal.pone.0096793
851 Wheeler WC, Coddington JA, Crowley LM, Dimitrov D, Goloboff PA, Griswold CE,
852 Hormiga G, Prendini L, Ramírez MJ, Sierwald P, Almeida-Silva L, Alvarez-Padilla
853 F, Arnedo MA, Benavides Silva LR, Benjamin SP, Bond JE, Grismado CJ, Hasan
854 E, Hedin M, Izquierdo MA, Labarque FM, Ledford J, Lopardo L, Maddison WP,
855 Miller JA, Piacentini LN, Platnick NI, Polotow D, Silva-Dávila D, Scharff N, Szűts T,
856 Ubick D, Vink CJ, Wood HM, Zhang J (2017) The spider tree of life: phylogeny of
857 Araneae based on target-gene analyses from an extensive taxon sampling.
858 Cladistics 33, 574-616. doi:10.1111/cla.12182.
859 Wood HM, Matzke NJ, Gillespie RG, Griswold CE (2012) Treating fossils as terminal
860 taxa in divergence time estimation reveals ancient vicariance patterns in the
861 palpimanoid spiders. Systematic Biology 62, 264-284.
862 Wood HM, González VL, Lloyd M, Coddington J, Scharff N (2018) Next-generation
863 museum genomics: Phylogenetic relationships among palpimanoid spiders using
864 sequence capture techniques (Araneae: Palpimanoidea). Molecular Phylogenetics
865 and Evolution 127, 907-918.
866 World Spider Catalog (2019). World Spider Catalog. Version 20. Natural History
867 Museum Bern, online at http://wsc.nmbe.ch, accessed on 25 Feb 2019. doi:
868 10.24436/2
869 Yang Z, Rannala B (2006). Bayesian estimation of species divergence times under a
870 molecular clock using multiple fossil calibrations with soft bounds. Molecular
871 Biology Evolution 23, 212-226.
872 Zerbino DR, Birney E (2008) Velvet: Algorithms for de novo short read assembly using
873 de Bruijn graphs. Genome Research 18, 821–829.
874

875 **Figure Legends**
876
877 **Figure 1.** Images of live animals and entrance web constructs, not to scale. A)
878 *Hexurella apachea*, Cochise County, AZ. MCH 18_029. B) *Mecicobothrium thorelli*,
879 image by G. Pompozzi. C) *Atypus karschi*. Honshu, Tottori, Japan. MCH 15_016. D)
880 *Megahexura fulva*, Fresno County, CA. MCH 09_018. E) *Aliatypus californicus*. Contra
881 Costa County, CA. MCH 10_031. F) *Hexura picea*. Lincoln County, OR. MCH 14_040.
882 G) *Antrodiaetus unicolor*, Jackson County, NC. H) *Atypoides (= Antrodiaetus) riversi*,
883 San Mateo County, CA. MCH 10_015. Arrows point to dorsal abdominal tergites in
884 images B, E and G. All photos (other than *Mecicobothrium*) by M. Hedin.
885
886 **Figure 2.** Summary of previous molecular phylogenetic analyses including members of
887 Atypoidea. References as in text.
888
889 **Figure 3.** Tree topology from RAxML concatenated analysis, based on filtered exon +
890 intron matrix. Support values for all nine analyses either indicated directly, or by circles
891 at nodes (when exceeding 95 or 0.95 for all). Support values within (*Namirea, Euagrus,*
892 *Bymainiella, Calisoga*) clade not shown, as these relationships vary across analyses
893 (see .tre files in **Data S1**). Also, two low support nodes not shown for PHYLUCE unfiltered
894 SVD results, as follows: *Antrodiaetus apachecus* + *A. roretzi* (67), *Antrodiaetus hadros*
895 + *A. gertschi* (68).
896
897 **Figure 4.** Chronogram derived from Phylobayes analyses, estimated using calibration
898 with minimum age for the root node of Atypoidea = 210 MYA. HPD values in brackets.
899 Results using the alternative calibration included as **Fig. S2**. Geological times from
900 http://www.geosociety.org/documents/gsa/timescale/timescl.pdf.
901
902 **Figure 5.** Ancestral character reconstruction for entrance web constructs. Proportional
903 likelihood values for funnel-and-sheet web shown at internal nodes.
904
905 **Figure 6.** Summary of new taxonomy and diagnostic morphological characters. See text
906 for references and explanation of terms.
907

Figure 1

Images of live animals and entrance web constructs

A) *Hexurella apachea*, Cochise County, AZ. MCH 18_029. B) *Mecicobothrium thorelli*, image by G. Pompozzi. C) *Atypus karschi*. Honshu, Tottori, Japan. MCH 15_016. D) *Megahexura fulva*, Fresno County, CA. MCH 09_018. E) *Aliatypus californicus*. Contra Costa County, CA. MCH 10_031. F) *Hexura picea*. Lincoln County, OR. MCH 14_040. G) *Antrodiaetus unicolor*, Jackson County, NC. H) *Atypoides (= Antrodiaetus) riversi*, San Mateo County, CA. MCH 10_015. Arrows point to dorsal abdominal tergites in images B, E and G. All photos (other than *Mecicobothrium*) by M. Hedin.

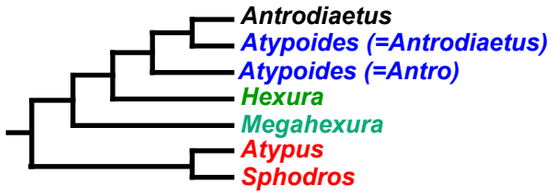

A) Hexurella
B) Mecicobothrium
C) Atypus
D) Megahexura
E) Aliatypus
F) Hexura
G) Antrodiaetus
H) Atypoides (=Antrodiaetus)

Figure 2(on next page)

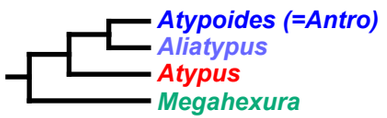
Summary of previous molecular phylogenetic analyses including members of Atypoidea.

References as in text.

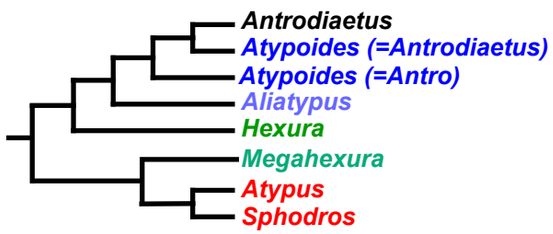
Hedin & Bond 2006, FIG 5



Ayoub et al 2007, FIG 3



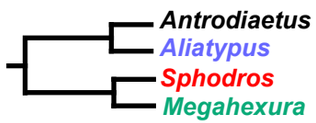
Bond et al 2012, FIG 1



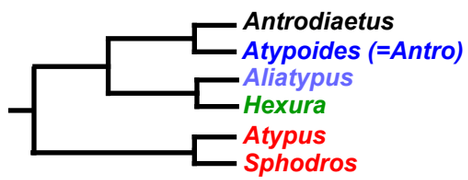
Starrett et al 2013, FIG 5

Hamilton et al 2016, FIG 4

Garrison et al 2016, FIG 2



Wheeler et al 2017, FIG 2



Hedin et al 2018a, FIG 2

Fernandez et al 2018, FIG 1

Opatova et al 2019, FIG 3

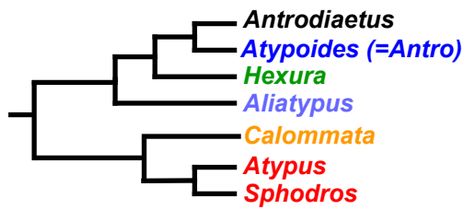


Figure 3(on next page)

Tree topology from RAxML concatenated analysis, based on filtered exon + intron matrix.

Support values for all nine analyses either indicated directly, or by circles at nodes (when exceeding 95 or 0.95 for all). Support values within (*Namirea, Euagrus, Bymainiella, Calisoga*) clade not shown, as these relationships vary across analyses (see .tre files in Data S1). Also, two low support nodes not shown for Phyluce unfiltered SVD results, as follows: *Antrodiaetus apachecus* + *A. roretzi* (67), *Antrodiaetus hadros* + *A. gertschi* (68).

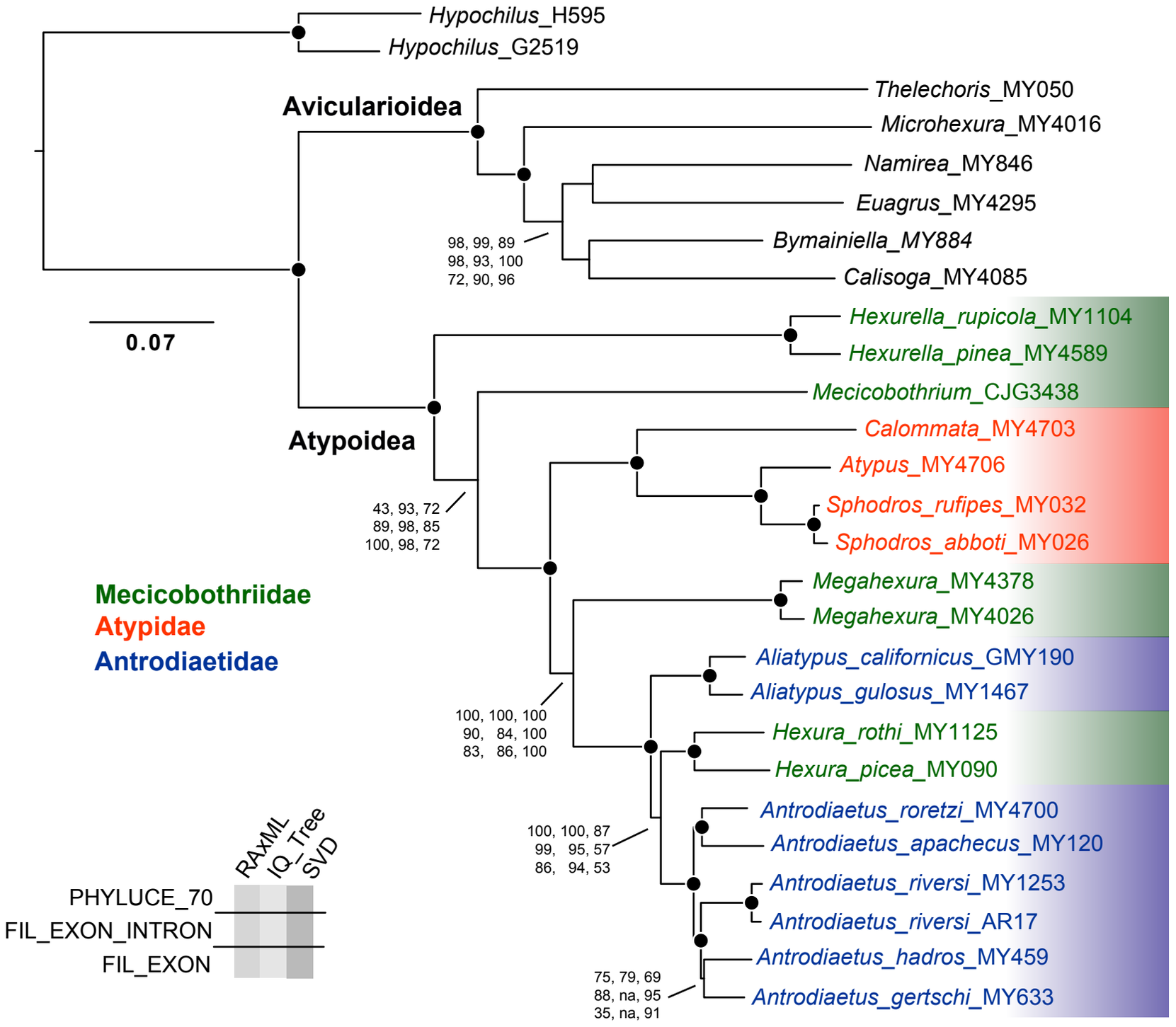

Hypochilus_H595
Hypochilus_G2519
Avicularioidea
Thelechoris_MY050
Microhexura_MY4016
Namirea_MY846
Euagrus_MY4295
98, 99, 89
98, 93, 100
72, 90, 96
Bymainiella_MY884
Calisoga_MY4085
0.07
Hexurella_rupicola_MY1104
Hexurella_pinea_MY4589
Mecicobothrium_CJG3438
Atypoidea
Calommata_MY4703
Atypus_MY4706
43, 93, 72
89, 98, 85
100, 98, 72
Sphodros_rufipes_MY032
Sphodros_abboti_MY026
Mecicobothriidae
Atypidae
Antrodiaetidae
Megahexura_MY4378
Megahexura_MY4026
Aliatypus_californicus_GMY190
Aliatypus_gulosus_MY1467
100, 100, 100
90, 84, 100
83, 86, 100
Hexura_rothi_MY1125
Hexura_picea_MY090
Antrodiaetus_roretzi_MY4700
100, 100, 87
99, 95, 57
86, 94, 53
Antrodiaetus_apachecus_MY120
Antrodiaetus_riversi_MY1253
Antrodiaetus_riversi_AR17
Antrodiaetus_hadros_MY459
75, 79, 69
88, na, 95
35, na, 91
Antrodiaetus_gertschi_MY633
RAxML
IQ_Tree
SVD
PHYLUCE_70
FIL_EXON_INTRON
FIL_EXON

Figure 4(on next page)

Chronogram derived from Phylobayes analyses, estimated using calibration with minimum age for the root node of Atypoidea = 210 MYA.

HPD values in brackets. Results using the alternative calibration included as Fig. S2.

Geological times from http://www.geosociety.org/documents/gsa/timescale/timescl.pdf .

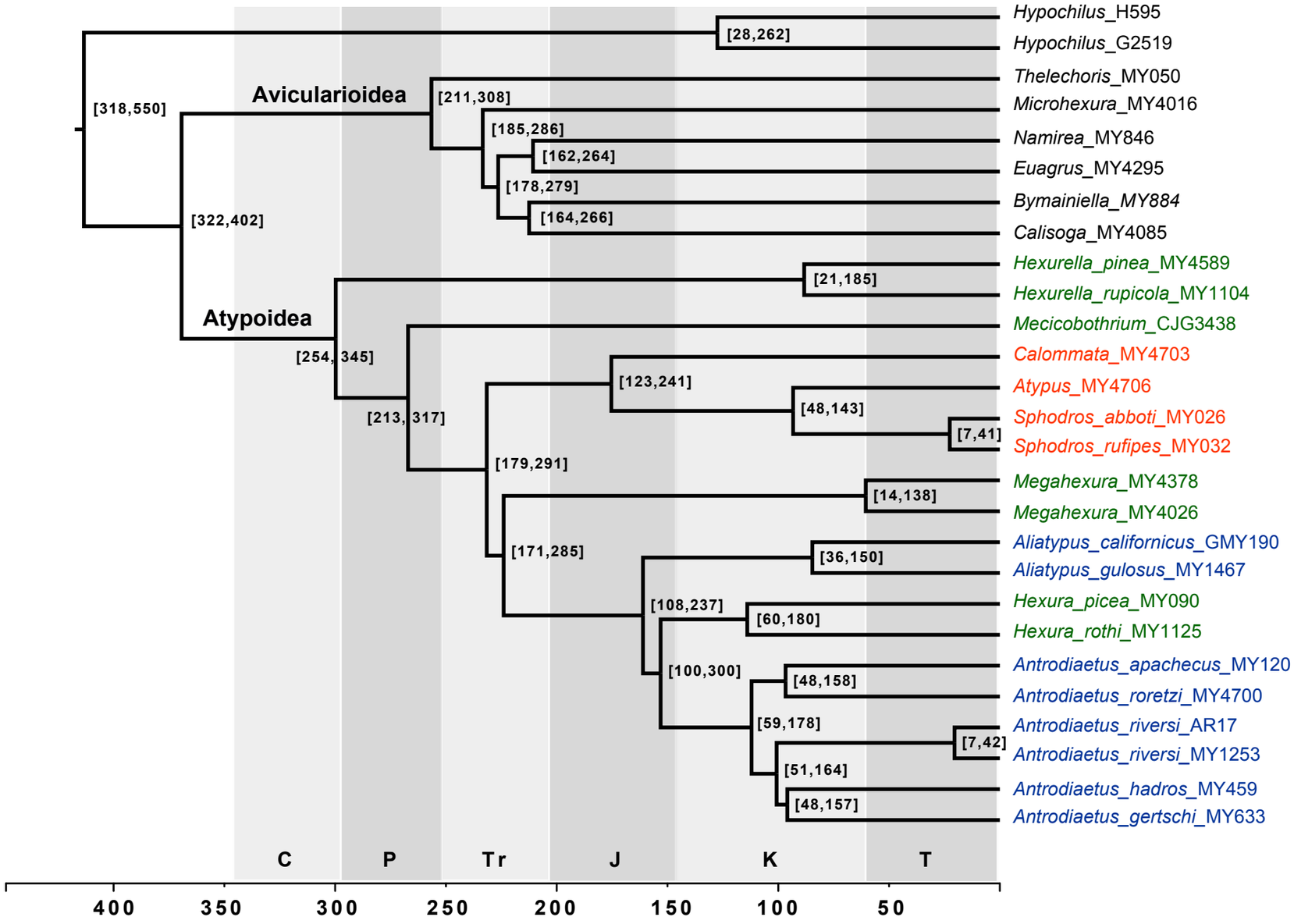

Avicularioidea
Atypoidea
[318,550]
[322,402]
[28,262]
[211,308]
[185,286]
[162,264]
[178,279]
[164,266]
[21,185]
[254, 345]
[213, 317]
[123,241]
[48,143]
[7,41]
[179,291]
[14,138]
[171,285]
[36,150]
[108,237]
[60,180]
[100,300]
[48,158]
[59,178]
[7,42]
[51,164]
[48,157]
Hypochilus_H595
Hypochilus_G2519
Thelechoris_MY050
Microhexura_MY4016
Namirea_MY846
Euagrus_MY4295
Bymainiella_MY884
Calisoga_MY4085
Hexurella_pinea_MY4589
Hexurella_rupicola_MY1104
Mecicobothrium_CJG3438
Calommata_MY4703
Atypus_MY4706
Sphodros_abboti_MY026
Sphodros_rufipes_MY032
Megahexura_MY4378
Megahexura_MY4026
Aliatypus_californicus_GMY190
Aliatypus_gulosus_MY1467
Hexura_picea_MY090
Hexura_rothi_MY1125
Antrodiaetus_apachecus_MY120
Antrodiaetus_roretzi_MY4700
Antrodiaetus_riversi_AR17
Antrodiaetus_riversi_MY1253
Antrodiaetus_hadros_MY459
Antrodiaetus_gertschi_MY633
C
P
Tr
J
K
T
400
350
300
250
200
150
100
50

Figure 5(on next page)

Ancestral character reconstruction for entrance web constructs.

Proportional likelihood values for funnel-and-sheet web shown at internal nodes.

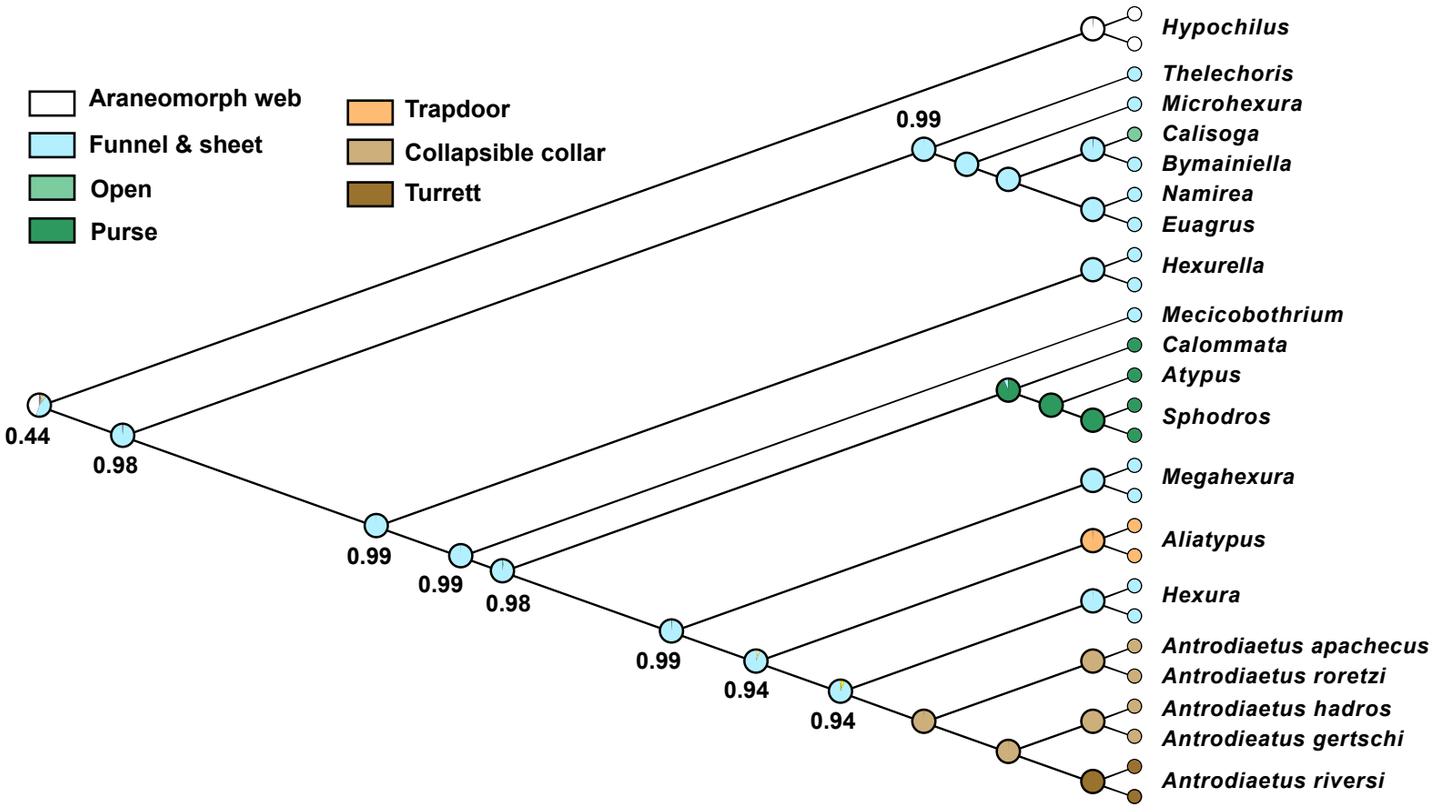
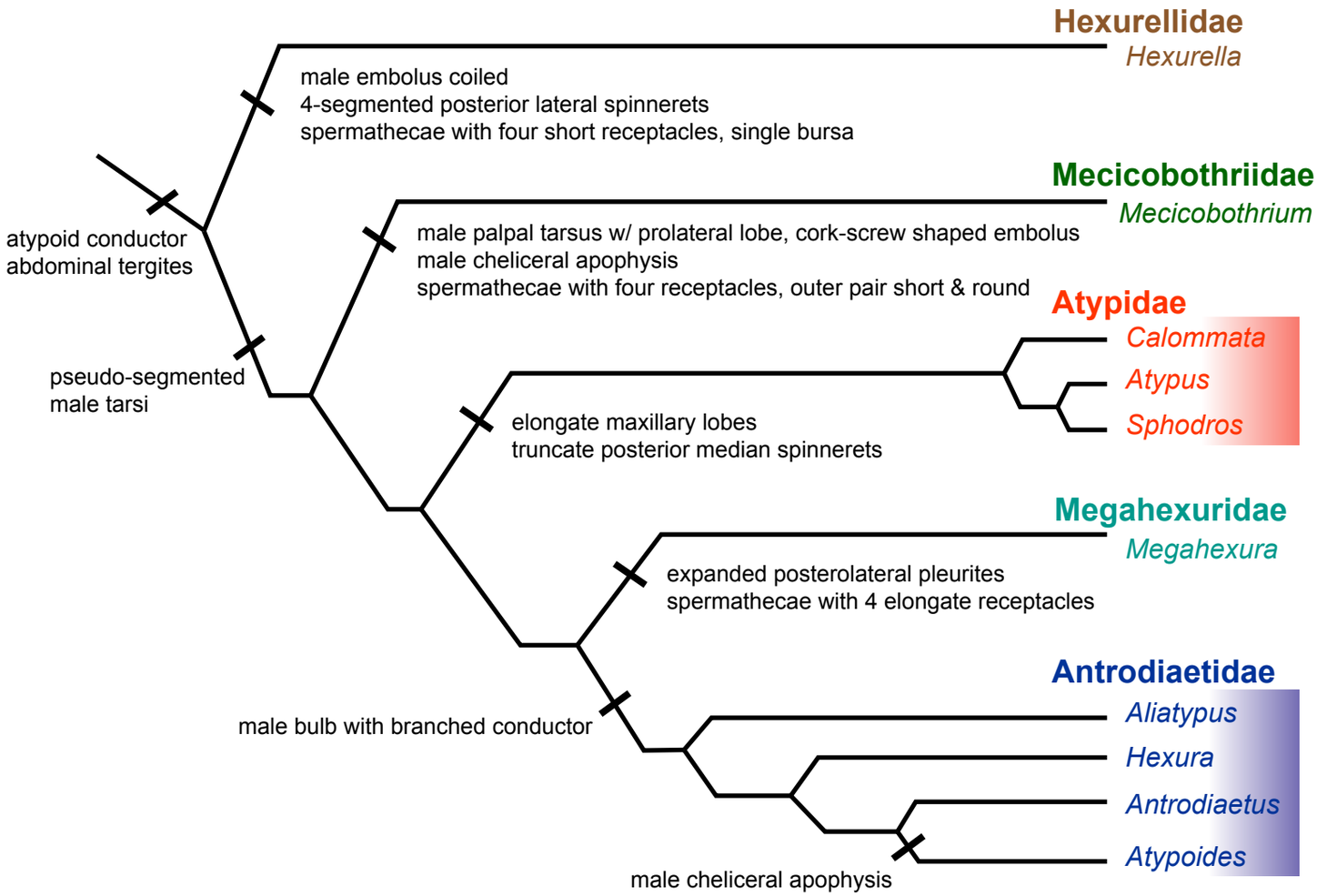

Araneomorph web
Funnel & sheet
Open
Purse
Trapdoor
Collapsible collar
Turrett
0.44
0.98
0.99
0.99
0.98
0.99
0.94
0.94
0.99
Hypochilus
Thelechoris
Microhexura
Calisoga
Bymainiella
Namirea
Euagrus
Hexurella
Mecicobothrium
Calommata
Atypus
Sphodros
Megahexura
Aliatypus
Hexura
Antrodiaetus apachecus
Antrodiaetus roretzi
Antrodiaetus hadros
Antrodieatus gertschi
Antrodiaetus riversi

Figure 6(on next page)

Summary of new taxonomy and diagnostic morphological characters.

See text for references and explanation of terms.


Hexurellidae
Hexurella
male embolus coiled
4-segmented posterior lateral spinnerets
spermathecae with four short receptacles, single bursa
Mecicobothriidae
Mecicobothrium
atypoid conductor
abdominal tergites
male palpal tarsus w/ prolateral lobe, cork-screw shaped embolus
male cheliceral apophysis
spermathecae with four receptacles, outer pair short & round
Atypidae
Calommata
Atypus
Sphodros
pseudo-segmented
male tarsi
elongate maxillary lobes
truncate posterior median spinnerets
Megahexuridae
Megahexura
expanded posterolateral pleurites
spermathecae with 4 elongate receptacles
Antrodiaetidae
male bulb with branched conductor
Aliatypus
Hexura
Antrodiaetus
Atypoides
male cheliceral apophysis