Spider phylogenomics: untangling the Spider Tree of Life

Nicole L Garrison, Juanita Rodriguez, Ingi Agnarsson, Jonathan A Coddington, Charles E Griswold, Christopher A Hamilton, Marshal Hedin, Kevin M Kocot, Joel M Ledford, Jason E Bond

Spiders (Order Araneae) are massively abundant generalist arthropod predators that are found in nearly every ecosystem on the planet and have persisted for over 380 million years. Spiders have long served as evolutionary models for studying complex mating and web spinning behaviors, key innovation and adaptive radiation hypotheses, and have been inspiration for important theories like sexual selection by female choice. Unfortunately, past major attempts to reconstruct spider phylogeny typically employing the “usual suspect” genes have been unable to produce a well-supported phylogenetic framework for the entire order. To further resolve spider evolutionary relationships we have assembled a transcriptome-based data set comprising 70 ingroup spider taxa. Using maximum likelihood and shortcut coalescence-based approaches, we analyze eight data sets, the largest of which contains 3,398 gene regions and 696,652 amino acid sites forming the largest phylogenomic analysis of spider relationships produced to date. Contrary to long held beliefs that the orb web is the crowning achievement of spider evolution, ancestral state reconstructions of web type support a phylogenetically ancient origin of the orb web and diversification analyses show that the mostly ground-dwelling, web-less RTA clade diversified faster than orb weavers. Consistent with molecular dating estimates we report herein, this may reflect a major increase in biomass of non-flying insects during the Cretaceous Tertiary Revolution 125-90 million years ago favoring diversification of spiders that feed on cursorial rather than flying prey. Our results also have major implications for our understanding of spider systematics. Phylogenomic analyses corroborate several well-accepted high level groupings: Opisthothele, Mygalomorphae, Atypoidina, Avicularioidea, Theraphosidina, Araneomorphae, Entelygynae, Araneoidea, the RTA – clade, Dionycha and the Lycosoidea. Alternatively, our results challenge the monophyly of Eresoidea, Orbiculariae, and Deinopoidea. The composition of the major Paleocribellate and Neocribellate clades, the basal divisions of Araneomorphae, appear to be falsified. Traditional Haplogynae, and even the new concept of Synspermiata, need revision after the departure of Filistatidae and Leptonetidae from the haplogyne clade. The sister pairing of filistatids with hypochilids, implies that some peculiar features of each family may in fact be synapomorphic for the pair. Leptonetids now are seen as a possible sister group to the Entelegynae, illustrating possible intermediates in the evolution of the more complex

entelegyne genitalic condition, spinning organs and respiratory organs.

Spider Phylogenomics: Untangling the Spider Tree of Life

Nicole L. Garrison[1], Juanita Rodriguez[1], Ingi Agnarsson[2], Jonathan A. Coddington[3], Charles E. Griswold[4], Chris A. Hamilton[1], Marshal Hedin[5], Kevin M. Kocot[6], Joel M. Ledford[7], and Jason E. Bond*[1]

[1]**Department of Biological Sciences and Auburn University Museum of Natural History, Auburn University**
[2]**Department of Biology, University of Vermont**
[3]**Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC**
[4]**Arachnology, California Academy of Sciences**
[5]**Department of Biology, San Diego State University**
[6]**Department of Biological Sciences, University of Alabama**
[7]**Department of Plant Biology, University of California Davis**

ABSTRACT

Spiders (Order Araneae) are massively abundant generalist arthropod predators that are found in nearly every ecosystem on the planet and have persisted for over 380 million years. Spiders have long served as evolutionary models for studying complex mating and web spinning behaviors, key innovation and adaptive radiation hypotheses, and have been inspiration for important theories like sexual selection by female choice. Unfortunately, past major attempts to reconstruct spider phylogeny typically employing the "usual suspect" genes have been unable to produce a well-supported phylogenetic framework for the entire order. To further resolve spider evolutionary relationships we have assembled a transcriptome-based data set comprising 70 ingroup spider taxa. Using maximum likelihood and shortcut coalescence-based approaches, we analyze eight data sets, the largest of which contains 3,398 gene regions and 696,652 amino acid sites forming the largest phylogenomic analysis of spider relationships produced to date. Contrary to long held beliefs that the orb web is the crowning achievement of spider evolution, ancestral state reconstructions of web type support a phylogenetically ancient origin of the orb web and diversification analyses show that the mostly ground-dwelling, web-less RTA clade diversified faster than orb weavers. Consistent with molecular dating estimates we report herein, this may reflect a major increase in biomass of non-flying insects during the Cretaceous Tertiary Revolution 125-90 million years ago favoring diversification of spiders that feed on cursorial rather than flying prey. Our results also have major implications for our understanding of spider systematics. Phylogenomic analyses corroborate several well-accepted high level groupings: Opisthothele, Mygalomorphae, Atypoidina, Avicularioidea, Theraphosidina, Araneomorphae, Entelygynae, Araneoidea, the RTA – clade, Dionycha and the Lycosoidea. Alternatively, our results challenge the monophyly of Eresoidea, Orbiculariae, and Deinopoidea. The composition of the major Paleocribellate and Neocribellate clades, the basal divisions of Araneomorphae, appear to be falsified. Traditional Haplogynae, and even the new concept of Synspermiata, need revision after the departure of Filistatidae and Leptonetidae from the haplogyne clade. The sister pairing of filistatids with hypochilids, implies that some peculiar features of each family may in fact be synapomorphic for the pair. Leptonetids now are seen as a possible sister group to the Entelegynae, illustrating possible intermediates in the evolution of the more complex entelegyne genitalic condition, spinning organs and respiratory organs.

Keywords: Arachnida, Araneae, Molecular Systematics, Spider phylogeny, Web evolution

Authorship Statement

Order of authorship following the second author (JR) determined alphabetically through senior corresponding author (JEB).

***Corresponding Author:** 101 Rouse Life Sciences, Auburn University , AL, 36849, USA, jbond@auburn.edu

INTRODUCTION

Spiders (Order Araneae; Figure 1) are a prototypical, hyperdiverse arthropod group comprising >45,000 described species (World Spider Catalog, 2015) distributed among 3,958 genera and 114 families; by some estimates the group may include >120,000 species (Agnarsson et al., 2013). Spiders are abundant, generalist predators that play dominant roles in almost every terrestrial ecosystem. The order represents an ancient group that has continued to diversify taxonomically and ecologically since the Devonian (>380mya). They are relatively easy to collect and identify, and are one of few large arthropod orders to have a complete online taxonomic catalog with synonymies and associated literature.

In addition to their remarkable ecology, diversity, and abundance, spiders are known for the production of extraordinary biomolecules like venoms and silks as well as their utility as models for behavioral and evolutionary studies (reviewed in Agnarsson et al., 2013). Stable and complex venoms have evolved over millions of years to target predators and prey alike. Although few are dangerous to humans, spider venoms hold enormous promise as economically important insecticides and therapeutics (Saez et al., 2010; King and Hardy, 2013). Moreover, no other animal lineage can claim a more varied and elegant use of silk. A single species may have as many as eight different silk glands, producing a variety of super-strong silks deployed in almost every aspect of a spider's life: safety lines, dispersal, reproduction (sperm webs, eggsacs, pheromone trails), and prey capture (Blackledge et al. 2011). Silken prey capture webs, particularly the orb, have long been considered a key characteristic contributing to the ecological and evolutionary success of this group (reviewed in Bond and Opell, 1998). Moreover, spider silks are promising biomaterials, already benefiting humans in myriad ways - understanding the phylogenetic basis of such super-materials will facilitate efforts to reproduce their properties in biomimetic materials like artificial nerve constructs, implant coatings, and drug delivery systems (Schacht and Scheibel, 2014; Blackledge et al., 2011).

The consensus on major spider clades has changed relatively little in the last two decades since the summary of Coddington and Levi (1991) and Coddington (2005). Under the classical view, Araneae comprises two clades (see Table 1 and Figure 1 for major taxa discussed throughout; node numbers (Figure 1) referenced parenthetically hereafter), Mesothelae (Node 2) and Opisthothelae (Node 3). Mesotheles are sister to all other spiders, possessing a plesiomorphic segmented abdomen and mid-ventral (as opposed to terminal) spinnerets. Opisthothelae contains two clades: Mygalomorphae (Node 4) and Araneomorphae (Node 8). Mygalomorphae is less diverse (6% of described Araneae diversity) and retains several plesiomorphic features (e.g. two pairs of book lungs, few and biomechanically 'weak' silks (Dicko et al., 2008; Starrett et al., 2012). Within Araneomorphae, Hypochilidae (Paleocribellatae; Node 9) is sister to Neocribellatae, within which Austrochiloidea are sister to the major clades Haplogynae (Node 10) and Entelegynae (Node 11), each weakly to moderately supported by few morphological features. Haplogynes have simple genitalia under muscular control whereas entelegynes have hydraulically activated, complex genitalia, with externally sclerotized female epigyna. Entelegynes comprise multiple, major, hyperdiverse groups, including the "RTA clade" (RTA = retrolateral tibial apophysis, Node 13), its subclade Dionycha (e.g., wolf, fishing, and jumping spiders; Ramírez, 2014, Node 14), and the Orbiculariae – the cribellate and ecribellate orb weavers and relatives (see Hormiga and Griswold, 2014).

Beginning with early higher-level molecular phylogenetic studies, it gradually became clear that major "stalwart" and presumably well-supported spider groups like the Neocribellatae, Haplogynae, Palpimanoidea, Orbiculariae, Lycosoidea, and others (generally only known to arachnologists) were questionable. Subsequent studies focusing on mygalomorph (Hedin and Bond, 2006; Bond et al., 2012) and araneomorph (Blackledge et al., 2009; Dimitrov et al., 2012) relationships continued to challenge the consensus view based largely on morphological data, finding polyphyletic families and ambivalent support for major clades, which were sometimes "rescued" by adding non-molecular data; molecular signal persistently contradicted past verities. In Agnarsson et al. (2013), a meta-analysis of available molecular data failed to recover several major groups such as Araneomorphae, Haplogynae, Orbiculariae, Lycosoidea, and others (Table 1). Although these authors criticized the available molecular data as insufficient, their results actually presaged current spider phylogenomic inferences (Bond et al., 2014). Incongruence between the traditional spider classification scheme and (non-phylogenomic) molecular systematics likely has one primary cause: too few data. Non-molecular datasets to date have been restricted to a relatively small set of morphological and/or behavioral characters whereas molecular analyses addressing deep spider relationships have largely employed relatively few, rapidly evolving loci (e.g., 28S and 18S rRNA genes, Histone 3, and a number of mitochondrial DNA markers).

The first analyses of spider relationships using genome-scale data, scored for 40 taxa by Bond et al. (2014) and for 14 taxa by Fernández et al. (2014), considerably refined understanding of spider phylogeny, the former explicitly calling into question long held notions regarding the tempo and mode of spider evolution. Using transcriptome-derived data, Bond et al. (2014) recovered the monophyly of some major groups (araneomorphs and mygalomorphs) but reshuffled several araneomorph lineages (haplogynes, paleocribellates, orbicularians, araneoids (Node 12) and the RTA clade). Notably, Bond et al. (2014) rejected Orbiculariae, which included both cribellate (Deinopoidea) and ecribellate orb weavers (Araneoidea). Instead they suggested either that the orb web arose multiple times, or, more parsimoniously, that it arose once and predated the major diversification of spiders. Despite major advances in understanding of spider phylogeny, only a small percentage of spider families were sampled and monophyly of individual families could not be tested. Denser taxon sampling is needed to warrant changes in higher classification and to more definitively address major questions about spider evolution.

Herein, we apply a spider-specific core ortholog approach with significantly increased taxon and gene sampling to produce a more complete and taxon specific set of alignments for phylogenetic reconstruction and assessment of spider evolutionary pattern and process. Existing genome-derived protein predictions and transcriptome sequences from a representative group of spiders and arachnid outgroups were used to create a custom core ortholog set specific to spiders. Taxon sampling was performed to broadly sample Araneae with an emphasis on lineages whose phylogenetic placement is uncertain and included previously sequenced transcriptomes, gene models from completely sequenced genomes, and novel transcriptome sequences generated by our research team. This resulted in a data set comprising 70 spider taxa plus five additional arachnid taxa as outgroups. We test long-held notions that the orb web, in conjunction with ecribellate adhesive threads, facilitated diversification among araneoids and present the most completely sampled phylogenomic data set for spiders to date using an extensive dataset of nearly 3,400 putative genes (~700K amino acids). Further, we test the hypothesis of a non-monophyletic Orbiculariae, assess diversification rate shifts across the spider phylogeny, and provide phylogenomic hypotheses for historically difficult to place spider families. Our results clearly demonstrate that our understanding of spider phylogeny and evolution requires major reconsideration and that several long-held and contemporary morphologically-derived hypotheses are likely destined for falsification.

MATERIALS & METHODS

Sampling, Extraction, Assembly

Spider sequence data representing all major lineages were collected from previously published transcriptomic and genomic resources (N=53) and supplemented with newly sequenced transcriptomes (N=22) to form the target taxon set for the current study. Existing sequence data were acquired via the NCBI SRA database (http://www.ncbi.nlm.nih.gov/sra). Raw transcriptome sequences were downloaded, converted to fastq file format, and assembled using Trinity (Grabherr et al., 2011). Genomic data sets in the form of predicted proteins were downloaded directly from the literature (Sanggaard et al., 2014) for downstream use in our pipeline. Newly sequenced spiders were collected from a variety of sources, extracted using the TRIzol total RNA extraction method, purified with the RNeasy mini kit (Qiagen) and sequenced in-house at the Auburn University Core Genetics and Sequencing Laboratory using an Illumina Hi-Seq 2500. This produced 100bp paired end reads for each newly sequenced spider transcriptome, which were then assembled using Trinity. Proteins were predicted from each transcriptome using the program TransDecoder (Haas et al., 2013).

Core Ortholog Approach and Data Processing

We employed a core ortholog approach for putative ortholog selection and implicitly compared the effect of using a common arthropod core ortholog set and one compiled for spiders; the arthropod core ortholog set was deployed as described in Bond et al. (2014). To generate the spider core ortholog set, we used an all-versus-all BLASTP method (Altschul, Stephen F. et al., 1990) to compare the transcripts of the amblypygid *Damon variegatus*, and the spiders *Acanthoscurria geniculata*, *Dolomedes triton*, *Ero leonina*, *Hypochilus pococki*, *Leucauge venusta*, *Liphistius malayanus*, *Megahexura fulva*, *Neoscona arabesca*, *Stegodyphus mimosarum*, and *Uloborus sp*. *Acanthoscurria geniculata* and *Stegodyphus mimosarum* were represented by predicted transcripts from completely sequenced genomes while the other taxa were represented by our new Illumina transcriptomes. An e-value cut-off of 10-5 was used. Next, based on

the BLASTP results, Markov clustering was conducted using OrthoMCL 2.0 (Li et al., 2003) with an inflation parameter of 2.1.

The resulting putatively orthologous groups (OGs) were processed with a modified version of the bioinformatics pipeline employed by Kocot et al. (2011). First, sequences shorter than 100 amino acids in length were discarded. Next, each candidate OG was aligned with MAFFT (Katoh, 2005) using the automatic alignment strategy with a maxiterate value of 1,000. To screen OGs for evidence of paralogy, an "approximately maximum likelihood tree" was inferred for each remaining alignment using FastTree 2 (Price et al., 2010). Briefly, this program constructs an initial neighbor-joining tree and improves it using minimum evolution with nearest neighbor interchange (NNI) subtree rearrangement. FastTree subsequently uses minimum evolution with subtree pruning regrafting (SPR) and maximum likelihood using NNI to further improve the tree. We used the "slow" and "gamma" options; "slow" specifies a more exhaustive NNI search, while "gamma" reports the likelihood under a discrete gamma approximation with 20 categories, after the final round of optimizing branch lengths. PhyloTreePruner (Kocot et al., 2013) was then employed as a tree-based approach to screen each candidate OG for evidence of paralogy. First, nodes with support values below 0.95 were collapsed into polytomies. Next, the maximally inclusive subtree was selected where all taxa were represented by no more than one sequence or, in cases where more than one sequence was present for any taxon, all sequences from that taxon formed a monophyletic group or were part of the same polytomy. Putative paralogs (sequences falling outside of this maximally inclusive subtree) were then deleted from the input alignment. In cases where multiple sequences from the same taxon formed a clade or were part of the same polytomy, all sequences but the longest were deleted. Lastly, in order to eliminate orthology groups with poor taxon sampling, all groups sampled for fewer than 7 of the 11 taxa and all groups not sampled for *Megahexura fulva* (taxon with greatest number of identified OGs) were discarded. The remaining alignments were used to build pHMMs for HaMStR with hmmbuild and hmmcalibrate from the HMMER package (Eddy, 2011).

For orthology inference, we employed HaMStR v13.2.3 (Ebersberger et al., 2009), which infers orthology based on predefined sets of orthologs. Translated transcripts for all taxa were searched against the new set of 4,934 spider-specific profile hidden Markov models (pHMMs; available for download from the Dryad Data Repository) and an arthropod core ortholog set previously employed in Bond et al. (2014). In the spider core ortholog analysis, the genome-derived *Acanthoscurria geniculata* OGs were used as the reference protein set for reciprocal best hit scoring. *Daphnia pulex* was used as the reference species for putative ortholog detection in the arthropod core ortholog analysis. Orthologs sharing a core identification number were pooled together for all taxa and processed using a modified version of the pipeline used to generate the custom spider ortholog set. In both analyses, sequences shorter than 75 amino acids were deleted first. OGs sampled for fewer than 10 taxa were then discarded. Redundant identical sequences were removed with the perl script uniqhaplo.pl (available at http://raven.iab.alaska.edu/ ntakebay/) leaving only unique sequences for each taxon. Next, in cases where one of the first or last 20 characters of an amino acid sequence was an X (corresponding to a codon with an ambiguity, gap, or missing data), all characters between the X and that end of the sequence were deleted and treated as missing data. Each OG was then aligned with MAFFT (mafft –auto –localpair –maxiterate 1000; Katoh (2005)). Alignments were then trimmed with ALISCORE (Misof and Misof, 2009) and ALICUT (Kück, 2009) to remove ambiguously aligned regions. Next, a consensus sequence was inferred for each alignment using the EMBOSS program infoalign (Rice et al., 2000). For each sequence in each single-gene amino acid alignment, the percentage of positions of that sequence that differed from the consensus of the alignment were calculated using infoalign's "change" calculation. Any sequence with a "change" value greater than 75 was deleted. Subsequently, a custom script was used to delete any mistranslated sequence regions of 20 or fewer amino acids in length surrounded by ten or more gaps on either side. This step was important, as sequence ends were occasionally mistranslated or misaligned. Alignment columns with fewer than four non-gap characters were subsequently deleted. At this point, alignments shorter than 75 amino acids in length were discarded. Lastly, we deleted sequences that did not overlap with all other sequences in the alignment by at least 20 amino acids, starting with the shortest sequence not meeting this criterion. This step was necessary for downstream single-gene phylogenetic tree reconstruction. As a final filtering step, OGs sampled for fewer than 10 taxa were discarded.

In some cases, a taxon was represented in an OG by two or more sequences (splice variants, lineage-specific gene duplications [=inparalogs], overlooked paralogs, or exogenous contamination). In order to select the best sequence for each taxon and exclude any overlooked paralogs or exogenous contamination,

we built trees in FastTree 2 (Price et al., 2010) and used PhyloTreePruner to select the best sequence for each taxon as described above. Remaining OGs were then concatenated using FASconCAT (Kück and Meusemann, 2010). The OGs selected by our bioinformatic pipeline were further screened in seven different ways (subsets listed in Table 2). OGs were first sorted based on amount of missing data; the half with the lowest levels was pulled out as matrix 2 (1699 genes). From matrix 2, a smaller subset of OGs optimized for gene occupancy was extracted, resulting in matrix 3 (850 genes). The full supermatrix was also optimized using the programs MARE (Meyer et al., 2011) and BaCoCa (Base Composition Calculator; Kück and Struck (2014)). MARE assesses the supermatrix by partition, providing a measure of tree-likeness for each gene and optimizes the supermatrix for information content. The full supermatrix was optimized with an alpha value of 5, to produce matrix 7 (1488 genes, 58 taxa). From the MARE-reduced matrix, genes having no missing partitions for any of the remaining taxa (n=50) were extracted to form a starting matrix for the BEAST analyses (details below). Matrix assessment was also conducted using BaCoCa, which provides a number of descriptive supermatrix statistics for evaluating bias in amino acid composition and patterns in missing data. This program was used to assess for patterns of non-random clusters of sequences in the data, which could potentially mislead phylogenetic analyses. Matrix 4 represents a 50% reduction of the full supermatrix using BaCoCa derived values for phylogenetically informative sites as a guide; essentially reducing missing data from absent partitions and gaps. This matrix is similar, but not identical to matrix 2. OGs for each matrix were concatenated using FASconCAT (Kück and Meusemann, 2010).

Phylogenetics

Table 2 summarizes run parameters of the seven individual maximum likelihood analyses conducted for each of the supermatrices. We selected the optimal tree for each supermatrix using the computer program ExaML ver. 3.0.1 (Kozlov et al., 2015). Models of amino acid substitution were selected using the AUTOF command in ExaML. Bootstrap data sets and starting parsimony trees for each matrix were generated using RAxML (Stamatakis, 2014) and each individually analyzed in ExaML. We generated 225-300 replicates for each matrix which were then used to construct a majority-rule bootstrap consensus bootstrap tree; a custom python script was used to automate the process and write a bash script to execute the analyses on a high performance computing (HPC) cluster. The arthropod core OG bootstrap analysis was conducted using RAxML. All analyses were conducted on the Auburn University CASIC HPC and Atrax (Bond Lab, Auburn University).

A coalescent-based method as implemented in ASTRAL (Accurate Species Tree Algorithm) (Mirarab et al., 2014) was used to infer a species tree from a series of unrooted gene trees. The ASTRAL approach is thought to be more robust to incomplete lineage sorting, or deep coalescence, than concatenation methods or other shortcut coalescent-based approaches (Mirarab et al., 2014). We first constructed individual gene trees for all partitions contained within matrix A. Gene trees were generated using ML based on 100 RAxML random addition sequence replicates followed by 100 bootstrap replicates (Table 2). Subsequent species tree estimation was inferred using ASTRAL v4.7.6, from all individual unrooted gene trees (and bootstrap replicates), under the multi-species coalescent model.

A chronogram was inferred in a Bayesian framework under an uncorrelated lognormal relaxed clock model (Drummond et al., 2006; Drummond and Rambaut, 2007) using Beast v1.8.1 (Drummond et al., 2012). For this analysis we used 43 partitions of a matrix which included complete partitions for all taxa derived from the MARE-optimized matrix 7. The model of protein evolution for each partition was determined using the perl script ProteinModelSelection.pl in RAxML. Beast analyses were run separately for each partition using eight calibration points based on fossil data. The most recent common ancestor (MRCA) of Mesothelae + all remaining spiders was given a lognormal prior of (mean in real space) 349 Ma (SD=0.1) based on the Mesothelae fossil *Palaeothele montceauensis* (Selden, 1996). The MRCA of extant araneomorphs was given a lognormal prior of (mean in real space) 267 Ma (SD=0.2) based on the fossil *Triassaraneus andersonorum* (Selden et al., 1999). The MRCA of extant mygalomorphs was given a lognormal prior of (mean in real space) 278 Ma (SD=0.1) based on the fossil *Rosamygale grauvogeli* (Selden and Gall, 1992). The MRCA of Haplogynae + Hypochilidae was given a lognormal prior of (mean in real space) 278 Ma (SD=0.1) based on the fossil Eoplectreurys gertschi (Selden and Penney, 2010). The MRCA of Deinopoidea (cribellate orb-weavers) was given a lognormal prior of (mean in real space) 195 Ma (SD=0.3) based on the fossil *Mongolarachne jurassica* (Selden et al., 2013). The MRCA of ecribellate orb-weavers was given a lognormal prior of (mean in real space) 168

Ma (SD=0.4) based on the fossil *Mesozygiella dunlopi* (Penney and Ortuño, 2006). The MRCA of Nemesiidae, excluding *Damarchus*, was given a lognormal prior of (mean in real space) 168 Ma (SD=0.4) based on the nemesiid fossil *Cretamygale chasei* (Selden, 2002). Finally, the MRCA of Antrodiaetidae was given a lognormal prior of (mean in real space) 168 Ma (SD=0.4) based on the fossil *Cretacattyma raveni* (Eskov and Zonstein, 1990). Two or more independent Markov Chain Monte Carlo (MCMC) searches were performed until a parameter effective sample size (ESS) >200 was achieved. ESS values were examined in Tracer v1.5. Independent runs for each partition were assembled with LogCombiner v1.7.5 and 10% percent of generations were discarded as burn-in. Tree files for each partition where then uniformly sampled to obtain 10,000 trees. A total of 430,000 trees (10,000 trees from each partition) were assembled with LogCombiner v1.7.5 and a consensus tree was produced using TreeAnnotator v1.8.1. A chronogram containing all taxa was generated using a penalized likelihood method in r8s v1.8 (Sanderson, 2002). The 95% highest posterior density dates obtained for the Beast analysis were incorporated as constraints for node ages of the eight fossil calibrated nodes. The analysis was performed using the TN algorithm, cross validation of branch-length variation and rate variation modeled as a gamma distribution with an alpha shape parameter.

To detect diversification rate shifts, we performed a Bayesian analysis of diversification in BAMM (Bayesian Analysis of Macroevolutionary Mixtures; Rabosky et al., 2014). For this analysis we used the chronogram obtained by the r8s analysis in order to maximize taxon sampling. To account for non-random missing speciation events, we quantified the percentage of taxa sampled per family and incorporated these into the analysis. We also accounted for missing families sampled at various taxonomic levels. The MCMC chain was run for 100,000,000 generations, with sampling every 10,000 generations. Convergence diagnostics were examined using coda (Plummer et al., 2006) in R. Ten percent of the runs were discarded as burn-in. The 95% credible set of shift configurations was plotted in the R package BAMMtools (Rabosky et al., 2014).

Character state reconstructions of web type following Blackledge et al. (2009) were performed using a maximum likelihood approach. The ML approach was implemented using the rayDISC command in the package corHMM (Beaulieu et al., 2013) in R (Ihaka and Gentleman, 1996). This method allows for multistate characters, unresolved nodes, and ambiguities (polymorphic taxa or missing data). Three models of character evolution were evaluated under the ML method: equal rates (ER), symmetrical (SYM) and all rates different (ARD). A likelihood-ratio test was performed to select among these varying models of character evolution.

RESULTS

Summary of Genomic Data

Twenty-one novel spider transcriptomes were sequenced, with an average of 72,487 assembled contigs (contiguous sequences) ranging from 6,816 (*Diguetia sp.*) to 191,839 (*Segestria sp.*); specimen data and transcriptome statistics for each sample are summarized in Supplemental Tables S1 and S2 respectively. Median contig length for the novel transcriptomes was 612 bp. The complete taxon set, including spider and outgroup transcriptomes from the SRA database, had an average contig number of 53,740 and a range of 5,158 (*Paratropis sp.*) to 202,311 (*Amaurobius ferox*) with a median contig length of 655. The newly constructed spider-specific core ortholog group (OG) set contained 4,934 OGs, more than three times the number of arthropod core orthologs used in prior spider analyses (Bond et al., 2014) and represents a significant step forward in generating a pool of reasonably well-vetted orthologs for spider phylogenomic analyses. The arthropod and spider core orthology sets had 749 groups in common; 4,185 OGs in the spider core were novel. Of the spider-core groups, 4,249 (86%) were present in the sequenced genome of our HaMSTR reference taxon of choice *Acanthoscurria geniculata* (Sanggaard et al., 2014) and were retained for use in downstream ortholog detection. The number of TransDecoder predicted proteins and ortholog detection success for each taxon is summarized in Table S2. Annotations for the arthropod set can be found in Bond et al. (2014); Supplemental Table S3 summarizes gene annotations for the spider core ortholog set generated for this study. Our new HaMStR spider core ortholog set and *Acanthoscurria geniculata* BLAST database file can be downloaded from the Dryad Data Repository at doi.xxxx.xxxxxx.

Phylogenetic Analyses

Seven super matrices were generated for downstream non time-calibrated analyses (Figure 2), one drawn from the arthropod core set and six using the spider core set. Data set sizes, summarized in Table 2, ranged

from a maximum of 3,398 OGs with a higher percentage of missing cells (38.5%), 850 OGs with 19.6% missing, to 549 OGs (arthropod core set) with 33% missing data. Two matrices were generated using automated filtering approaches implemented by BaCoCa (Kück and Struck, 2014) and MARE (Meyer et al., 2011). In BaCoCa we sorted partitions using number of informative sites, capturing the top half (1700 OGs) of the matrix containing the most informative sites. RCFV values generated by BaCoCa were <0.05 for all taxa in all partitions for each of the matrices, indicating homogeneity in base composition. Additionally, there was no perceptible taxonomic bias observed in shared missing data (Supplemental Figures S1-S6). The MARE optimized matrix comprised 58 taxa and 1,488 genes with 19.6% missing data. For graphical representations of gene occupancy for each matrix, see Supplemental Figures S7-S12. Blast2GO (Conesa et al., 2005) gene ontology distributions of molecular function for OGs recovered from both the spider and arthropod ortholog sets (Suplemental Figures S13 and S14) can be found in the supplemental materials.

Our phylogenetic analyses (see Table 2 and Discussion), the results of which are summarized in Figure 2, consistently recover many well-supported monophyletic groups: Araneae, Mygalomorphae, Araneomorphae, Haplogynae (excluding Filistatidae and Leptonetidae), Entelegynae, the RTA clade, Dionycha, and Lycosoidea. Within Mygalomorphae, Atypoidina and Avicularioidea are monophyletic; Nemesiidae is polyphyletic. Filistatidae (*Kukulcania*) is removed from other haplogynes and emerges as the sister group to *Hypochilus*. Interestingly, Leptonetidae emerges as the sister group to Entelegynae. Eresidae, controversially, is sister to Araneoidea. Deinopoidea is polyphyletic. Oecobiidae is sister to Uloboridae, which are together sister to Deinopidae plus the RTA clade. Homalonychidae (previously unplaced by phylogenomics) and by implication the entire Zodarioidea (Miller et al., 2010), is sister to Dionycha plus Lycosoidea. Hahniidae, represented by the cryphoecine *Calymmaria*, is sister to Dictynidae. Thomisidae belongs in Lycosoidea as proposed by Homann (1971) and Polotow et al. (2015).

Coalescent-based species-tree analysis in ASTRAL employed unrooted gene trees based on the 3,398 gene matrix as input and inferred a well-supported tree (most nodes >95% bs; Figure 2). With few exceptions the topology recovered using this approach was congruent with the likelihood-based supermatrix analysis. Conflicting nodes, some corresponding to key araneomorph lineages, which were moderately to weakly supported in concatenated analyses, are summarized in Figure 2.

A chronogram based on 43 partitions with no missing data (matrix 7, see Table 2) is shown in Figure 4. Divergence time estimates are summarized in Table 3: Mesothelae - Opisthothelae at 340 Ma (287-398 95% CI); Mygalomorphae - Araneomorphae at 308 Ma (258-365 95% CI); Haplogynae + Hypochilidae - Entelegynae at 276 Ma (223-330 95% CI); RTA + Deinopoidea - *Stegodyphus* + Araneoidea at 214 Ma (154-280 95% CI); RTA - Dionycha at 138.8 Ma (Figure 4).

Diversification rate shift analysis estimated three instances of significant diversification shifts within spiders (95% credibility). The highest rate shift is within the RTA + Dionycha + Lycosoidea (Figure 5) followed by Avicularioidea and within Araneoidea (f = 0.23; 0.21; Figure 5).

Maximum likelihood ancestral state reconstruction of web type (Figure 6) shows that the spider common ancestor likely foraged from a subterranean burrow, sometimes sealed by a trapdoor. The ancestral condition for araneomorphs may have been a stereotypical aerial sheet. Entelegynae ancestors probably spun orbs, which were subsequently lost at least three times. RTA taxa largely abandoned webs to become hunting spiders. Precise location of these character state shifts depends upon sufficient sampling; denser sampling reduces the number of unobserved evolutionary events. While this analysis contains only 47 of 114 spider families, the sequence and overall mapping to the spider backbone phylogeny is strongly supported.

DISCUSSION

Our phylogenomic analyses represent the largest assessment of spider phylogeny to date using genomic data, both in terms of taxa and number of orthologs sampled. Our results are largely congruent with our earlier work (Bond et al., 2014): we recover all of the major backbone lineages (Mygalomorphae, Araneomorphae, RTA, etc.), but reiterate that our understanding of spider evolutionary pattern and process needs thorough reconsideration. This expanded study reinforces the ancient origin of the orb web hypothesis and shows that rates of spider species diversification appear to be associated with web change or loss – or with modification of the male palp rather than the origin of the orb web. It shows that the Haplogynae are polyphyletic with Filistatidae as sister to Hypochilidae and Leptonetidae as sister to

Entelegynae. It also suggests a position for two enigmatic families – Hahniidae and Homalonychidae – and provides an alternate view of RTA relationships and the contents of Dionycha clade.

Data Characteristics and Development of Spider Core Orthologs

Transcriptome analyses are unquestionably data rich. Thousands of assembled sequences emerge from even modest RNA-seq experiments, providing, among other things, a basis for identifying phylogenetically informative orthologs. This bounty comes with a few caveats. Isoforms, paralogous sequences, and assembly artifacts (chimeric contigs) can mislead inference of single-copy orthologous genes. The data represent one snapshot – a specific organism, point in time, and combination of tissues – that can lead to gaps in downstream supermatrices due to stochastic sampling issues. Large amounts of missing data, due to missing loci and indels introduced during alignment, can arise post-assembly in the ortholog detection and filtering stages of phylogenomic analyses (compare Bond et al. (2014) to Fernández et al. (2014)). Lemmon et al. (2009) and a number of other authors (Roure et al., 2013; Dell'Ampio et al., 2014; Xia, 2014) have discussed the potential negative effects of such missing data in large phylogenomic (transcriptome-based) datasets. Phylogeny inference may also be misled by recombination (Springer and Gatesy, 2016), especially in genes with exons separated by long introns; spider genomes published to date are intron-rich (Sanggaard et al., 2014). Recent studies argue that the phylogenetic signal from transcriptomes can conflict with alternative reduced representation approaches like targeted sequence capture (Jarvis et al., 2014; Brandley et al., 2015; Prum et al., 2015). From vast amounts of bird genome protein-coding data, Jarvis et al. (2014) concluded that these loci were not only insufficient (low support values), but also misleading due to convergence and high levels of incomplete lineage sorting during rapid radiations.

Simulation studies now predict that 10's-100's of loci will resolve most phylogenies, albeit sensitive to factors such as population size or speciation tempos (Knowles and Kubatko, 2011; Leache and Rannala, 2011; Liu and Yu, 2011). To mitigate the impacts of paralogy, incomplete lineage sorting, and missing data, we developed *a priori* a set of spider core orthologs that comprise a database consisting of over 4,500 genes that are expected to be recovered from most whole spider RNA extractions and are likely orthologous. We summarize the annotations for each of the genes in the HaMStR pHMM file in Supplemental table S3.

Our approach enhances repeatability, downstream assessment, scalability (taxon addition), and data quality. Studies that employ pure clustering approaches like OMA stand-alone (Altenhoff et al., 2013) may produce more data (i.e., more "genes") on the front end; however, they are not as scalable, lack comparability, and, if not carefully curated or filtered, will retain some low quality loci. Although adding more genes is one strategy (e.g., (Lopardo and Hormiga, 2015), it is increasingly clear that taxon sampling and data quality are more important than quantity (Lemmon and Lemmon, 2013; Bond et al., 2014).

A Modified View of Spider Evolution and Key Innovations

Once considered the "crowning achievement of aerial spiders" (Gertsch, 1979), the orb web and consequent adaptive radiation of araneoid spiders (ecribellate orb weavers and their relatives) captured the imagination of spider researchers for over a century. The evolution of adhesive threads and the vertical orientation of the orb web, positioned to intercept and retain flying insects, has been long considered a "key innovation" that allowed spiders to inhabit a new adaptive zone (Bond and Opell, 1998). It is important to note that several prior authors speculated about orb web adaptive value, such as Levi (1980), Opell (1979, 1983), and Coddington (1986), although Bond and Opell (1998) quantified the pattern in a formal phylogenetic framework. Over 25% of all spider species are araneoids. Given orb weaver monophyly on quantitative phylogenies, rigorous empirical studies tended to confirm the orb as a prime cause of spider diversification. Nevertheless, a lack of correlation of the orb web and species richness has been apparent for some time. Griswold et al. (1998) noted that over 50% of Araneoidea no longer build recognizable orb webs and suggested that "the orb web has been an evolutionary base camp rather than a summit."

Bond et al. (2014) tested two alternative evolutionary scenarios for orb web evolution, reflecting different analytical results; parsimony implied multiple independent origins, and maximum likelihood implied one origin and subsequent multiple losses. The current study (Figure 6) favors the latter: the orb evolves at the base of the araneoid + deinopoid + RTA clade, but is lost at least three times independently. Large amounts of morphological and behavioral data (albeit often correlated with features essential to the orb) still support the single origin hypothesis (Coddington, 1986, 1991; Scharff and Coddington, 1997; Griswold et al., 1998; Agnarsson et al., 2013). Our results suggest both that the orb web originated earlier

than previously supposed, and that heretofore-unsuspected clades of spiders descend from orb weavers. In a sense, this ancient origin hypothesis reconciles the implications of genomic data with the classical evidence for multiple, homologous, complex, co-adapted character systems.

Recent discoveries of large, cribellate orb web-weaving taxa from the late Triassic agree with our molecular dates. Diverse Mesozoic deinopoids (Selden et al., 2015) are consistent with the "orb web node" at 213 mya (Figure 4, Table 3). Under this view, modern uloborids and deinopids are distinct remnants of this diverse group. Selden et al. (2015) previously noted that if other extant taxa "emerged from the deinopoid stem or crown group it would render the whole-group Deinopoidea paraphyletic"; we discuss this scenario in detail below.

Contrary to the contemporary paradigm that the evolution of the orb web and adhesive sticky threads elevated rates of diversification among the araneoid spiders, our BAMM analysis (Figure 5) indicates that the highest rates of diversification likely occurred among the RTA spiders followed by mygalomorphs and then araneoids as a distant third, the latter driven–in part–by the secondarily non-orb weaving theridiids and linyphiids. These results imply that other foraging strategies (e.g. cursorial hunting and irregular sheets) were a more "successful" strategy than the orb. Indeed, the point estimate for the RTA node during the early Cretaceous (138.8 mya; Figure 4, Table 3) precedes the subsequent diversification of the RTA clade at 100-125 mya.

This date coincides with the Cretaceous Tertiary Revolution (KTR). Angiosperms radiated extensively at 125-90 Ma (Crane, 1987; Wang et al., 2013), as did various plant-dependent insect lineages, including weevils (McKenna et al., 2009), lepidopterans (Wahlberg et al., 2013), ants (Moreau, 2006), and holometabolous insects in general (Misof et al., 2014), although some insect lineages do not show a pulse (e.g., darkling beetles; Kergoat et al., 2014). Spiders, as important insect predators, may also have diversified rapidly along with their prey (e.g., Penney et al., 2003; Penalver, 2006; Selden and Penney, 2010). The fossil and phylogenomic data presented here show that most spider lineages predate the KTR (Selden and Penney, 2010; Bond et al., 2014). Among these, the RTA clade especially, but also mygalomorphs and araneoids, diversified in response to the KTR insect pulse. That aerial web spinners specialized on rapidly radiating clades of flying insects is hardly surprising. Similarly, if forest litter habitats became more complex and spurred insect diversification (Moreau, 2006), ground-dwelling spiders may also have diversified at unusual rates. Perhaps the most dramatic change in insect abundances occurred with the origin and early diversification of the social insects that today dominate animal biomass on the planet: the ants and the termites (Hölldobler and Wilson, 1990). Both groups date back to 150-125 my and diversified during the KTR (LaPolla et al., 2013; Ward, 2014; Legendre et al., 2015). A major increase in biomass of non-flying insects may have favoured spiders that feed on cursorial prey and thus could help explain the concurrent increase in diversification in the RTA clade, mygalomorphs, and non-orb weaving araneoids such as cobweb weavers (Dziki et al., 2015).

Taken together, this new evidence on character evolution, divergence estimates, and rates of diversification indicates that previous conclusions regarding the timing and rate of spider evolution were imprecise, if not faulty. Our data support an ancient orb web hypothesis that is further bolstered by a wealth of fossil data showing that a cribellate deinopoid stem group likely diversified during the early Mesozoic. Molecular divergence clock estimates are consistent with the placement of the orb web further down the tree as well as suggesting that some of the greatest rates of species diversification coincided with the KTR. The latter suggests that spiders took advantage of increased abundance of cursorial prey.

These findings likely diminish the hypothesis proposed by Bond and Opell (1998) that the vertically oriented orb web represented a key innovation, particularly in light of the fact that over half of araneoid species do not build an orb web (e.g. Theridiidae and Linyphiidae; noted by Griswold et al., 1998; Fernández et al., 2014). We already knew that major orb web-weaving groups are very successful in spite of abandoning the orb (Blackledge et al., 2009).

Spider Systematics

Although our results show that many classical ideas in spider systematics require revision (e.g. mygalomorph families, Haplogynae, paleocribellates, higher araneoids, and RTA + dionychan lineages), they also robustly support many classical taxonomic concepts.

Mygalomorphae relationships.

Since Raven (1985), Mygalomorphae (Table 1, Node 4) has continuously represented a challenge to spider systematics. As discussed by Hedin and Bond (2006) and Bond et al. (2012), nearly half the

families are probably non-monophyletic. While our sampling here and previously (Bond et al., 2014) is far greater than any other published phylogenomic study (e.g., Fernández et al. (2014) included just one theraphosid), taxon sampling remains insufficient to address major issues aside from deeper level phylogenetic problems. However, the data (Figure 2) support Euctenizidae as a monophyletic family, but not Nemesiidae. As indicated in Bond et al. (2014), the once controversial Atypoidina (Node 5) consistently has strong statistical support in all analyses. Alternatively, the placement of paratropidids, ctenizids, and idiopids remains questionable and warrants further sampling.

Haplogynae relationships.

The traditional view of spider phylogeny (Coddington, 2005) places Paleocribellatae and Austrochiloidea (Table 1) as sister groups to all the remaining Araneomorphae taxa – Haplogynae and Entelegynae. Our current tree (Figure 2) is congruent with Bond et al. (2014) in placing Paleocribellatae (Table 1, *Hypochilus*; Figure 1, Node 9) as sister to Neocribellatae. Filistatidae (*Kukulcania*), formerly placed as sister to all other haplogynes, pairs with Hypochilus in a novel arrangement. This arrangement suggests that characters formerly considered "primitive" to araneomorphs, for example, mobile leg three cribellate silk carding and an M-shaped midgut, might instead be synapomorphies for the new hypochilid-filistatid clade. Remaining haplogyne relationships are somewhat congruent with previously published analyses (Ramírez, 2000). However, one of the more intriguing results is the placement of the morphologically intermediate "haplogyne" (Table 1) *Calileptoneta*, (Leptonetidae) as sister to Entelegynae, suggesting that leptonetids may represent intermediate genitalic forms between haplogyne and the relatively more complex entelegyne condition (Ledford and Griswold, 2010). As outlined by Ledford and Griswold (2010), a number of previous analyses (Platnick et al., 1991; Ramírez, 2000; Griswold et al., 2005) discussed the "rampant" homoplasy required to place leptonetids (sister to Telemidae) among haplogynes and suggest two possible scenarios – leptonetids are proto-entelegynes, or they are the sister group to the remaining Haplogynae. Our phylogenomic analyses support the former hypothesis favored by Ledford and Griswold (2010), and puts the discovery of the cribellate *Archoleptoneta* into better phylogenetic context. These results provide a novel and robust phylogenetic framework for understanding the evolution of entelegyne genitalia.

Araneoidea relationships.

Our reconstruction of araneoid relationships departs dramatically from the traditional classification scheme and a number of recently published molecular systematic studies (e.g., Blackledge et al., 2009; Dimitrov et al., 2012). Theridiidae (cobweb spiders) is sister to the remaining araneoids as opposed to occupying a more derived position within that clade. Comparisons to Dimitrov et al. (2012) should be viewed with caution: that analysis had a large suite of taxa not included here, and many results of that analysis had only weak support. Nevertheless, our phylogenomic data agree in supporting the close relationship between Mysmenidae, Mimetidae, and Tetragnathidae. Unlike that study, we recover nesticids sister to linyphioids (Pimoidae plus Linyphiidae) rather than theridiids: Theridioid (Theridiidae and Nesticidae) diphyly is a surprising result, which has already been shown with standard markers by Agnarsson et al. (2013). Theridioids have strikingly similar spinning organs and tarsus IV comb for throwing silk, but are otherwise genitalically distinct. We retain the more inclusive linyphioids as close relatives of Araneidae + Nephilidae. Clearly relationships among the derived araneoids require more intensive sampling, especially of missing families (Theridiosomatidae, Malkaridae, Anapidae, etc.) to adequately resolve their phylogeny.

Deinopoidea relationships.

The addition of nearly 30 terminals to the Bond et al. (2014) dataset corroborates the non-monophyly of the classically defined Orbiculariae, although the orb and its behavioral, morphological, and structural constituents may be homologous. Deinopoidea, with these data, is polyphyletic. Instead, a new clade, Uloboridae + Oecobiidae, is sister to Araneoidea + Deinopidae + the RTA clade. Bootstrap support was consistently low for this node in all analyses except matrix 6 (Figure 2), which omits the eresid exemplar *Stegodyphus*. The placement of the two eresoid taxa (Table 1), *Stegodyphus* and *Oecobius* continues to present difficulties here as in previous published phylogenomic studies. Fernández et al. (2014) found alternative placements for *Oecobius* (their only eresoid) whereas Bond et al. (2014) typically recovered *Stegodyphus* as the sister group to all entelegynes (recovered here as the sister group to araneoids) and *Oecobius* as a member of a clade comprising uloborid and deinopid exemplars, but with notably lower support. Disparities between the two analyses may be attributed to differences in taxon

sampling, which, as noted above, was far greater in Bond et al. (2014). On the other hand, increased taxon sampling across the tree diminished node support in some places. However, it is worth noting that support was very strong in the ASTRAL species tree analysis, suggesting that while there may be some conflict among individual data partitions there is an overwhelming amount of signal in the data for a Deinopoidea + RTA relationship. This trend was noted by Bond et al. (2014) who found that only 2.4% of all bootstrap replicates recovered a monophyletic Orbiculariae. Based on these data and the putative rapid diversification that occurred once the orb web was abandoned, it is clear that resolving relationships at this point in spider evolutionary history remains a challenge. Finally, Bond et al. (2014) and Agnarsson et al. (2013) recovered an unexpected relationship between eresoid taxa and deinopoids that consistently rendered the Deinopoidea paraphyletic or polyphyletic if *Oecobius* was included in the analysis. Our results, here including an additional uloborid exemplar, still confirm Deinopoidea polyphyly. Perhaps careful examination of *Oecobius* web morphology and spinning behavior will provide independent corroboration of this molecular signal.

RTA and Dionycha relationships.

Although all of our analyses recover a monophyletic RTA clade, relationships among its members reflect some departure from the traditional view of RTA phylogeny but are largely consistent with a more recent morphology-based study. First, we recover a clade that comprises a mix of agelenoids (Agelenidae, Desidae, and Amphinectidae) as a sister group to Dictynidae + Hahniidae and Amaurobiidae. The taxonomic composition of Dictynidae, Hahniidae and Amaurobiidae, as well as their phylogenetic placement, remains problematic and in a state of flux (Coddington, 2005; Spagna et al., 2010; Miller et al., 2010). The typical hahniine hahniids have been difficult to place due to their long branches (Spagna and Gillespie, 2008; Miller et al., 2010). *Calymmaria*, has been moved into "Cybaeidae s.l." by Spagna et al. (2010), suggesting that the relationships among hahniids, cybaeids, and dictynids need further scrutiny.

Second, these data indicate a novel placement for Amaurobiidae. Amaurobiids have also been hard to place, though this is in part because Amaurobiidae are a moving target. The term "Amaurobiids" needs to be clarified, as most of nine subfamilies discussed in Lehtinen (1967) are now placed elsewhere. We use *Callobius*, from the type subfamily of the family. Our amaurobiid placement, basal to an agelenoid and dictynoid grouping corroborates previous findings (Miller et al., 2010; Spagna et al., 2010). Dictynids on the other hand were considered one of the unresolved sister groups to amaurobioids, zodarioids, and dionychans (Spagna et al., 2010). Here the placement of our dictynid exemplar *Cicurina* is more precise: sister group to the hahniid *Calymmaria*.

Third, we recover Homalonychidae (representing Zodarioidea) as the sister group to dionychans and lycosoids, once again, mirroring the results of Agnarsson et al. (2013). Previously Zodarioidea was placed closer to the base of the RTA clade (Miller et al., 2010). Dionychans here include salticids, anyphaenids, corinnids, and gnaphosids whereas crab spiders (Thomisidae) nest with the lycosoids containing a paraphyletic Pisauridae. Placement of Thomisidae within Lycosoidea goes back at least to Homann (1971) and was formally established by the total evidence analysis of Polotow et al. (2015). Although Ramírez (2014) placed Thomisidae outside of Lycosoidea, in one of his slightly suboptimal results thomisids were included in Lycosoidea. The relationships we recover among dionychan and lycosoid taxa are largely congruent with those inferred by Ramírez (2014) in a massive morphological study of Dionycha and RTA exemplars. Given the general incongruence among previous morphological and molecular spider systematic studies, it will be interesting to see how Ramírez (2014) phylogeny and familial-level reevaluations compare as phylogenomic studies expand. Raven (1985) was a landmark study for mygalomorphs; perhaps Ramírez (2014) may serve in the same capacity for one of the most diverse branches on the spider tree of life.

CONCLUSIONS

Following Coddington and Levi (1991), higher-level spider systematics underwent a series of challenges from quantitative studies of morphology, producing provocative but weakly-supported hypotheses. (Griswold et al., 1998, 2005). Total evidence studies, for example, Wood et al. (2012a,b) for Palpimanoidea, Polotow et al. (2015) for Lycosoidea, appear to have settled some local arrangements, but much of the backbone of the spider tree of life remains an open question. Phylogenomics has already brought data-rich, convincing solutions to long standing controversies, for example, phylogeny of the orb web (Bond et al., 2014; Fernández et al., 2014). Phylogenomics portends a new and exciting period for spider evolutionary

biology. Recent advances in digital imaging, proteomics, silk biology and major fossil discoveries mean that our understanding of spider evolution will likely accelerate by leaps and bounds in the coming years. The tempo and mode of spider evolution is likely different than previously thought. At this point it seems reasonably clear that the orb web evolved earlier phylogenetically than previously thought, only to be subsequently lost at least three times independently during the Cretaceous. While the orb web has certainly been successful, a likely dramatic increase in the abundances of cursorial insects during the KTR, including the emergence of most social insects, also impacted the success of other foraging strategies, including webless hunting. Our results and that of others like Ramírez (2014) show that spider systematics still remains a work in progress.

ADDITIONAL INFORMATION AND DECLARATIONS

Acknowledgments

This is contribution 7XX of the Auburn University Museum of Natural History.

Competing Interests

The authors declare there are no competing interests

Author Contributions

Jason E. Bond, Nicole L. Garrison, and Juanita Rodriguez designed the study, analyzed the data and wrote the paper. Kevin Kocot and Chris A. Hamilton played integral roles in the development of the spider core ortholog data set and gene tree analyses respectively. All remaining authors contributed specimens and/or transcriptomes, assisted in writing manuscript sections, and reviewed drafts of the paper.

Accession Numbers

Illumina transcriptome sequence data are available from NCBI database archive under accession numbers SAMNXXXXX-SAMNXXXXX. Phylogenomics data matrices were deposited on XX November 2015 in the Dryad Digital Repository at http://dx.doi.org/xx.xxxx/drayd.xxxxx.

Supplemental Information

Supplemental information, figures and tables, can be found online at http://dx.doi.org/xx.xxxx/peerj.

REFERENCES

Agnarsson, I., Coddington, J. A., and Kuntner, M. (2013). Systematics - progress in the study of spider diversity and evolution. In Penney, D., editor, *Spider Research in the 21st Century*, pages 58–111. Manchester.

Altenhoff, A. M., Gil, M., Gonnet, G. H., and Dessimoz, C. (2013). Inferring Hierarchical Orthologous Groups from Orthologous Gene Pairs. *PLoS ONE*, 8(1):e53786.

Altschul, Stephen F., Gish, Warren, Miller, Webb, Myers, Eugene W., and Lipman, David J. (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215:403–410.

Beaulieu, J. M., O'Meara, B. C., and Donoghue, M. J. (2013). Identifying Hidden Rate Changes in the Evolution of a Binary Morphological Character: The Evolution of Plant Habit in Campanulid Angiosperms. *Systematic Biology*, 62(5):725–737.

Blackledge, T. A., Kuntner, M., and Agnarsson, I. (2011). The form and function of spider orb webs: evolution from silk to ecosystems. *Advances in Insect Physiology*, 41:175.

Blackledge, T. A., Scharff, N., Coddington, J. A., Szüts, T., Wenzel, J. W., Hayashi, C. Y., and Agnarsson, I. (2009). Reconstructing web evolution and spider diversification in the molecular era. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13):5229–5234.

Bond, J. E., Garrison, N. L., Hamilton, C. A., Godwin, R. L., Hedin, M., and Agnarsson, I. (2014). Phylogenomics Resolves a Spider Backbone Phylogeny and Rejects a Prevailing Paradigm for Orb Web Evolution. *Current Biology*, 24(15):1765–1771.

Bond, J. E., Hendrixson, B. E., Hamilton, C. A., and 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.

Bond, J. E. and Opell, B. D. (1998). Testing Adaptive Radiation and Key Innovation Hypotheses in Spiders. *Evolution*, 52(2):403.

606 Brandley, M. C., Bragg, J. G., Singhal, S., Chapple, D. G., Jennings, C. K., Lemmon, A. R., Lemmon,
607 E. M., Thompson, M. B., and Moritz, C. (2015). Evaluating the performance of anchored hybrid
608 enrichment at the tips of the tree of life: a phylogenetic analysis of Australian Eugongylus group scincid
609 lizards. *BMC Evolutionary Biology*, 15(62).
610 Coddington, J. (1986). The monophyletic origin of the orb web. In Shear, W., editor, *Spiders: Webs,*
611 *Behavior, and Evolution*, pages 319–363. Stanford University Press, Stanford, California.
612 Coddington, J. A. (1991). Cladistics and spider classification: araneomorph phylogeny and the monophyly
613 of orbweavers (Araneae: Araneomorphae; Orbiculariae). *Acta Zoologica Fennica*, 190:75–87.
614 Coddington, J. A. (2005). Phylogeny and classification of spiders. In Ubick, P., Paquin, P., Cushing, P.,
615 and Roth, V., editors, *Spiders of North America: an identification manual*, pages 18–24. American
616 Arachnological Society.
617 Coddington, J. A. and Levi, H. W. (1991). Systematics and Evolution of Spiders (Araneae). *Annual*
618 *Review of Ecology and Systematics*, 22:565–592.
619 Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2go: a
620 universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*,
621 21(18):3674–3676.
622 Crane, P. (1987). The origin of angiosperms and their biological consequences. In Friis, E., Chaloner, W.,
623 and Crane, P., editors, *Vegetational consequences of the angiosperm diversification*, pages 105–144.
624 Cambridge University Press, Cambridge.
625 Dell'Ampio, E., Meusemann, K., Szucsich, N. U., Peters, R. S., Meyer, B., Borner, J., Petersen, M.,
626 Aberer, A. J., Stamatakis, A., Walzl, M. G., Minh, B. Q., von Haeseler, A., Ebersberger, I., Pass, G.,
627 and Misof, B. (2014). Decisive Data Sets in Phylogenomics: Lessons from Studies on the Phylogenetic
628 Relationships of Primarily Wingless Insects. *Molecular Biology and Evolution*, 31(1):239–249.
629 Dicko, C., Porter, D., Bond, J., Kenney, J. M., and Vollrath, F. (2008). Structural Disorder in Silk Proteins
630 Reveals the Emergence of Elastomericity. *Biomacromolecules*, 9(1):216–221.
631 Dimitrov, D., Lopardo, L., Giribet, G., Arnedo, M. A., Alvarez-Padilla, F., and Hormiga, G. (2012).
632 Tangled in a sparse spider web: single origin of orb weavers and their spinning work unravelled by
633 denser taxonomic sampling. *Proceedings of the Royal Society B: Biological Sciences*, 279(1732):1341–
634 1350.
635 Drummond, A. J., Ho, S. Y. W., Phillips, M. J., and Rambaut, A. (2006). Relaxed Phylogenetics and
636 Dating with Confidence. *PLoS Biology*, 4(5):e88.
637 Drummond, A. J. and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees.
638 *BMC Evolutionary Biology*, 7(1):214.
639 Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with
640 BEAUti and the BEAST 1.7. *Molecular biology and evolution*, 29(8):1969–1973.
641 Dziki, A., Binford, G., Coddington, J. A., and Agnarsson, I. (2015). Spintharus flavidus in the caribbean–a
642 30 million year biogeographical history and radiation of a 'widespread species'. Technical report, PeerJ
643 PrePrints.
644 Ebersberger, I., Strauss, S., and von Haeseler, A. (2009). HaMStR: Profile hidden markov model based
645 search for orthologs in ESTs. *BMC Evolutionary Biology*, 9(1):157.
646 Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Computational Biology*, 7(10):e1002195.
647 Eskov, K. Y. and Zonstein, S. (1990). First Mesozoic mygalomorph spiders from the Lower Cretaceous
648 of Siberia and Mongolia, with notes on the system and evolution of the infraorder Mygalomorphae
649 (Chelicerata: Araneae). *Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen*, 178:325–368.
650 Fernández, R., Hormiga, G., and Giribet, G. (2014). Phylogenomic Analysis of Spiders Reveals Non-
651 monophyly of Orb Weavers. *Current Biology*, 24(15):1772–1777.
652 Gertsch, W. J. (1979). *American spiders.* Number Ed. 2. Van Nostrand Reinhold Co.
653 Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L.,
654 Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma,
655 F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length
656 transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*,
657 29(7):644–652.
658 Griswold, C. E., Coddington, J. A., Hormiga, G., and Scharff, N. (1998). Phylogeny of the orb-web
659 building spiders (Araneae, Orbiculariae: Deinopoidea, Araneoidea). *Zoological Journal of the Linnean*
660 *Society*, 123(1):1–99.

Griswold, C. E., Ramírez, M., Coddington, J., and Platnick, N. (2005). Atlas of Phylogenetic Data for Entelegyne Spiders (Araneae: Araneomorphae: Entelegynae), with Comments on Their Phylogeny. *Procceedings of the California Academy of Sciences*, 56:1–324.

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., and Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8):1494–1512.

Hedin, M. and 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. *Molecular Phylogenetics and Evolution*, 41(2):454–471.

Homann, H. (1971). Die Augen der Araneae. *Zeitschrift für Morphologie der Tiere*, 69(3):201–272.

Hormiga, G. and Griswold, C. E. (2014). Systematics, Phylogeny, and Evolution of Orb-Weaving Spiders. *Annual Review of Entomology*, 59(1):487–512.

Hölldobler, B. and Wilson, E. O. (1990). *The ants*. Harvard University Press.

Ihaka, R. and Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, 5(3):299.

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., Ho, S. Y., Faircloth, B. C., Nabholz, B., Howard, J. T., and others (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346(6215):1320–1331.

Katoh, K. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, 33(2):511–518.

Kergoat, G. J., Soldati, L., Anne-Laure Clamens, Jourdan, H., Jabbour-Zahab, R., Genson, G., Bouchard, P., and Condamine, F. L. (2014). Higher level molecular phylogeny of darkling beetles (Coleoptera: Tenebrionidae): Darkling beetle phylogeny. *Systematic Entomology*, 39(3):486–499.

King, G. F. and Hardy, M. C. (2013). Spider-Venom Peptides: Structure, Pharmacology, and Potential for Control of Insect Pests. *Annual Review of Entomology*, 58(1):475–496.

Knowles, L. L. and Kubatko, L. S. (2011). *Estimating species trees: practical and theoretical aspects*. John Wiley and Sons.

Kocot, Moroz, L., Citarella, M., and Halanych, K. (2013). PhyloTreePruner: A Phylogenetic Tree-Based Approach for Selection of Orthologous Sequences for Phylogenomics. *Evolutionary Bioinformatics*, page 429.

Kocot, K. M., Cannon, J. T., Todt, C., Citarella, M. R., Kohn, A. B., Meyer, A., Santos, S. R., Schander, C., Moroz, L. L., Lieb, B., and Halanych, K. M. (2011). Phylogenomics reveals deep molluscan relationships. *Nature*, 477(7365):452–456.

Kozlov, A. M., Aberer, A. J., and Stamatakis, A. (2015). ExaML version 3: a tool for phylogenomic analyses on supercomputers. *Bioinformatics*, 31(15):2577–2579.

Kück, P. (2009). ALICUT: a Perlscript which cuts ALISCORE identified RSS. *Department of Bioinformatics, Zoologisches Forschungsmuseum A. Koenig (ZFMK), Bonn, Germany, version*, 2.

Kück, P. and Meusemann, K. (2010). FASconCAT: Convenient handling of data matrices. *Molecular Phylogenetics and Evolution*, 56(3):1115–1118.

Kück, P. and Struck, T. H. (2014). BaCoCa – A heuristic software tool for the parallel assessment of sequence biases in hundreds of gene and taxon partitions. *Molecular Phylogenetics and Evolution*, 70:94–98.

LaPolla, J. S., Dlussky, G. M., and Perrichot, V. (2013). Ants and the Fossil Record. *Annual Review of Entomology*, 58(1):609–630.

Leache, A. D. and Rannala, B. (2011). The Accuracy of Species Tree Estimation under Simulation: A Comparison of Methods. *Systematic Biology*, 60(2):126–137.

Ledford, J. M. and Griswold, C. E. (2010). A study of the subfamily Archoleptonetinae (Araneae, Leptonetidae) with a review of the morphology and relationships for the Leptonetidae. *Zootaxa*, 2391:1–32.

Legendre, F., Nel, A., Svenson, G. J., Robillard, T., Pellens, R., and Grandcolas, P. (2015). Phylogeny of Dictyoptera: Dating the Origin of Cockroaches, Praying Mantises and Termites with Molecular Data and Controlled Fossil Evidence. *PLOS ONE*, 10(7):e0130127.

Lehtinen, P. T. (1967). Classification of the cribellate spiders and some allied families, with notes on

the evolution of the suborder Araneomorpha. In *Annales zoologici fennici*, pages 199–468. Societas Zoologica Botanica Fennica Vanamo.

Lemmon, A. R., Brown, J. M., Stanger-Hall, K., and Lemmon, E. M. (2009). The Effect of Ambiguous Data on Phylogenetic Estimates Obtained by Maximum Likelihood and Bayesian Inference. *Systematic Biology*, 58(1):130–145.

Lemmon, E. M. and Lemmon, A. R. (2013). High-Throughput Genomic Data in Systematics and Phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 44(1):99–121.

Levi, H. W. (1980). Orb-webs: primitive or specialized. In Gruber, J., editor, *Proceedings of the 8th International Congress of Arachnology*, pages 367–370.

Li, L., Stoeckert, C. J., and Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome research*, 13(9):2178–2189.

Liu, L. and Yu, L. (2011). Estimating Species Trees from Unrooted Gene Trees. *Systematic Biology*, 60(5):661–667.

Lopardo, L. and Hormiga, G. (2015). Out of the twilight zone: phylogeny and evolutionary morphology of the orb-weaving spider family Mysmenidae, with a focus on spinneret spigot morphology in symphytognathoids (Araneae, Araneoidea): Phylogeny and Morphology of Mysmenidae. *Zoological Journal of the Linnean Society*, 173(3):527–786.

McKenna, D. D., Sequeira, A. S., Marvaldi, A. E., and Farrell, B. D. (2009). Temporal lags and overlap in the diversification of weevils and flowering plants. *Proceedings of the National Academy of Sciences*, 106(17):7083–7088.

Meyer, B., Meusemann, K., and Misof, B. (2011). MARE: MAtrix REduction—a tool to select optimized data subsets from supermatrices for phylogenetic inference. *Bonn (Germany): Zentrum fuur molekulare Biodiversitätsforschung (zmb) am ZFMK*.

Miller, J. A., Carmichael, A., Ramírez, M. J., Spagna, J. C., Haddad, C. R., Řezáč, M., Johannesen, J., Král, J., Wang, X.-P., and Griswold, C. E. (2010). Phylogeny of entelegyne spiders: Affinities of the family Penestomidae (NEW RANK), generic phylogeny of Eresidae, and asymmetric rates of change in spinning organ evolution (Araneae, Araneoidea, Entelegynae). *Molecular Phylogenetics and Evolution*, 55(3):786–804.

Mirarab, S., Reaz, R., Bayzid, M. S., Zimmermann, T., Swenson, M. S., and Warnow, T. (2014). ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17):i541–i548.

Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., Frandsen, P. B., Ware, J., Flouri, T., Beutel, R. G., and others (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210):763–767.

Misof, B. and Misof, K. (2009). A Monte Carlo Approach Successfully Identifies Randomness in Multiple Sequence Alignments : A More Objective Means of Data Exclusion. *Systematic Biology*, 58(1):21–34.

Moreau, C. S. (2006). Phylogeny of the Ants: Diversification in the Age of Angiosperms. *Science*, 312(5770):101–104.

Opell, B. (1979). Revision of the genera and tropical American species of the spider family Uloboridae. Revisión de los géneros de las especies americanas tropicales de arañas de la familia Uloboridae. *Bulletin of the Museum of Comparative Zoology.*, 148(10):443–549.

Opell, B. D. (1983). A review of the genus Tangaroa (Araneae, Uloboridae). *Journal of Arachnology*, pages 287–295.

Penalver, E. (2006). Early Cretaceous Spider Web with Its Prey. *Science*, 312(5781):1761–1761.

Penney, D. and Ortuño, V. M. (2006). Oldest true orb-weaving spider (Araneae: Araneidae). *Biology Letters*, 2(3):447–450.

Penney, D., Wheater, C. P., and Selden, P. A. (2003). Resistance of spiders to Cretaceous-Tertiary extinction events. *Evolution*, 57(11):2599–2607.

Platnick, N. I., Coddington, J. A., Forster, R. R., and Griswold, C. E. (1991). Spinneret Morphology and the Phylogeny of Haplogyne Spiders (Araneae, Araneomorphae). *American Museum noviates*, 3016.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R news*, 6(1):7–11.

Polotow, D., Carmichael, A., and Griswold, C. E. (2015). Total evidence analysis of the phylogenetic relationships of Lycosoidea spiders (Araneae, Entelegynae). *Invertebrate Systematics*, 29(2):124.

Price, M. N., Dehal, P. S., Arkin, A. P., and others (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS one*, 5(3):e9490.

Prum, R. O., Berv, J. S., Dornburg, A., Field, D. J., Townsend, J. P., Lemmon, E. M., and Lemmon, A. R. (2015). A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature*, 526(7574):569–573.

Rabosky, D. L., Donnellan, S. C., Grundler, M., and Lovette, I. J. (2014). Analysis and Visualization of Complex Macroevolutionary Dynamics: An Example from Australian Scincid Lizards. *Systematic Biology*, 63(4):610–627.

Ramírez, M. J. (2000). Respiratory system morphology and the phylogeny of haplogyne spiders (Araneae, Araneomorphae). *Journal of Arachnology*, 28(2):149–157.

Ramírez, M. J. (2014). The morphology and phylogeny of dionychan spiders (Araneae: Araneomorphae). *Bulletin of the American Museum of Natural History*, 390(1):1–374.

Raven, R. J. (1985). The Spider Infraorder Mygalomorphae (Araneae): Cladistics and Systematics. *Bulletin of the American Museum of Natural History*, 182(1):1–184.

Rice, P., Longden, I., Bleasby, A., and others (2000). EMBOSS: the European molecular biology open software suite. *Trends in genetics*, 16(6):276–277.

Roure, B., Baurain, D., and Philippe, H. (2013). Impact of Missing Data on Phylogenies Inferred from Empirical Phylogenomic Data Sets. *Molecular Biology and Evolution*, 30(1):197–214.

Saez, N. J., Senff, S., Jensen, J. E., Er, S. Y., Herzig, V., Rash, L. D., and King, G. F. (2010). Spider-Venom Peptides as Therapeutics. *Toxins*, 2(12):2851–2871.

Sanderson, M. J. (2002). Estimating Absolute Rates of Molecular Evolution and Divergence Times: A Penalized Likelihood Approach. *Molecular Biology and Evolution*, 19(1):101 –109.

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., Funch, P., Enghild, J. J., Schauser, L., Andersen, S. U., Villesen, P., Schierup, M. H., Bilde, T., and Wang, J. (2014). Spider genomes provide insight into composition and evolution of venom and silk. *Nature Communications*, 5.

Schacht, K. and Scheibel, T. (2014). Processing of recombinant spider silk proteins into tailor-made materials for biomaterials applications. *Current Opinion in Biotechnology*, 29:62–69.

Scharff, N. and Coddington, J. A. (1997). A phylogenetic analysis of the orb-weaving spider family Araneidae (Arachnida, Araneae). *Zoological Journal of the Linnean Society*, 120(4):355–434.

Selden, P. A. (1996). First fossil mesothele spider, from the Carboniferous of France. *Revue suisse de Zoologie*, 2:585–96.

Selden, P. A. (2002). First British Mesozoic spider, from Cretaceous amber of the Isle of Wight, southern England. *Palaeontology*, 45:973–983.

Selden, P. A., Anderson, J. M., Anderson, H. M., and Fraser, N. C. (1999). Fossil araneomorph spiders from the Triassic of South Africa and Virginia. *Journal of Arachnology*, pages 401–414.

Selden, P. A., Dunlop, J. A., and Garwood, R. J. (2015). Carboniferous araneomorph spiders reinterpreted as long-bodied harvestmen. *Journal of Systematic Palaeontology*, (ahead-of-print):1–11.

Selden, P. A. and Gall, J.-C. (1992). A Triassic mygalomorph spider from the northern Vosges, France. *Palaeontology*, 35:211–235.

Selden, P. A. and Penney, D. (2010). Fossil spiders. *Biological Reviews*, 85(1):171–206.

Selden, P. A., Shih, C., and Ren, D. (2013). A giant spider from the Jurassic of China reveals greater diversity of the orbicularian stem group. *Naturwissenschaften*, 100(12):1171–1181.

Spagna, J. C., Crews, S. C., and Gillespie, R. G. (2010). Patterns of habitat affinity and Austral/Holarctic parallelism in dictynoid spiders (Araneae:Entelegynae). *Invertebrate Systematics*, 24(3):238.

Spagna, J. C. and Gillespie, R. G. (2008). More data, fewer shifts: Molecular insights into the evolution of the spinning apparatus in non-orb-weaving spiders. *Molecular Phylogenetics and Evolution*, 46(1):347–368.

Springer, M. S. and Gatesy, J. (2016). The gene tree delusion. *Molecular Phylogenetics and Evolution*, 94:1–33.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313.

Starrett, J., Garb, J. E., Kuelbs, A., Azubuike, U. O., and Hayashi, C. Y. (2012). Early events in the evolution of spider silk genes. *PLoS One*, 7(6):e38084.

Wahlberg, N., Wheat, C. W., and Peña, C. (2013). Timing and Patterns in the Taxonomic Diversification of Lepidoptera (Butterflies and Moths). *PLoS ONE*, 8(11):e80875.

826 Wang, B., Zhang, H., and Jarzembowski, E. A. (2013). Early Cretaceous angiosperms and beetle evolution.
827 *Frontiers in Plant Science*, 4.
828 Ward, P. S. (2014). The Phylogeny and Evolution of Ants. *Annual Review of Ecology, Evolution, and*
829 *Systematics*, 45(1):23–43.
830 Wood, H. M., Griswold, C. E., and Gillespie, R. G. (2012a). Phylogenetic placement of pelican spiders
831 (Archaeidae, Araneae), with insight into evolution of the "neck" and predatory behaviours of the
832 superfamily Palpimanoidea. *Cladistics*, 28(6):598–626.
833 Wood, H. M., Matzke, N. J., Gillespie, R. G., and Griswold, C. E. (2012b). Treating fossils as terminal
834 taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders.
835 *Systematic Biology*, page sys092.
836 World Spider Catalog (2015). World Spider Catalog, Version 16.5. Natural History Museum Bern.
837 Xia, X. (2014). Phylogenetic Bias in the Likelihood Method Caused by Missing Data Coupled with
838 Among-Site Rate Variation: An Analytical Approach. In Hutchison, D., Kanade, T., Kittler, J., Klein-
839 berg, J. M., Kobsa, A., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen,
840 B., Terzopoulos, D., Tygar, D., Weikum, G., Basu, M., Pan, Y., and Wang, J., editors, *Bioinformatics*
841 *Research and Applications*, volume 8492, pages 12–23. Springer International Publishing, Cham.

842 **Table 1.** Summary of all phylogenomic analyses. Data matrix numbers correspond to Figure 2, inset.

Lineage	Composition and Placement	Description/Characteristics
[1] Araneae	All spiders	Cosmopolitan; cheliceral venom glands, ability to produce silk from abdominal appendages (spinnerets); male pedipalps modified for sperm transfer
[2] Mesothele	Plesiomorphic sister group to all living spiders	SE Asia; mid ventrally positioned spinnerets; distinct dorsal abdominal tergites, very narrow sternum
[3] Opisthothele	The two major spider lineages	Typical terminal spinneret placement and sternal morphology
[4] Mygalomorphae	Trapdoor, baboon and funnel spiders, tarantulas, and their kin	Paraxial chelicerae with venom glands; most lead sedentary lives in burrows; lack anterior median spinnerets; often large and hirsute; two pairs of book lungs
[5] Atypoidina	Sister group to remaining mygalomorphs	Most species with vestigial abdominal tergites and unique modifications to male pedipalp
[6] Avicularioidea	All remaining mygalomorph taxa	Includes major mygalomorph families, nearly half of which are likely not monophyletic
[7] Theraphosoidina	Comprises families Theraphosidae and Barychelidae	Includes the typically large and hirsute tarantulas and baboon spiders
[8] Araneomorphae	Over 90% of all spider diversity	Anterior median spinnerets fused to form a cribellum (later lost multiple times)
[9] Paleocribellatae	Comprises single family Hypochilidae; hypothesized sister group to all other araneomorphs	Hypochilid synapomorphies, e.g., cheliceral depression; also retain a number of primitive traits including two pairs of booklungs
Neocribellatae	Remaining spider lineages	Paracribellum (complimentary spinning field to cribellum); extension of venom gland into prosoma
Austrochiliodea	Families Austrochilidae and Gradungulidae; sister group to all other neocribellate lineages	Gondwanan taxa with notched tarsal organs; typically with two pairs of booklungs – posterior pair modified as tracheae in some taxa
[10] Haplogynae	Neocribellate lineage with simple genitalia; includes spitting spiders and cellar spiders	Spinnerets lack tartipores; mating with palps inserted simultaneously; in some taxa female genital opening lacks an epigynum; chelicerae fused at base, synspermia, male palpal organ simple
[11] Entelegynae	Comprises all remaining spider lineages with complex genitalia	Female genitalia with a "flow through system" of separate copulatory and fertilization ducts; male palpal organ typically under hydraulic control

Table 1 – continued from previous page

Lineage	Composition and Placement	Description/Characteristics
Palpimanoidea	Comprises a number of enigmatic families	Araneophages with lateral scopulae on anterior legs
Eresoidea	Includes 3 families: Eresidae, Hersiliidae, Oecobiidae; sister to remaining entelegynes	Controversial superfamily; oecobiids and hersiliids share a unique attack behavior
Orbiculariae	Comprises the Deinopoidea and Araneoidea	Members of this lineage include cribellate and ecribellate orb-web weavers as well as derived araneoids that use adhesive threads to construct sheet and cob-webs
Deinopoidea	Includes the cribellate orbicularian families Uloboridae and Deinopidae	Construct cribellate orb web; long considered sister group to adhesive orb web weavers on basis of behavioral web construction data
[12] Araneoidea	Spider superfamily that includes adhesive orb web weaving taxa and others	Members of this lineage all use adhesive threads; monophyly supported by a number of spinning and other morphological characteristics
[13] RTA	Large diverse lineage of spiders that includes wolf, jumping, running, fishing, and crab spiders	Defined primarily by the presence of a projection on the male palp – the retrolateral tibial apophysis (RTA)
[14] Dionycha	Subclade of the RTA lineage, comprises about 1/3 of all spider diversity	Defined as a group based on their two clawed condition with flanking tufts of setae for adhesion to smooth surfaces
Lycosoidea	Large superfamily comprising 10 families including fishing and wolf spiders	Monophyly of this superfamily is based on a number of morphological features (not universal) including a grate-shaped tapetum, an oval-shaped calamistrum, and male palpal features

Table 2. Summary of all phylogenomic analyses. Data matrix numbers correspond to Figure 2, inset.

Data Set	#OGs	#AAs	% missing	#reps	Log Likelihood	Notes
(1) All genes	3,398	696,652	38.5%	225	-20949310.821967	ExaML AUTOF
(2) 1st reduce	1,699	410,717	26.0%	300	-14297508.033111	ExaML AUTOF
(3) 2nd reduce	850	230,582	19.6%	300	-8098715.107390	ExaML AUTOF
(4) BCC	1,699	311,756	33.6%	300	-10017456.343941	ExaML AUTOF
(5) Arthropod core OG	549	107,307	33.0%	1000	-2729523.038858	ExaML AUTOF bs in RAxML
(6) 74 taxa (-Stegodyphus)	3,398	629,566	38.8%	300	-20569138.970981	ExaML AUTOF
(7) MARE (58 taxa, 55 in-group)	1,488	351,333	19.6%	295	-9227466.065087	ExaML AUTOF
(8) ASTRAL	3,398			100		100 bootstrap reps per partition

Table 3. Posterior probabilities (PP), ages (Ma), and 95% confidence intervals (CI) for the highest posterior density (HPD) recovered by the BEAST analysis. Node numbers correspond to Figure 5. Node numbers in bold correspond to numbers in Figure 1 and Table 1.

Node	Age	HPD 95% CI	Taxonomic Group
1	340	287-398	Araneae
3	309	258-365	Opistothele
4	261	218-307	Mygalomorphae
5	108	49-192	Atypoidina
6	114	57-197	Avicularoidea
7	47	2-125	Theraphosoidina
8	276	223-330	Opistothelae
10	190	121-262	Haplogynae
11	214	154-280	Entelegynae
12	170	114-233	Araneoidea
13	139	83-201	RTA
14	86	40-139	Dionycha
15	218	53-389	
16	37	2-109	
17	79	18-163	
18	162	85-257	
19	93	47-151	
20	71	25-127	
21	48	35217	Ctenizidae
22	232	165-299	
23	160	49-254	
24	158	85-232	
25	101	28-179	
26	81	23-148	Pholcidae
27	197	137-263	
28	92	26-172	Theridiidae
29	148	96-208	
30	127	75-186	
31	100	44-160	
32	64	15-123	Tetragnathidae
33	130	81-186	

Table 3 – continued from previous page

Node	Age	HPD 95% CI	Taxonomic Group
34	107	52-165	
35	76	25-131	
36	94	49-149	
37	61	22-116	Araneidae
38	33	29312	
39	41	33420	
40	191	134-258	
41	152	64-228	
42	21	28126	Uloboridae
43	174	117-242	
44	112	60-174	
45	44	4-113	
46	92	44-149	
47	74	29-126	
48	47	34243	
49	120	68-182	
50	104	57-160	
51	71	28-121	
52	52	36130	
53	70	28-120	Lycosoidea
54	50	35735	
55	49	15-93	
56	37	27211	

848

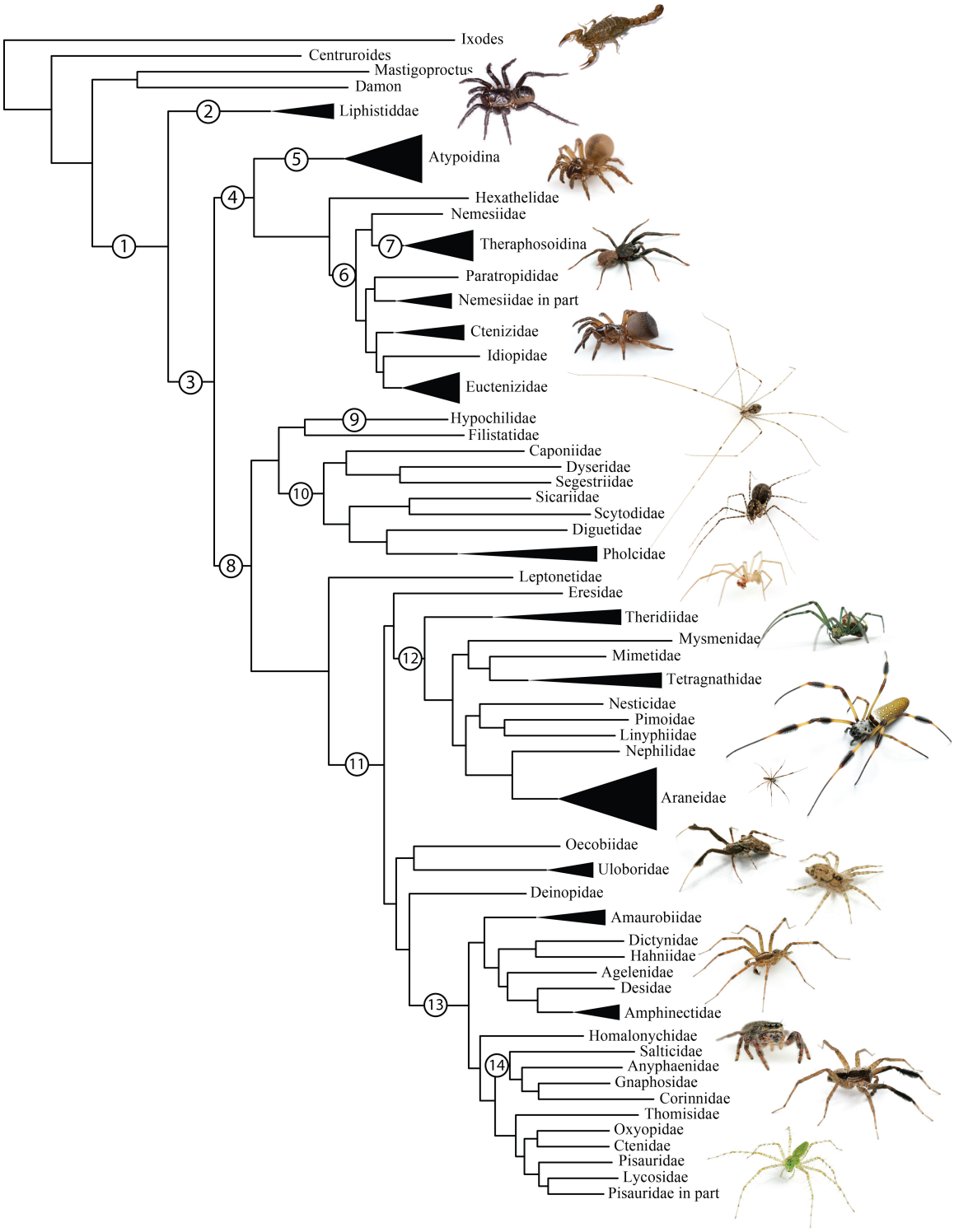

Figure 1. Summary, preferred tree, of spider relationships based on phylogenomic analyses shown in Figure 2. Numbers at nodes correspond to superscripts in Table 1. Images in descending order: Scorpion, Mesothelae, Antrodiaetidae, Paratropopididae, Ctenizidae, Pholcidae, Scytodidae, Theridiidae, Tetragnathidae, Nephilidae, Uloboridae, Amaurobiidae, Agelenidae, Salticidae, Lycosidae, Oxyopidae.

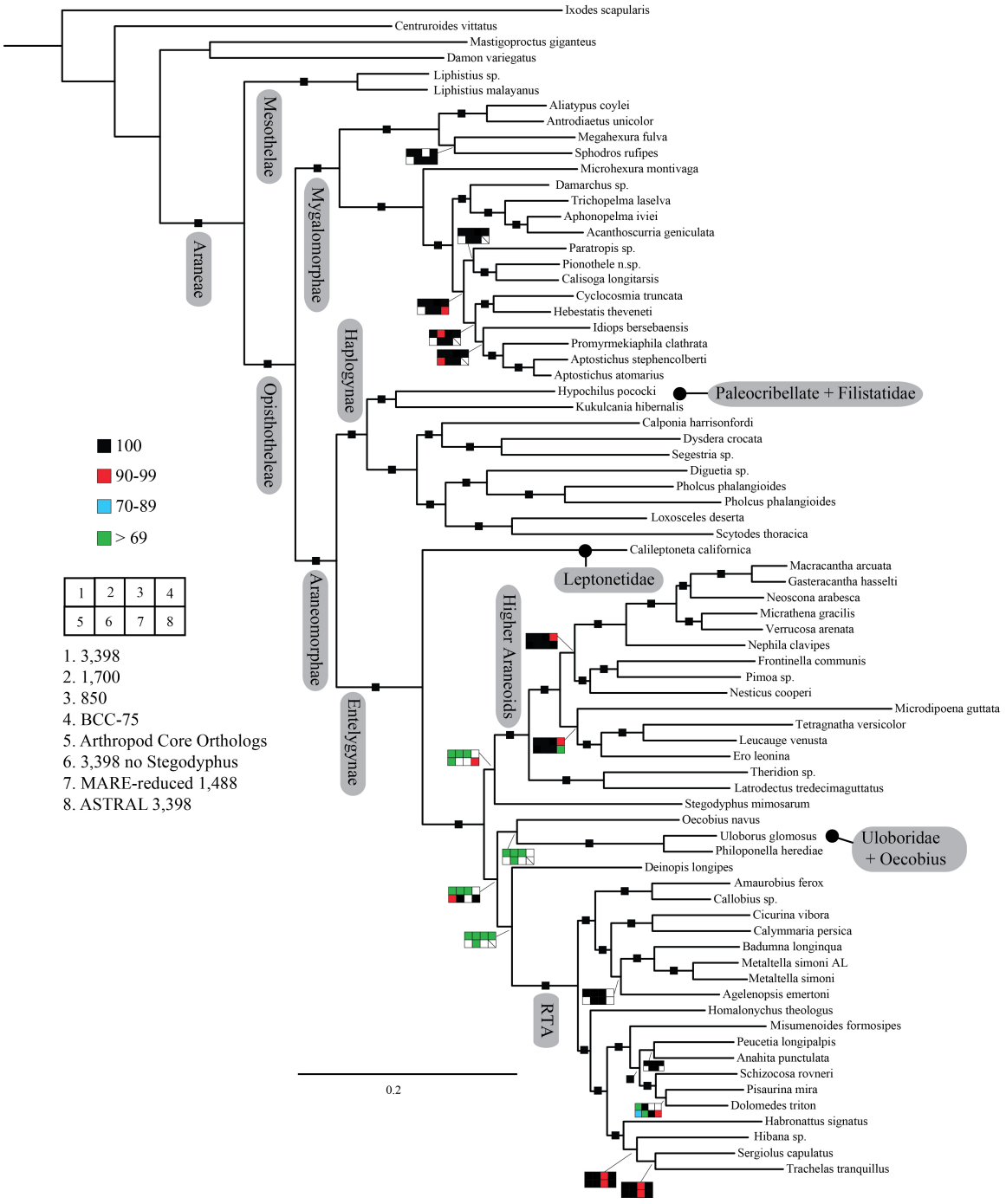

Figure 2. Summary of phylogenomic analyses (different matrices outlined in Table 2) on the phylogenetic hypothesis based on ExaML analysis of dataset 1 (3,398 OGs). Box plots indicate bootstrap value ranges for each node across matrices 1-7; single solid blocks indicate bootstrap values of 100% in all analyses.

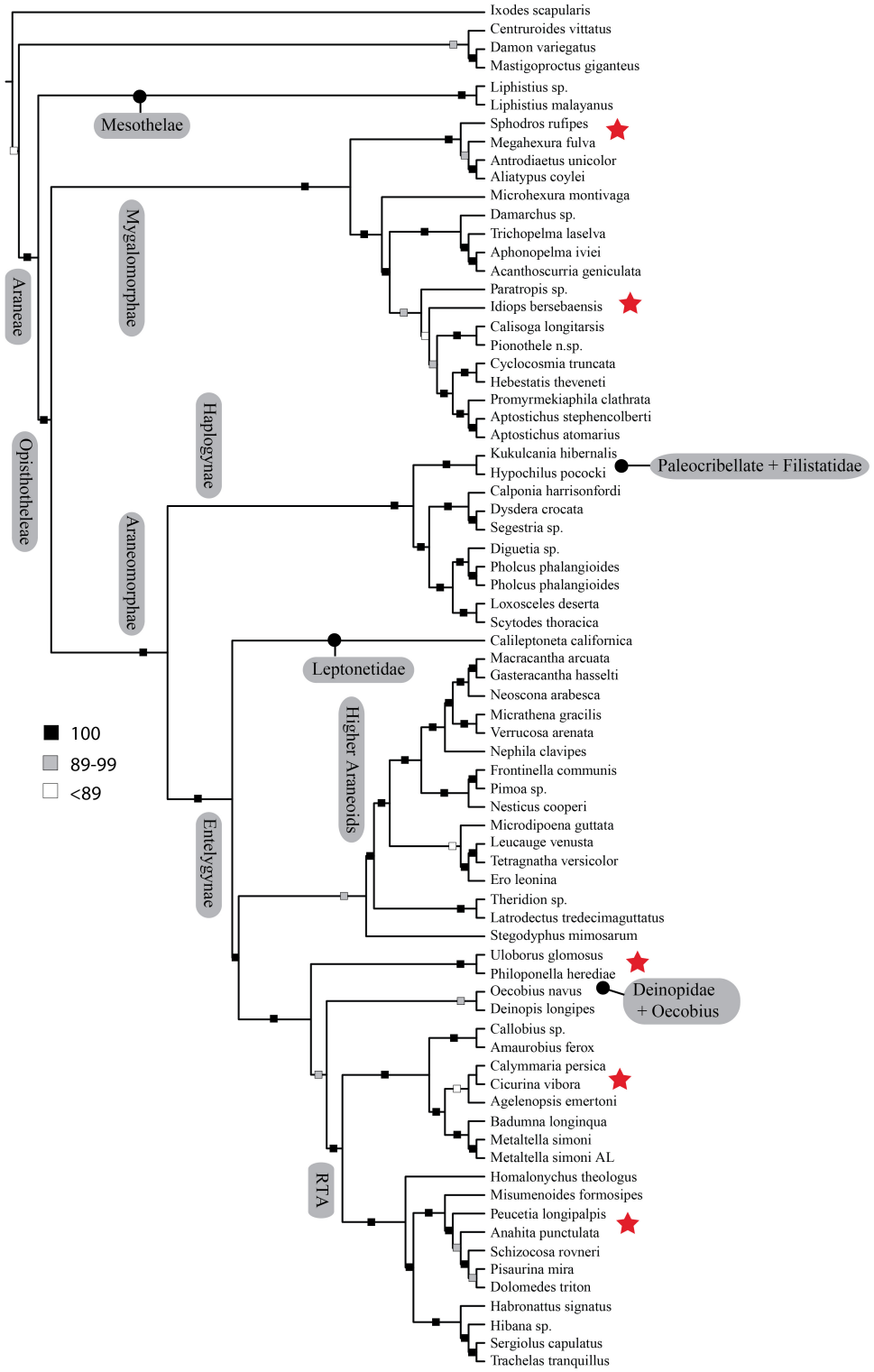

Figure 3. ASTRAL gene tree analysis of spider relationships based on 3,398 genes. Relative support value ranges reported at each node (inset legend); red stars indicate branches not congruent with tree shown in Figures 1, 2.

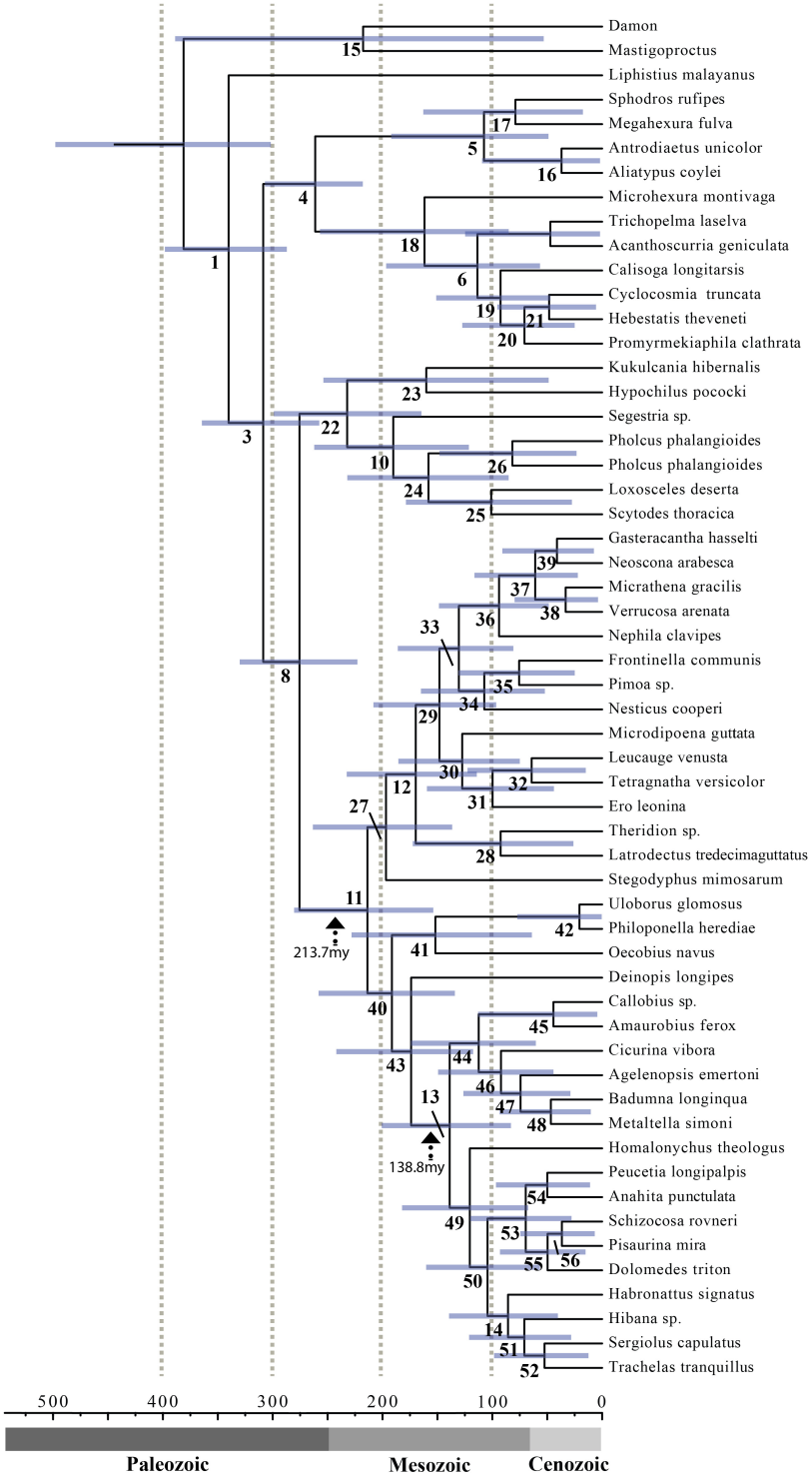

Figure 4. Chronogram resulting from two Bayesian MCMC runs performed in BEAST showing estimated divergence time for major spider lineages. Time scale on x axis; node point estimates and 95% confidence intervals (blue bars) are reported in Table 2.

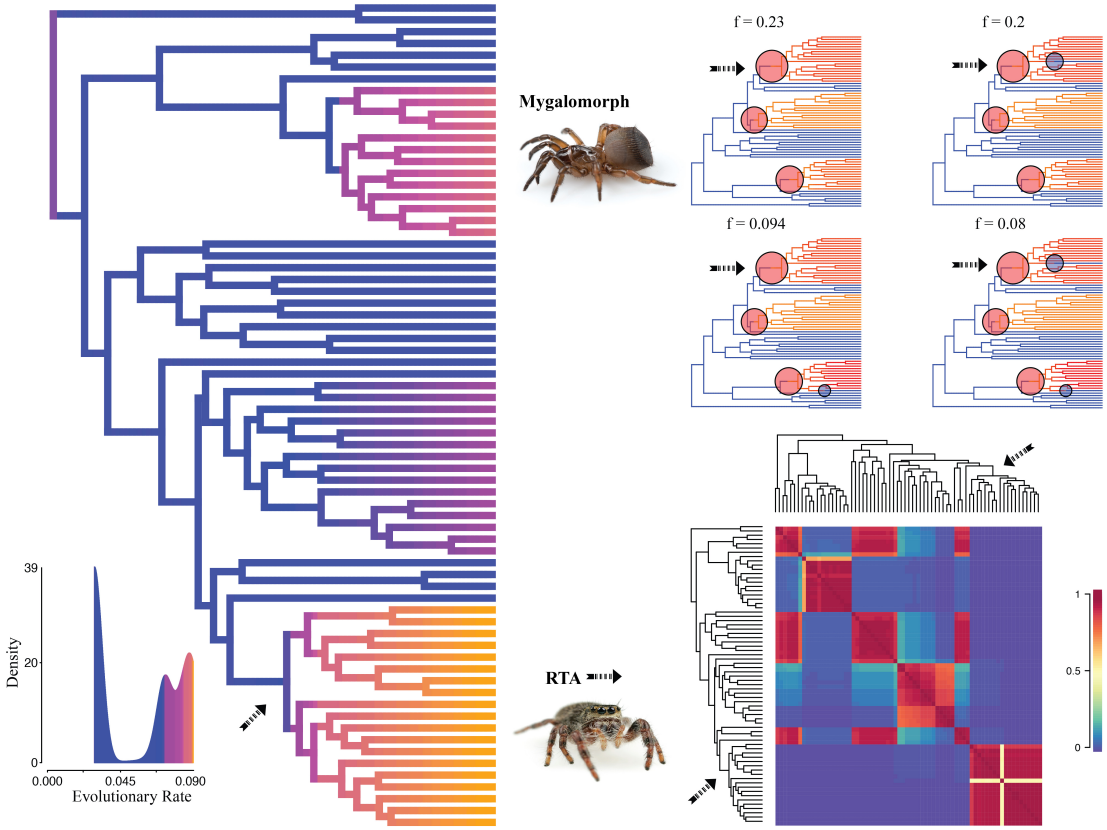

Figure 5. Time-calibrated phylogeny of spiders with branches colored by reconstructed net diversification rates (lower left). Rates on branches are means of the marginal densities of branch-specific rates. Inset histogram shows posterior density of speciation rates. Smaller phylogenies (top right) show the four distinct shift configurations with the highest posterior probability. For each distinct shift configuration, the locations of rate shifts are shown as red (rate increases) and blue (rate decreases) circles, with circle size proportional to the marginal probability of the shift. The macroevolutionary cohort analysis (lower right) displays the pairwise probability that any two species share a common macroevolutionary rate dynamic. Dashed arrow indicates position of RTA clade on each tree.

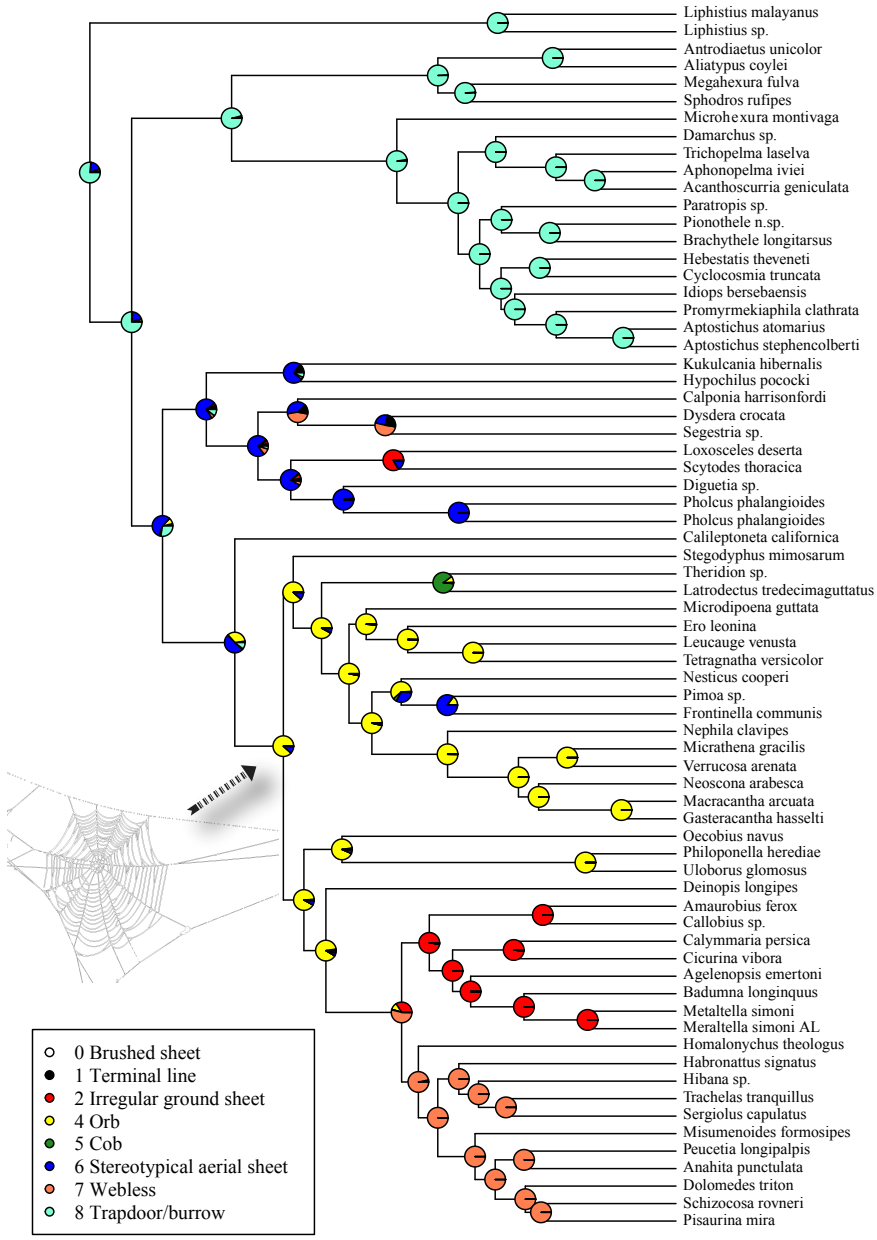

Figure 6. ML ancestral state reconstructions of web type on the time-calibrated phylogeny of spiders. Circle areas correspond to probability of ancestral states. The arrow points to one of the main diversification rate shift reconstructed by BAMM at the MRCA of the RTA clade.

SUPPLEMENTAL MATERIALS

Table S1. Specimen locality data.

Family	Genus	Specific Epithet	Specimen ID	Locality	Latitude/Longitude
Agelenidae	*Agelenopsis*	*emertoni*	AUMS5739	Auburn, Lee Co, AL, USA	32.6028°N 85.4554°W
Amaurobiidae	*Callobius*	*sp.*	AUMS11930	Muir Woods, Marin Co, CA, USA	37.8922°N 122.5717°W
Amphinectidae	*Metaltella*	*simoni*	AUMS11905	Auburn University, Lee Co, AL, USA	32.5986°N 85.4833°W
Amphinectidae	*Metaltella*	*simoni*	AUMS11915	SanDiego Co, CA, USA	NA
Antrodiaetidae	*Aliatypus*	*coylei*	AUMS56	near Monterey, Monterey Co, CA, USA	36.5714°N 121.9043°W
Antrodiaetidae	*Antrodiaetus*	*unicolor*	MY2335	near Hendersonville, Transylvania Co, NC, USA	35.2192°N 82.7794°W
Anyphaenidae	Hibana	sp.	AUMS11902	Louise Kreher Forest Ecology Preserve, Auburn, Lee Co, AL, USA	32.6603°W 85.4836°W
Araneidae	Micrathena	gracilis	AUMS5620	Dadeville, Co, AL, USA	32.8318°N 85.7636°W
Araneidae	Verrucosa	arenata	AUMS11901	Louise Kreher Forest Ecology Preserve, Auburn, Lee Co, AL, USA	32.6603°N 85.4836°W
Atypidae	Sphodros	rufipes	AUMS146	EV Smith Research Station, Macon Co, AL, USA	32.4257°N 85.9015°W
Barychelidae	Trichopelma	laselva	AUMS8485	La Selva, CR	10.4295°N 84.0095°W
Caponiidae	Calponia	harrisonfordi	1347-JML-001	Russian Ridge Open Space Park, San Mateo Co, CA, USA	37.3149°N 122.1872°W
Ctenidae	Anahita	punctulata	AUMS11932	Auburn, Lee Co, AL, USA	32.6028°N 85.4554°W

Supplemental Table 1 – continued from previous page

Family	Genus	Specific Epithet	Specimen ID	Locality	Latitude/Longitude
Ctenizidae	Cyclocosmia	truncata	AUMS120	Grove Hill, Auburn, Lee Co, AL, USA	32.5841°N 85.4604°W
Ctenizidae	Hebestatis	theveneti	AUMS152	Tuolumne Co, CA, USA	38.0453°N 119.9747°W
Deinopidae	Deinopis	longipes	AUMS8783	La Selva, CR	10.4295°N 84.0095°W
Desidae	Badumna	longinquus	AUMS11914	Petaluma, Sonoma Co, CA, USA	38.2247°N 122.6264°W
Dictynidae	Cicurina	vibora	HED003	Temples of Thor Cave, Williamson Co, TX, USA	NA
Diguetidae	Diguetia	sp.	AUMS11919	Iperial Co, CA, USA	NA
Dipluridae	Microhexura	montivaga	HED001	Mt Gibbes, Mt Mitchell SP, Yancey Co, NC, USA	35.7394°N 82.2850°W
Dysderidae	Dysdera	crocata	AUMS11924	Golden Gate Park, San Francisco Co, CA, USA	37.7686°N 122.4659°W
Euctenizidae	Aptostichus	atomarius	MY4002	Winchester, Riverside Co, CA, USA	33.7069°N 117.0844°W
Euctenizidae	Aptostichus	stephencolberti	AUMS20	Marina Dunes State Park, Monterey Co, CA, USA	36.7035°N 121.8068°W
Euctenizidae	Promyrmekiaphila	atathrata	AUMS5761	near Cazadero, Sonoma Co, CA, USA	38.5231°N 123.0845°W
Filistatidae	Kukulcania	hibernalis	AUMS8597	Auburn, Lee Co, AL, USA	32.6094°N 85.4544°W
Gnaphosidae	Sergiolus	capulatus	AUMS5674	Opelika, Lee Co, AL, USA	32.6887°N 85.4012°W
Hahniidae	Calymmaria	persica	AUMS11926	Petaluma, Sonoma Co, CA, USA	38.2247°N 122.6264°W
Homalonychidae	Homalonychus	theologus	AUMS11917	Imperial Co, CA, USA	NA
Hypochilidae	Hypochilus	pococki	AUMS155	Laurel Falls Trail, Sevier Co, TN, USA	35.6782°N 83.5929°W
Idiopidae	Idiops	bersebaensis	AUMS6746	Namibia, Africa	17.0021°S 13.2445°E

Supplemental Table 1 – continued from previous page

Family	Genus	Specific Epithet	Specimen ID	Locality	Latitude/Longitude
Leptonetidae	Calileptoneta	californica	1348-JML-001	Mount Diablo State Park, Contra Co, CA, USA	37.8745°N 121.9616°W
Linyphiidae	Frontinella	communis	AUMS5733	Auburn University, Lee Co, AL, USA	32.5987°N 85.4835°W
Liphistiidae	Liphistius	sp.	AUMS5743	Thailand	NA
Lycosidae	Schizocosa	rovneri	AUMS5122	Oxford, Lafayette Co, MS, USA	32.4116°N 89.4183°W
Mecicobothriidae	Megahexura	fulva	AUMS154	East of Groveland, Tuolumne Co, CA, USA	38.0453°N 119.9747°W
Mimetidae	Ero	leonina	AUMS8817	Tuskegee National Forest, Macon Co, AL, USA	32.4418°N 85.6328°W
Nemesiidae	Brachythele	longitarsis	AUMS47	Jasper Ridge, San Mateo Co, CA, USA	37.4132°N 122.2050°W
Nemesiidae	Damarchus	sp.	AUMS5127	Ramnagar Nepal	NA
Nemesiidae	Pionothele	n.sp.	AUMS6718	Namibia, Africa	23.5698°S 15.0398°E
Nephilidae	Nephila	clavipes	AUMS5736	Donald E. Davis Arboretum, Auburn University, Lee Co, AL, USA	32.5950°N 85.4834°W
Nesticidae	Nesticus	cooperi	HED002	Nantahala River Gorge, Swain Co, NC, USA	35.4111°N 83.5220°W
Oecobiidae	Oecobius	navus	AUMS5741	Auburn University, Lee Co, AL, USA	32.6000°N 85.4823°W
Oxyopidae	Peucetia	longipalpis	AUMS5740	Opelika, Lee Co, AL, USA	32.6887°N 85.4012°W
Paratropididae	Paratropis	sp.	AUMS148	Socotá, Departamento Boyaca, Columbia	6.08976°N 72.6198°W
Pholcidae	Pholcus	phalangioides	AUMS5732	Auburn University, Lee Co, AL, USA	32.6000°N 85.4823°W

Supplemental Table 1 – continued from previous page

Family	Genus	Specific Epithet	Specimen ID	Locality	Latitude/Longitude
Pholcidae	Pholcus	phalangioides	1304-JML-001	Petaluma, Sonoma Co, CA, USA	38.2247°N 122.6264°W
Pimoidae	Pimoa	sp.	AUMS14951	Del Norte Co, CA	NA
Pisauridae	Dolomedes	triton	AUMS11906	Opelika, Lee Co, AL, USA	32.6544°N 85.3973°W
Salticidae	Habronattus	signatus	HED004	Ocotillo, Imperial Co, CA, USA	32.7386°N 115.9941°W
Scytodidae	Scytodes	thoracica	AUMS5673	Grove Hill, Lee Co, AL, USA	32.5841°N 85.4604°W
Segestriidae	Segestria	sp.	AUMS11925	Golden Gate Park, San Francisco, CA, USA	37.7686°N 122.4659°W
Sicariidae	Loxosceles	deserta	1346-JML-001	Yarnell, Yavapai Co, AZ	32.0756°N 110.6258°W
Tetragnathidae	Leucauge	venusta	AUMS11903	Chewacla State Park, Auburn, Lee Co, AL, USA	32.5536°N 85.4845°W
Tetragnathidae	Tetragnatha	versicolor	AUMS5738	Donald E. Davis Arboretum, Auburn University, Lee Co, AL, USA	32.5950°N 85.4834°W
Theraphosidae	Aphonopelma	iviei	APH2038	East of Plymouth, Amador Co, CA, USA	38.4759°N 120.8230°W
Theridiidae	Theridion	sp.	AUMS5737	Donald E. Davis Arboretum, Auburn University, AL, USA	32.5950°N 85.4834°W
Thomisidae	Misumenoides	formosipes	AUMS6454	Opelika, Lee Co, AL, USA	32.6887°N 85.4012°W
Uloboridae	Philopenella	herediae	AUMS8784	La Selva, CR	10.4295°N 84.0095°W
Uloboridae	Uloborus	glomosus	AUMS11904	Chewacla State Park, Auburn, Lee Co, AL, USA	32.5536°N 85.4845°W

851 **Table S2.** Transcriptome sequence and assembly data.

Family	Species	Reads	Contigs	Ave. Len.	TransDecoder	ARTH	SPID
Tetragnathidae	Leucauge venusta	15567091	127727	919	51004	988	3972
Pimoidae	Pimoa sp.	46352103	160889	978.3	47496	998	3915
Araneidae	Verrucosa arenata	14468378	94721	911.9	41450	976	3993
Araneidae	Gasteracantha hasselti	12564452	50554	858.2	19944	978	4020
Theridiidae	Latrodectus tredecimagut-tatus	27605467	10259	1110.6	9266	904	3399
Araneidae	Macracantha arcuata	17523883	34411	787.9	16560	949	3826
Mysmenidae	Microdipoena guttata	16972695	156439	617.5	70785	998	3940
Araneidae	Neoscona arabesca	28551664	161311	745.6	47096	992	3834
Linyphiidae	Frontinella communis	28476743	27195	757.1	11753	862	3126
Araneidae	Micrathena gracillis	56963267	38210	856.6	16269	935	3719
Nephilidae	Nephila clavipes	34853551	32305	753.9	15200	935	3653
Nesticidae	Nesticus cooperi	20188741	41169	714	15744	945	3715
Tetragnathidae	Tetragnatha versicolor	33465090	34328	815.4	15917	938	3672
Theridiidae	Theridion sp.	37459365	24669	733.7	10013	800	2810
Uloboridae	Uloborus glomosus	12362545	114137	657	40782	970	3810
Deinopidae	Deinopis longipes	39222056	36976	705.7	12764	880	3375
Uloboridae	Philoponella herediae	47234871	40659	655.9	16468	947	3656
Eresidae	Stegodyphus mimosarum	NA	NA	NA	26888	965	4055
Oecobiidae	Oecobius navus	25031200	24989	695.4	12582	881	3183
Leptonetidae	Calileptoneta californica	59806212	37641	582	11687	875	3240
Caponiidae	Calponia harrisonfordi	59921781	14397	542.1	4297	551	1604
Diguetidae	Diguetia sp.	19838746	6816	306.9	1368	309	457
Dysderidae	Dysdera crocata	1718572	18286	452.8	5058	566	932

Supplemental Table 2 – continued from previous page

Family	Species	Reads	Contigs	Ave. Len.	TransDecoder	ARTH	SPID
Sicariidae	Loxosceles deserta	61963685	40702	582.5	11078	876	3256
Pholcidae	Pholcus phalangioides	58019035	40396	760.8	11883	902	2129
Segestriidae	Segestria sp.	38407502	191839	708.3	54753	987	3804
Filistatidae	Kukulcania hibernalis	42693292	29491	702.4	11121	835	3141
Pholcidae	Pholcus phalangioides	24861584	17870	556.9	6698	703	3183
Scytodidae	Scytodes thoracica	30924460	59599	855.7	18223	962	3657
Hypochilidae	Hypochilus pococki	25747925	19793	540.2	7561	796	2860
Mimetidae	Ero leonina	32363996	78785	801.7	24565	980	3971
Amphinectidae	Metaltella simoni	14728368	42301	574.7	18716	889	3039
Ctenidae	Anahita punctulata	44538794	75653	483.4	23111	961	3407
Anyphaenidae	Hibana sp.	40346328	55291	615.1	15212	909	2789
Desidae	Badumna longinquus	16455974	85746	661.9	34737	968	3557
Amaurobiidae	Callobius sp.	26240373	62055	499	21591	919	3401
Pisauridae	Dolomedes triton	13590617	81075	625.3	32520	951	3209
Amphinectidae	Metaltella simoni	12935049	62734	612.4	23086	950	3174
Thomisidae	Misumenoides formosipes	25351927	58396	570.4	21546	810	2161
Amaurobiidae	Amaurobius ferox	12549070	202311	530.8	35361	976	3247
Pisauridae	Pisaurina mira	6887470	129530	563.8	27116	958	3219
Corrinidae	Trachelas tranquillus	15506968	118533	467.9	30037	950	3546
Agelenidae	Agelenopsis emertoni	27264400	20517	719.7	8192	751	2613
Dictynidae	Cicurina vibora	29071083	175943	412.7	20256	852	3293
Salticidae	Habronattus signatus	75391275	26276	574.9	10447	843	3259
Oxyopidae	Peucetia longipalpis	23273514	18810	668.4	7994	737	2536
Lycosidae	Schizocosa rovneri	132349831	42744	871.6	14965	911	3584
Gnaphosidae	Sergiolus capulatus	32765239	28757	725.6	11340	820	2989
Hahniidae	Calymmaria persica	19286137	110707	644.9	20501	982	2165

Supplemental Table 2 – continued from previous page

Family	Species	Reads	Contigs	Ave. Len.	TransDecoder	ARTH	SPID
Homolonychidae	Homalonychus theologus	31165362	40734	462.1	13258	933	3167
Theraphosidae	Acanthoscurria geniculata	NA	NA	NA	76237	933	4249
Antrodiaetidae	Aliatypus coylei	29958173	23447	744.6	6098	695	2222
Antrodiaetidae	Antrodiaetus unicolor	32624239	14062	550.2	9850	756	2745
Theraphosidae	Aphonopelma iviei	36326210	13442	605.1	4717	596	2035
Euctenizidae	Aptostichus atomarius	27431535	14152	709.2	5795	643	2149
Euctenizidae	Aptostichus stephencol-berti	30904990	13267	779.5	5344	589	1944
Nemesiidae	Brachythele longitarsis	30773715	20721	574.8	7674	743	2635
Ctenizidae	Cyclocosmia truncata	33664901	26408	604.7	8446	736	2679
Nemesiidae	Damarchus sp.	21876221	13519	707.8	5558	623	2089
Ctenizidae	Hebestatis theveneti	40097804	16167	650	6711	713	2647
Idiopidae	Idiops berse-baensis	23040778	6270	649.1	2897	421	1151
Mecicobothriidae	Megahexura fulva	59599533	40526	671.9	15303	964	4147
Dipluridae	Microhexura montivaga	24680385	19680	635.8	8286	761	2695
Paratropidae	Paratropis sp.	18409810	9021	605	3694	473	1463
Nemesiidae	Pionothele n.sp.	20155275	5158	631.7	2283	368	970
Euctenizidae	Promyrmekiaphila clathrata	24733435	22423	669.4	8445	758	2646
Atypidae	Sphodros ru-fipes	51968533	27266	715.5	11504	915	3704
Barychelidae	Trichopelma laselva	27264400	33544	665.4	9061	807	2861
Liphistiidae	Liphistius sp.	54043289	7830	370.5	1938	333	667
Liphistiidae	Liphistius malayanus	62897982	83669	515.4	19784	941	3568
Ixodidae	Ixodes scapu-laris	NA	18810	668.4	17799	815	2726
Phrynichidae	Damon varie-gatus	64733221	83669	515.4	27304	944	3327
Buthidae	Centruroides vittatus	45691843	17788	378.2	4854	660	1099

Supplemental Table 2 – continued from previous page

Family	Species	Reads	Contigs	Ave. Len.	TransDecoder	ARTH	SPID
Thelyphonidae	Mastigoproctus giganteus	25983006	157263	623.7	43626	994	3785
Tetranychidae	Tetranychus cinnabarinus	26040173	30288	963.8	17083	938	3063

852

853 **Supplemental Table 3**

854 See /Supplemental_Material/AnnotationTable_S3.tex :too large to compile within main.tex

855

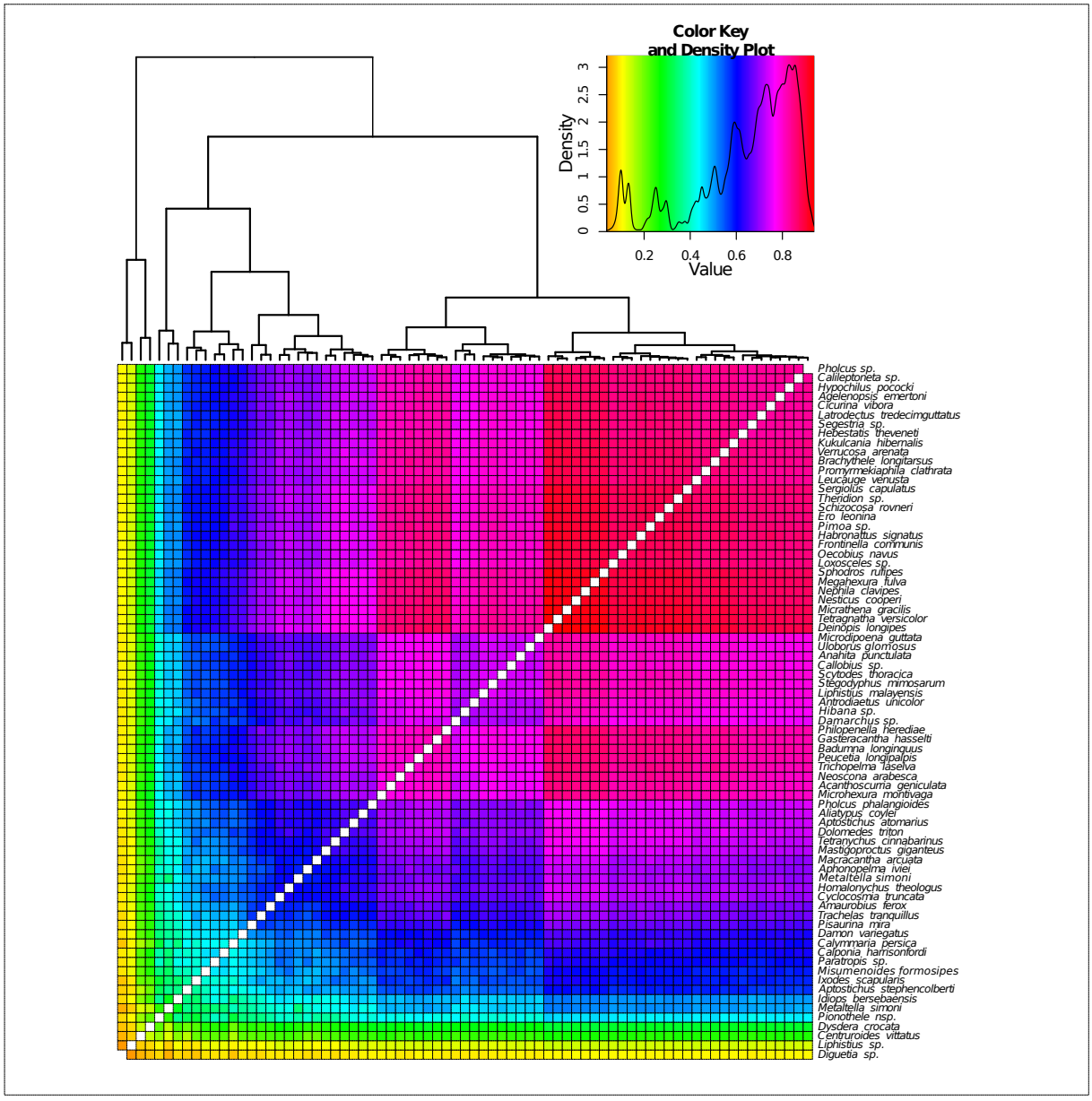

Figure S1. Clustering/heatmap analysis depicting degree of shared data between each species pair; missing data reduction (matrix 3, Table 2). Degree of positive data overlap indicated by a color-coded heatmap (yellow = low, red = high); species order from right to left in the same order as listed from top to bottom on right side of figure. Lack of phylogenetic clustering indicates bias from shared data does not explain relationships seen in phylogenomic analyses.

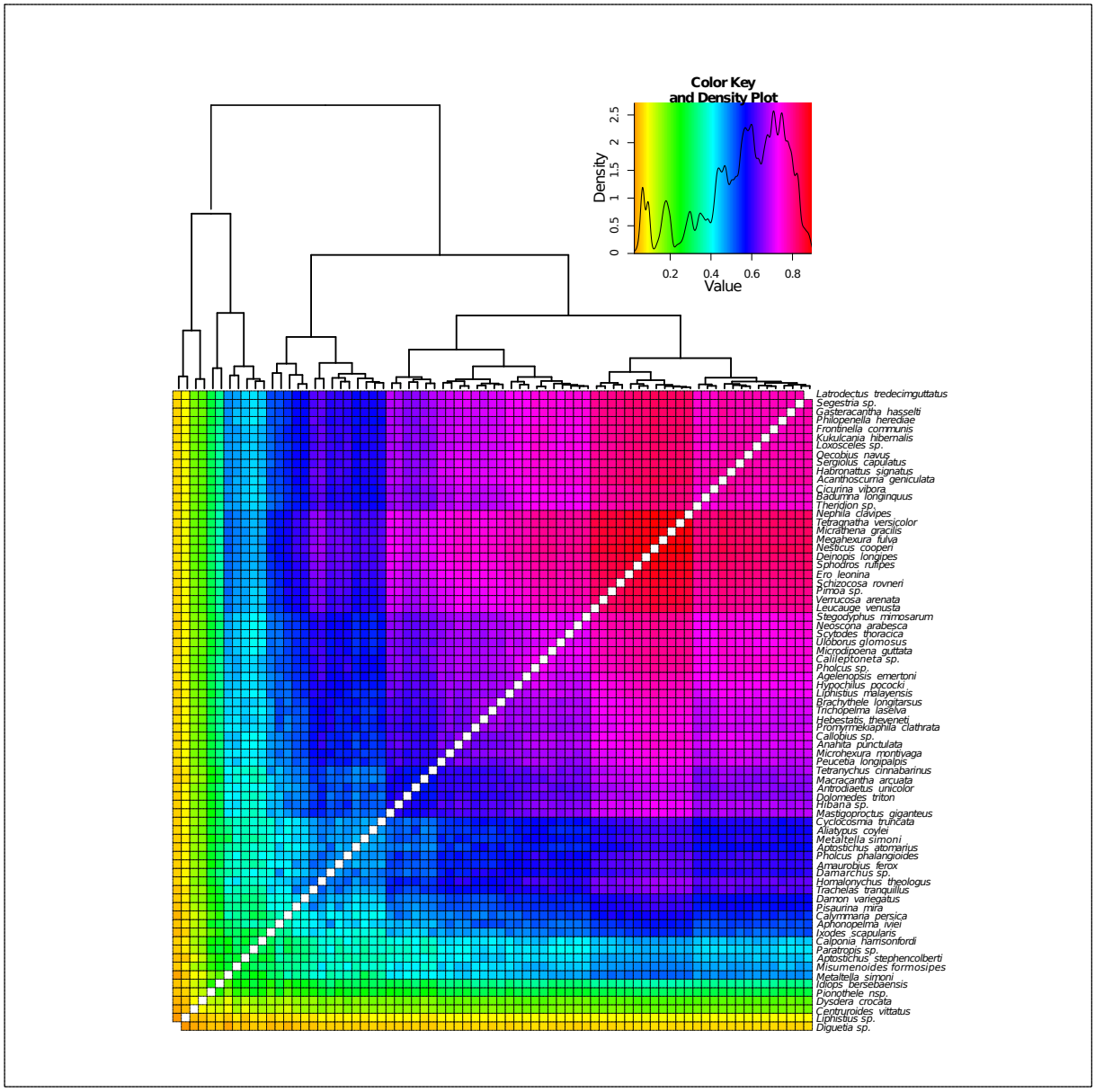

Figure S2. Clustering/heatmap analysis depicting degree of shared data between each species pair; missing data reduction (matrix 2, Table 2). Degree of positive data overlap indicated by a color-coded heatmap (yellow = low, red = high); species order from right to left in the same order as listed from top to bottom on right side of figure. Lack of phylogenetic clustering indicates bias from shared data does not explain relationships seen in phylogenomic analyses.

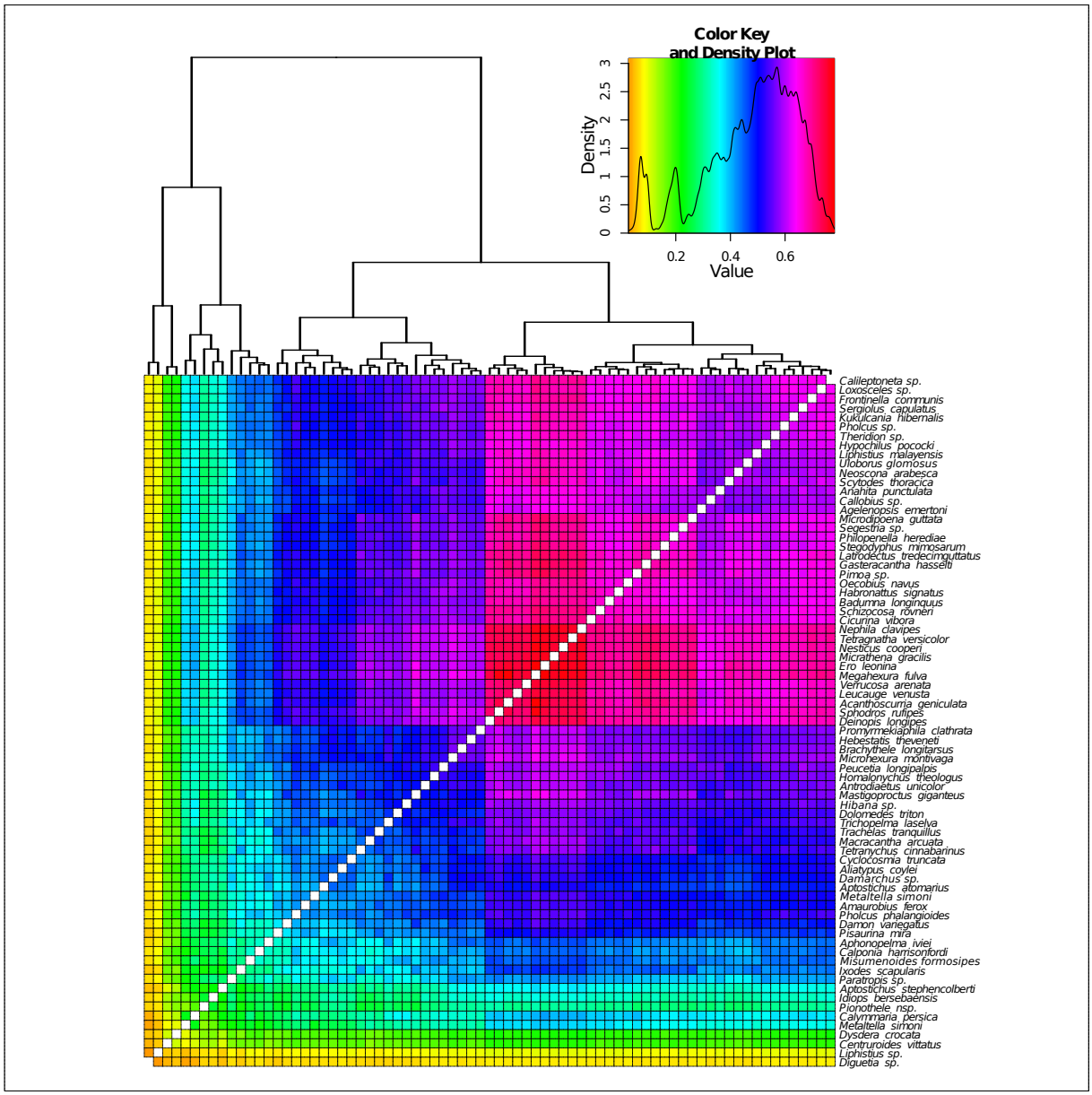

Figure S3. Clustering/heatmap analysis depicting degree of shared data between each species pair; BaCoCa reduced matrix (matrix 4, Table 2). Degree of positive data overlap indicated by a color-coded heatmap (yellow = low, red = high); species order from right to left in the same order as listed from top to bottom on right side of figure. Lack of phylogenetic clustering indicates bias from shared data does not explain relationships seen in phylogenomic analyses.

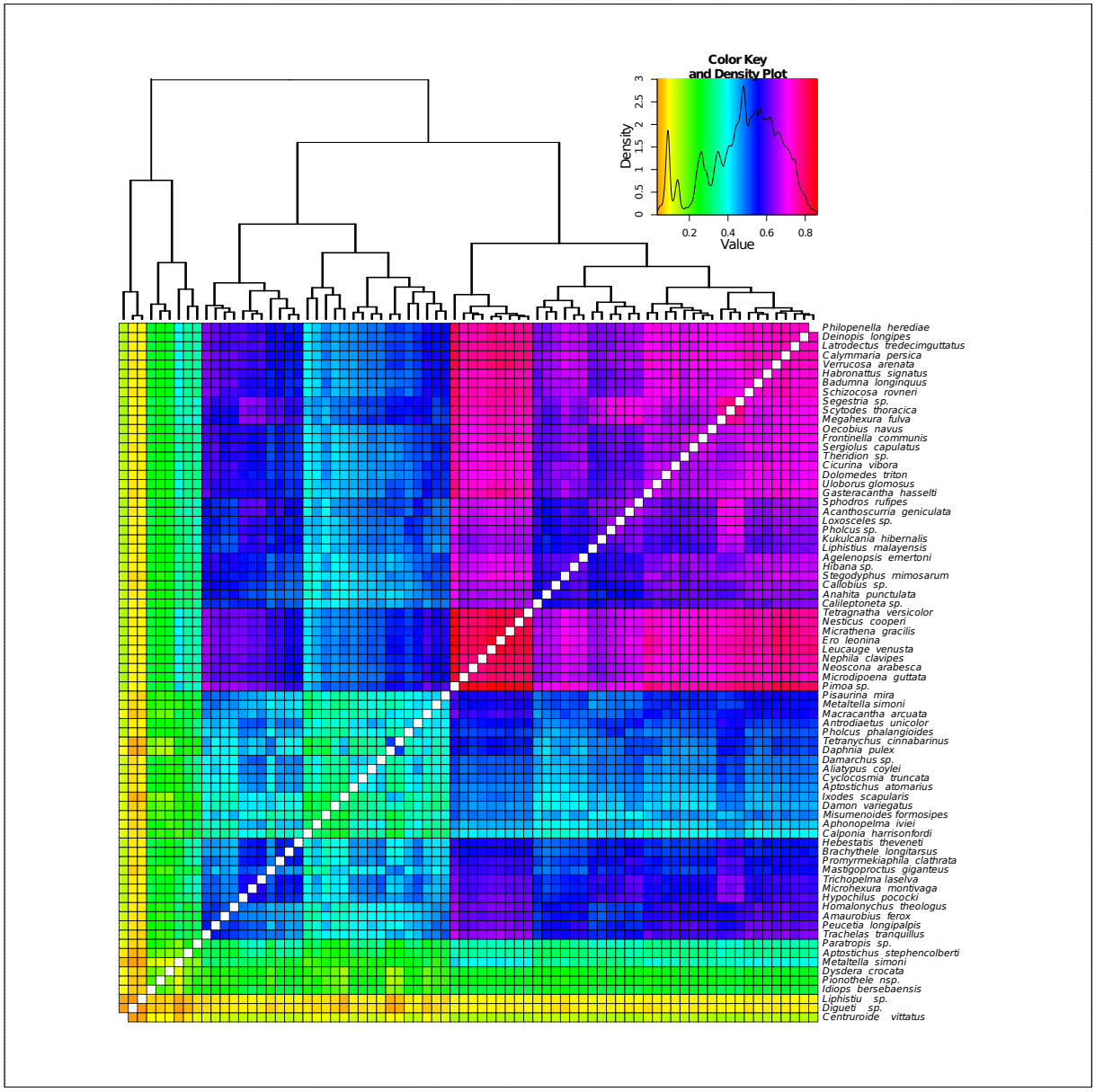

Figure S4. Clustering/heatmap analysis depicting degree of shared data between each species pair; arthropod core ortholog matrix (matrix 5, Table 2). Degree of positive data overlap indicated by a color-coded heatmap (yellow = low, red = high); species order from right to left in the same order as listed from top to bottom on right side of figure. Lack of phylogenetic clustering indicates bias from shared data does not explain relationships seen in phylogenomic analyses.

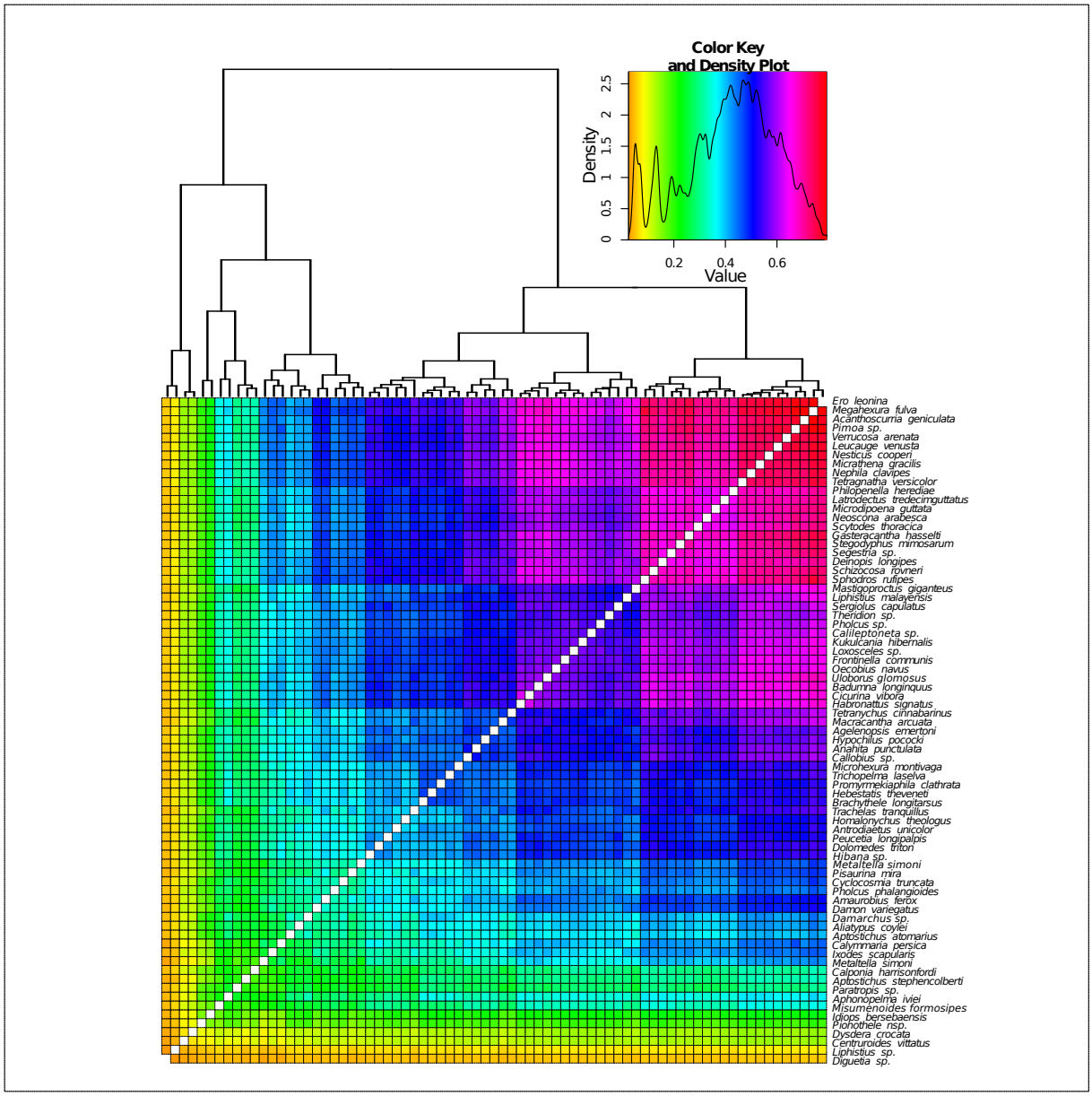

Figure S5. Clustering/heatmap analysis depicting degree of shared data between each species pair; full spider ortholog matrix (matrix 1, Table 2). Degree of positive data overlap indicated by a color-coded heatmap (yellow = low, red = high); species order from right to left in the same order as listed from top to bottom on right side of figure. Lack of phylogenetic clustering indicates bias from shared data does not explain relationships seen in phylogenomic analyses.

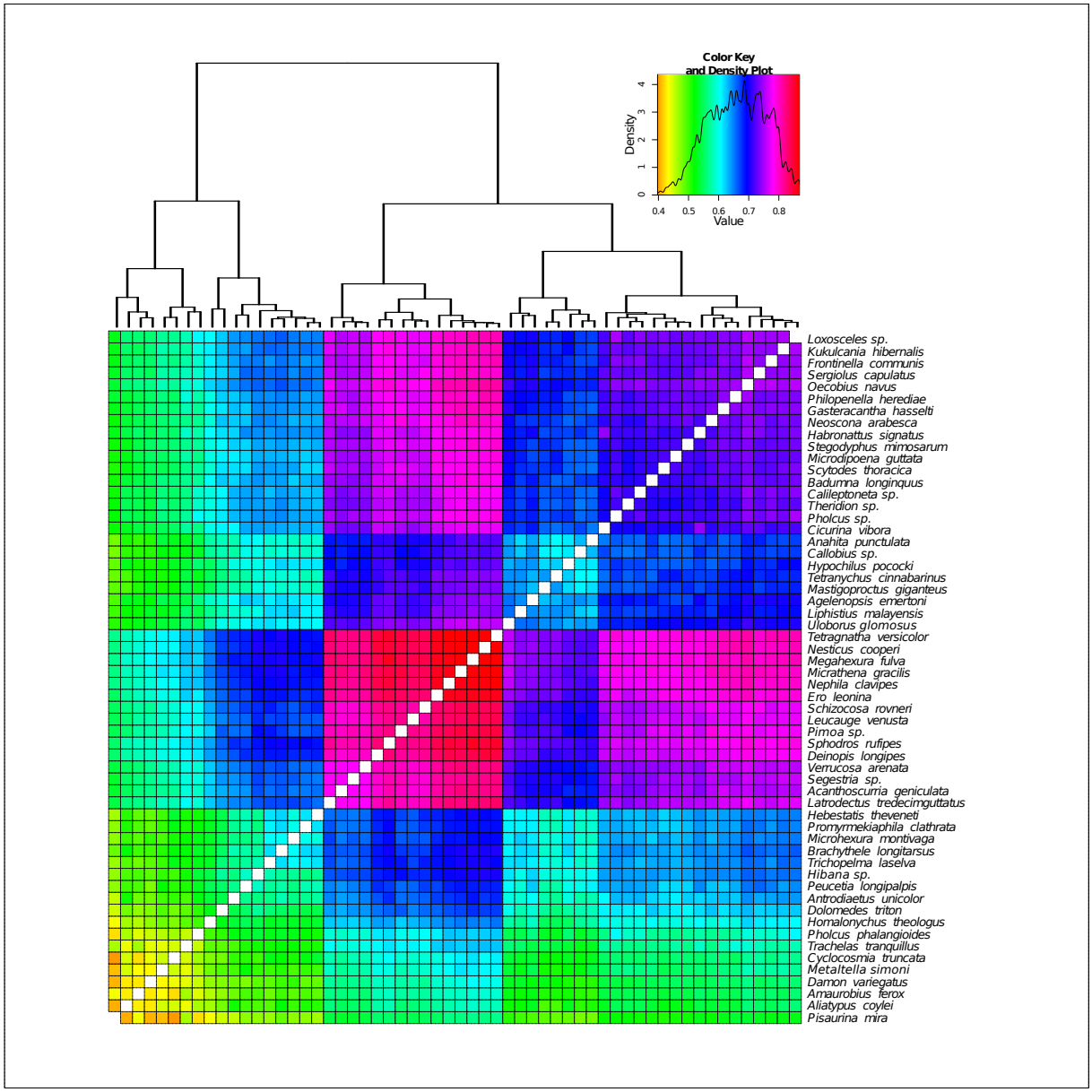

Figure S6. Clustering/heatmap analysis depicting degree of shared data between each species pair; MARE matrix (matrix 7, Table 2). Degree of positive data overlap indicated by a color-coded heatmap (yellow = low, red = high); species order from right to left in the same order as listed from top to bottom on right side of figure. Lack of phylogenetic clustering indicates bias from shared data does not explain relationships seen in phylogenomic analyses.

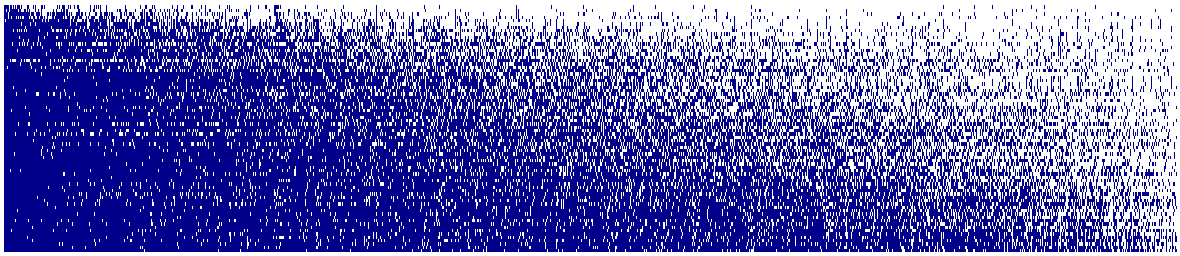

Figure S7. Gene occupancy of matrix 1 (see Table 2). Colored squares represent partitions present in matrix for each OTU (x-axis, in descending order of OTU representation from left to right) and each partition or gene (y-axis, in ascending order of partition representation).

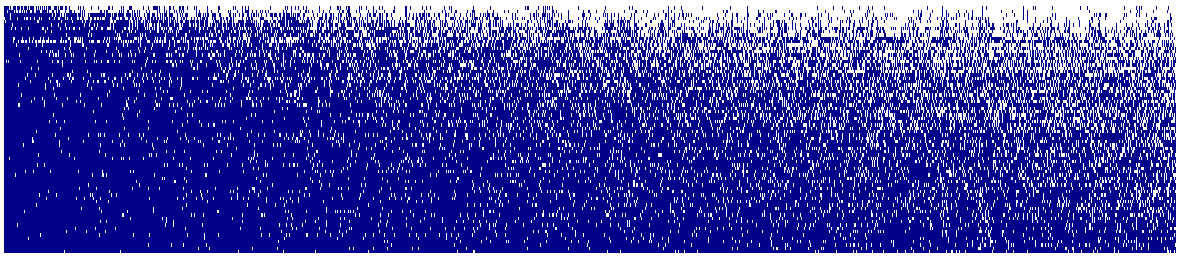

Figure S8. Gene occupancy of matrix 2 (see Table 2). Colored squares represent partitions present in matrix for each OTU (x-axis, in descending order of OTU representation from left to right) and each partition or gene (y-axis, in ascending order of partition representation).

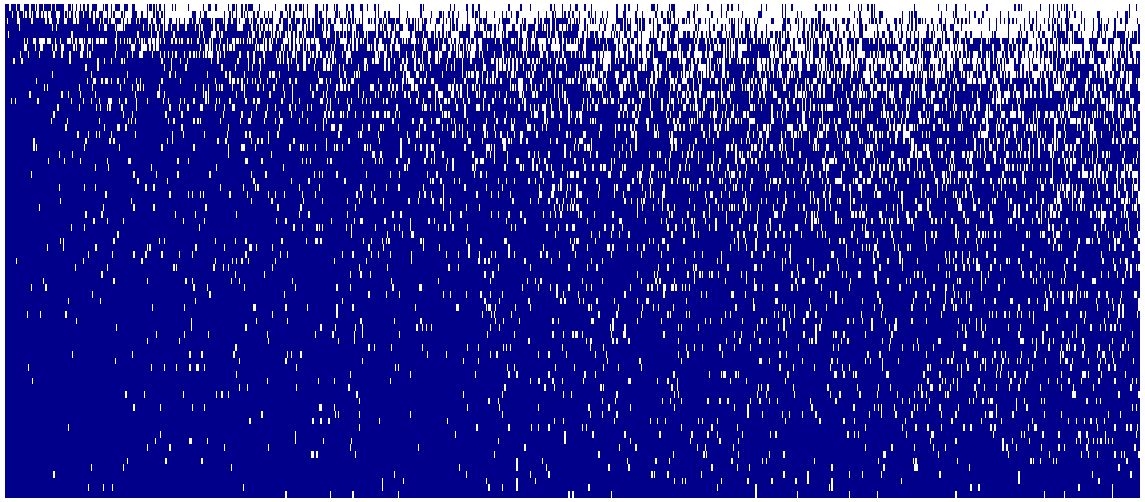

Figure S9. Gene occupancy of matrix 3 (see Table 2). Colored squares represent partitions present in matrix for each OTU (y-axis, in descending order of OTU representation from bottom to top) and each partition or gene (x-axis, in descending order of partition representation from left to right).

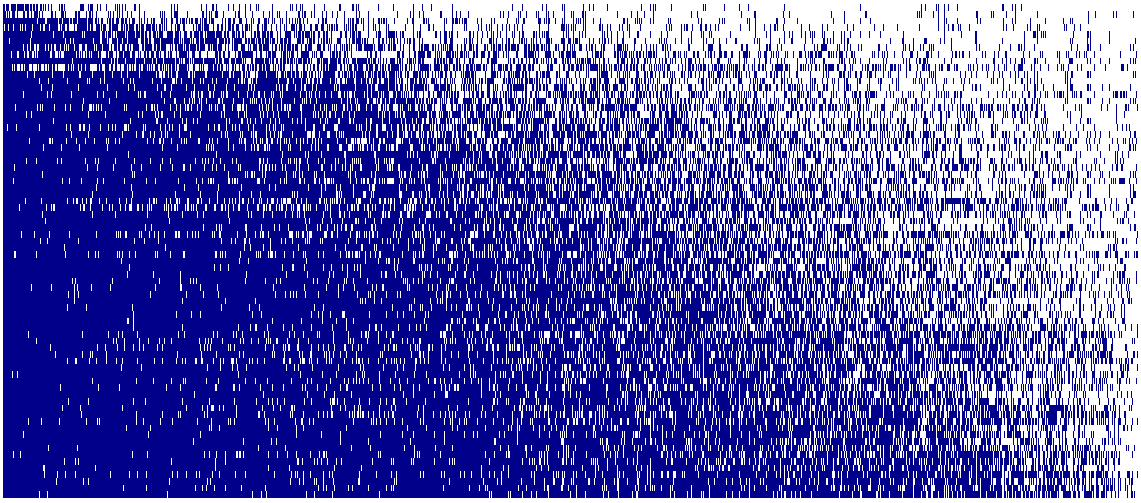

Figure S10. Gene occupancy of matrix 4 (see Table 2). Colored squares represent partitions present in matrix for each OTU (y-axis, in descending order of OTU representation from bottom to top) and each partition or gene (x-axis, in descending order of partition representation from left to right).

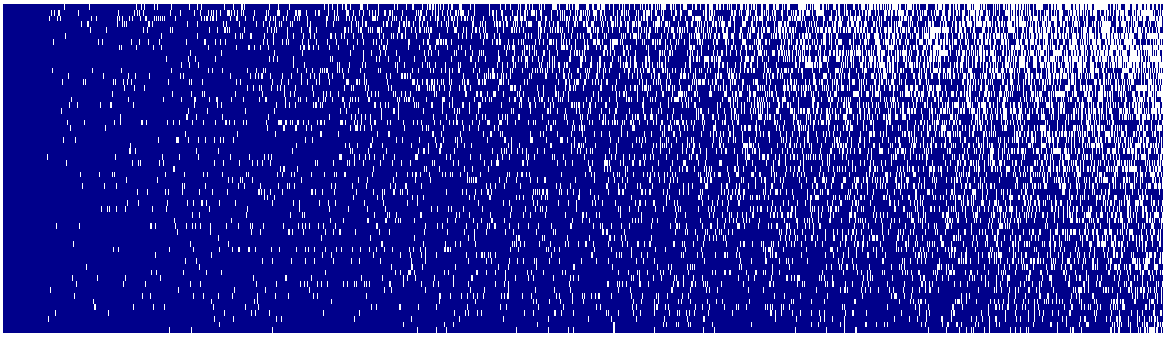

Figure S11. Gene occupancy of matrix 7 (see Table 2). Colored squares represent partitions present in matrix for each OTU (y-axis, in descending order of OTU representation from bottom to top) and each partition or gene (x-axis, in descending order of partition representation from left to right).

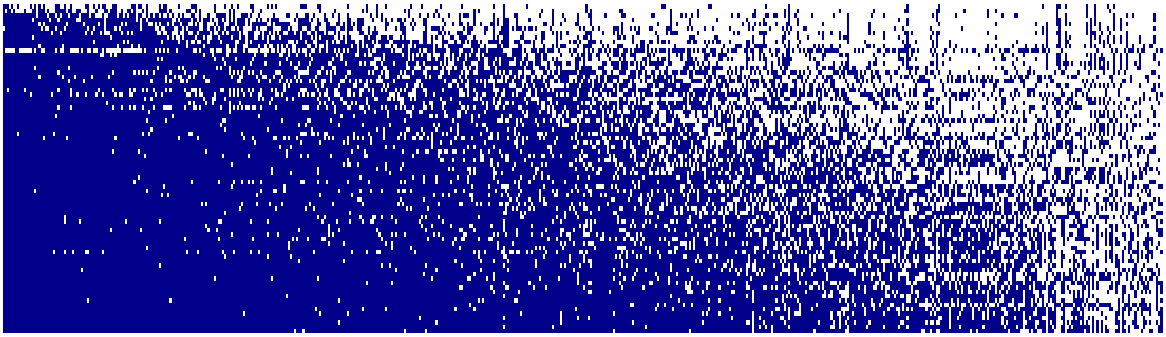

Figure S12. Gene occupancy of matrix 5 (see Table 2). Colored squares represent partitions present in matrix for each OTU (y-axis, in descending order of OTU representation from bottom to top) and each partition or gene (x-axis, in descending order of partition representation from left to right).

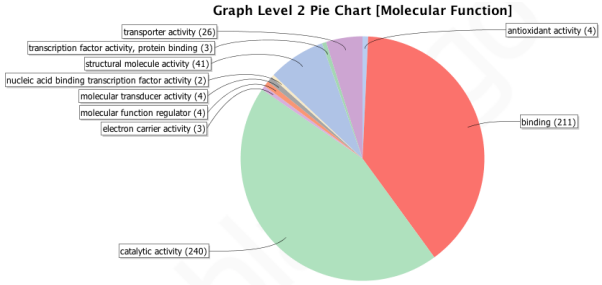

Figure S13. Gene Ontology molecular functions, levels 2 for OGs shared by Arthropod and Spider Core sets. Figures generated by Blast2GO analysis.

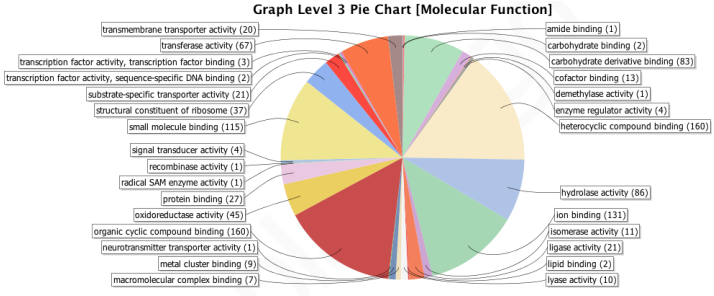

Figure S14. Gene Ontology molecular functions, level 3 for OGs shared by Arthropod and Spider Core sets. Figures generated by Blast2GO analysis.