**Supplementary File 1. Methods, Results, and Additional Analyses for:**

**Sequence-capture phylogenomics of true spiders reveals convergent evolution of respiratory systems**

Martín J. Ramírez1, Ivan L.F. Magalhaes1, Shahan Derkarabetian2, Joel Ledford3, Charles E. Griswold4, Hannah M. Wood5, Marshal Hedin6

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

*2 - Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA*

*3 - Department of Plant Biology, University of California, Davis, CA 95616 USA*

*4 - Entomology, California Academy of Sciences, San Francisco, CA, 94118, USA*

*5 - National Museum of Natural History, Smithsonian Institution, Washington DC 20560-0188, USA*

*6 - Department of Biology, San Diego State University, San Diego, CA 92182-4614, USA*

**Materials & Methods**

*Taxon sampling and DNA extraction***.** We used ultra-conserved element (UCE) target-capture data, building from the results of Wood et al. (2018) for early-diverging araneomorphs. Hedin et al. (2019) recently showed that the conserved baits of the arachnid UCE probeset (Faircloth 2017; Starrett et al. 2017) target exons, so this method is effectively exon capture. The contention that UCE capture is equivalent to exon capture applies particularly to ancient arachnid divergences, where rapidly-evolving (and alignment ambiguous) flanking introns might best be removed from analysis using bioinformatic filtering.

A mesothele and two mygalomorphs were used as outgroup taxa. We included, for the first time in a molecular phylogenetic analysis, the genus *Ectatosticta*, hypothesized as the sister genus to *Hypochilus*. We sampled all described families of Synspermiata (except for the Ochyroceratidae, recently separated from Psilodercidae), in most cases including at least two genera per family; the Synspermiata sample had two Telemidae exemplars, including the type genus *Telema*. We included all three primary austrochiloid lineages, including *Hickmania*, multiple grandugulids (both cribellate and ecribellate), and both described austrochiline genera. We included a large sample of leptonetids, with both described archoleptonetine genera (cribellate and ecribellate) and many leptonetine genera, including the type species of the family. Finally, we included a representative sample of palpimanoids, and a sparse sample of the primary Entelegynae lineages.

Most specimens were preserved for DNA studies (preserved in >95% EtOH at -80°C), and genomic DNA was extracted from leg tissue using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA). For a handful of specimens preserved in 70–80% EtOH we used standard phenol/chloroform extractions with 24-hour incubation for lysis. Extractions were quantified using a Qubit Fluorometer (Life Technologies, Inc.) and quality was assessed via agarose gel electrophoresis.

*UCE data collection, matrix assembly & filtering.* Data were collected in multiple library preparation and sequencing experiments. DNA sonication and library preparation followed methods previously used for arachnids (e.g., Starrett et al. 2017; Derkarabetian et al. 2018; Hedin et al. 2019). Target enrichment was performed using the MYbaits Arachnida 1.1K version 1 kit (Arbor Biosciences; Faircloth 2017) following the Target Enrichment of Illumina Libraries v. 1.5 protocol (http://ultraconserved.org/#protocols). Libraries were sequenced with an Illumina HiSeq 2500 with 150 bp PE reads (Brigham Young University DNA Sequencing Center), or on an Illumina HiSeq 2500 with 125 bp PE reads at the Bauer Core Facility at Harvard University.

Raw demultiplexed reads were processed with the Phyluce pipeline (Faircloth 2016). Quality control and adapter removal were conducted with the Illumiprocessor wrapper (Faircloth 2013). Assemblies were created with Velvet (Zerbino et al. 2008) and/or Trinity (Grabherr et al. 2011), both at default settings. When contigs from both assemblies were available, these were combined for probe matching, retrieving assembly-specific UCEs and increasing the number of UCEs per sample relative to using only a single assembly method. Contigs were matched to probes using 75\_75 minimum coverage\_ minimum identity values. UCE loci were aligned with MAFFT (Katoh and Standley 2013) and trimmed with Gblocks (Castresana 2000; Talavera and Castresana 2007), using conservative settings (-b1 0.5 -b2 0.85 -b3 4 -b4 8) in the Phyluce pipeline. These conservative settings should remove most flanking intron regions from individual locus alignments.

Phyluce alignments with at least 50% taxon occupancy were imported into Geneious. Alignments were sorted by percent sequence identity, and for alignments with 25% identity or less we reconstructed individual locus RAxML gene trees (Stamatakis 2004; raxmlHPC-PTHREADS-SSE3 -f a -m GTRGAMMA -N 200). Visual inspection of these gene trees confirmed that almost all such alignments included divergent single sequences or clades of sequences, which we hypothesize represent paralogous sequences (.tre files available on **Dryad**). As a preliminary paralogy filter, we only included the 578 loci with sequence identity values above 25%. We then used the arachnid set UCE annotations of Hedin et al. (2019) to flag additional potentially paralogous loci - those 44 loci flagged as including potential paralogs in Hedin et al. (2019) were excluded in a final 534\_25\_noP matrix.

*Phylogenomic Analyses*

Multiple phylogenomic analyses were conducted to explore impacts of analysis method, data partitioning, and coalescent-based approaches. We also discovered large differences in GC base composition across taxa, and conducted analyses to attempt to control for this issue.

*Maximum likelihood.* Partitioned and unpartitioned maximum likelihood (ML) analyses were conducted using IQ-TREE (Nguyen et al. 2015; Chernomor et al. 2016). Partitions and models were determined using ModelFinder (Kalyaanamoorthy et al. 2017; commands for partitions -m TESTMERGEONLY -rclusterf 10), with support estimated via 1000 ultrafast bootstrap replicates (Hoang et al. 2018). Gene- and site- concordance factors were calculated with IQ-TREE v.2.0-rc1; gene trees for individual loci were computed with IQ-TREE after model selection using ModelFinder.

*Parsimony.*Analysis of the concatenated dataset was performed in TNT under equal weights parsimony (Goloboff et al. 2008), with 1000 replicates of symmetric jackknifing.

*Testing for bias from GC content.*Weinvestigated the possible bias from heterogeneous GC content across taxa (e.g., Bossert et al. 2017). GC content was calculated with AMAS (Borowiec 2016) foreach terminal in the concatenated dataset and for each UCE locus over all terminals. Loci were sorted by GC content and divided in thirds. Loci of middle GC content were discarded. The low GC third of loci were concatenated (with AMAS) in an AT-rich dataset, and the high GC third in a GC-rich dataset. Other analyses were performed replacing all Gs and Cs with “?” in selected taxa, or recoding as RY, and analyzed with IQ-TREE and TNT.

*Testing for bias from long branch attraction.* A few selected terminals of disputed placement were removed from the dataset (see details below), with matrices re-analyzed with IQ-TREE.

*Coalescent-based species tree.* A species trees was inferred using the multispecies coalescent model implemented in SVDquartets (Chifman and Kubatko 2014), part of PAUP\* version 4.0a (Swofford 2002). This quartets-based method was originally developed for unlinked SNP data, where each site has an independent genealogy drawn from a coalescent model. Under both simulation and for empirical datasets, the method has also been shown to perform well for linked data (Chifman and Kubatko 2014), as used here. SVDquartets was used to construct a “lineage” tree, relating individual sequences; we evaluated all possible quartets (evalq=all, quartet assembly algorithm = QFM), and conducted a non-parametric bootstrap analysis (nreps=100).

*Bayesian inference.* After completion of all exploratory analyses described above, a “preferred” dataset version (GC sites for *Trogloraptor* replaced by ?) was analyzed with the multi-threaded MPI hybrid variant of ExaBayes version 1.5 (Aberer et al. 2014), with parameter statistics analyzed using Tracer version 1.7 (Rambaut et al. 2018). We used the GTR + gamma model under default parameters. The analysis reached 0.54% standard deviation of split frequencies after 106 generations, and all parameters had effective sampling sizes ESS > 200.

*Character Evolution Analyses*

*Respiratory system character scoring.* The opisthosoma or entire spiders were treated with KOH or pancreatin solution (see Álvarez-Padilla and Hormiga 2007; Ramírez 2014), which digests soft tissues while leaving the cuticular structures of the respiratory system unharmed. Structures were observed under compound or stereo microscope (**Fig. 1**). Additional scorings were obtained from published sources (see **Table S2**).

The homology of posterior book lungs with lateral tracheae, and of the third abdominal apodemes with the median tracheae, follows the embryological studies of Purcell (1909) and Ramírez (2000, 2014). This ontogenetic evidence also suggests that the single lamella of Filistatidae, the single lamella plus a tube of Austrochilinae, and the tracheae of other araneomorph spiders may be independently derived from book lungs. To allow these transformations, the organization of the posterior respiratory system (PRS) was scored as a multistate character: (1) book lungs (**Figs. 1, 3**), (2) tubular tracheae (**Figs. 1, 3**), (3) single lamella (**Figs. 1, 3**), (4) tube plus single lamella (**Figs. 1, 3**), and (5) PRS absent. To explore the effects of including “absence” as a fifth state, we also ran the analysis coding absences as missing data or pruning terminals without PRS. Alternatively, we also coded PRS as a simpler character disregarding ontogenetic evidence and assuming homology of all non-book lung configurations: (1) book lungs, (2) tracheae, (3) NA = PRS absent.

*Web and silk gland character scoring.* To trace the evolution of aerial webs we focused on a stringent definition as foraging webs from which the spiders hang in an inverted position. The spider posture in relation to webs is a classic descriptor of spider webs and behavior (Bristowe 1930; Kaston 1964; Shear 1986). It has been used as character in phylogenetic analyses of higher-level systematics of araneomorph spiders (as “web posture: 0, inverted; 1, erect” - Griswold et al. 1999; Griswold et al. 2005), and as a biomechanical predictor of morphological evolution (Moya-Laraño et al. 2008). We classified the web of *Hypochilus* as aerial because the frame lines of the web usually extend very far in the air (Shear 1969; Fergusson 1972), and the spider spins hanging from the web (personal observation of MJR on *Hypochilus kastoni*, voucher in MACN-Ar MJR-2138).

Ampullate silk and the strong piriform anchorages are essential to build thin strong cables that spiders manipulate precisely and can hang from (Foelix 1970; Wolff et al. 2019). All other silken structures (silken tubes, silk-lined burrows, sensing lines, funnel webs to walk on top, etc.) were scored as ‘substrate webs’. Because many spiders do not make a foraging web, we explored different scorings schemes, with webless spiders as NA or as a third state = ‘webless’.

*Character Mapping.* For the reconstruction of ancestral states of the respiratory system, we tested whether the character is better correlated with an ultrametric tree or with a phylogram with branch lengths derived from the sequence dataset (see Litsios et al. 2012a). Since we used a wide sample of genetic markers, it is plausible that phenotypic traits are correlated with molecular branch lengths via common processes (e.g., longer generation times result in shorter branches, etc.; see references in Litsios et al. 2012a). Litsios et al. (2012a) proposed to compare the phylogenetic signal of the phenotypic traits on an ultrametric tree (‘chronogram’) versus a phylogram, and then choose the tree with the higher phylogenetic signal (see Litsios et al. 2012b; Cusimano and Renner 2014; McCann et al. 2016). As our trait is discrete, we measured phylogenetic signal with the delta parameter for categorical traits (Borges et al. 2019) implemented as in <https://github.com/mrborges23/delta_statistic>. To produce an ultrametric tree, we used a penalized likelihood approach (Kim and Sanderon 2008; Paradis 2013) implemented in the *chronos* function of the R package *ape* (Paradis and Schliep 2018). We started from the preferred Bayesian tree (**Fig. 2**) and used a set of 13 calibration points from the fossil literature to produce a time-calibrated ultrametric tree which was used for comparison with the phylogram. Fossils used as calibrations are *Rosamygale grauvogeli* Selden & Gall, *Eoplectreurys gertschi* Selden & Huang, *Montsecarachne amicorum* Selden, *Burmorchestina acuminata* Wunderlich, *Priscalecrercera paucispinae* Wunderlich, *Seppo koponeni* Selden & Dunlop, *Patarchaea muralis* Selden et al., "Linyphiinae” indet., *Zamilia aculeopectens* Wunderlich, *Almolinus ligula* Wunderlich, *Dasumiana emicans* Wunderlich, *Scytodes weitschati* Wunderlich, and *Furcembolus armatura* Wunderlich following the placements and phylogenetic justifications of Magalhaes et al. (2020).

Ancestral state estimates were carried out using the ace function in the R package *ape* (Paradis and Schliep 2018). Optimizations were carried out using distinct models of character evolution: equal rates (ER; a single global rate of transition between states), symmetric (SYM; a symmetric forward-and-reverse rate for each pair of states), and all rates different (ARD; a different transition rate estimated for all possible transformations) (**Tables S4, S5**). To test a scenario where book lung morphology is forced to be ancestral, we built a custom “irreversible” model (IRV) which disallowed (1) transitions from any state to book lungs and (2) transitions from absence to other states (**Table S5**). The fit of the different reconstructions was compared using the Akaike information criterion (AIC). The number of parameters for all models for reconstructions of the respiratory system are detailed in **Table S5**. Character codings and models for reconstructions of other character systems (webs, silk glands and tracheae passing to the prosoma) are more straightforward and follow the outlines above.

*Correlated evolution.* To test the correlation between traits we used the threshold model of quantitative genetics as modified by Felsenstein (2012), implemented in the threshBayes function of the R package *phytools* (Revell 2014). We also used Pagel’s (1994) test for correlated evolution between discrete characters, implemented in the *correl* package of Mesquite (Midford and Maddison 2006).

**Results**

We successfully gathered original UCE data for 53 specimens**,** with additional data for 43 specimens taken from previous studies (Starrett et al. 2017, Hedin et al. 2018; Wood et al. 2018; Hedin et al. 2019**; Table S1**). We recovered multiple hundreds of loci for most taxa, except for older museum specimens of *Ectatosticta* (**Table S1)**; we note however that the phylogenetic placement of *Ectatosticta* is always as expected (see Results below).

*Phylogenomic analyses*

Theunpartitioned ML analysis resulted in a tree with high bootstrap support for most branches (**Fig. S2**). The selected model was GTR+F+R10 (weight BIC = weight AIC = 1). A simpler GTR+I+G model produced an identical tree, with only branch lengths differing slightly (not shown). This indicates that the analysis is robust to subtle changes in already complex models.

An analysis with ModelFinder detected five partitions (partition 1, 284 loci, model, GTR+F+I+G4; partition 2, 186 loci, GTR+F+I+G4; partition 3, 39 loci, GTR+F+I+G4; partition 4, 24 loci, TVM+F+I+G4; partition 5, 1 locus, TPM2u+F+I+G4). The partitioned ML analysis (**Fig. S3**) produced the same tree as for the unpartitioned dataset, indicating that the analysis is robust to data partitioning.

The TNT parsimony analysis produced a tree similar to the ML analyses (**Fig. S4**), but with lower support on several branches. ML versus parsimony topological differences (only groups of bootstrap / jackknife >90 are mentioned) are in the placement of *Trogloraptor* + Telemidae together with the psilodercid *Leclercera*, and the placement of Pholcidae (*Physocyclus, Pholcophora*). The former difference is discussed below; the latter is of no consequence for character evolution analyses, since pholcids are still placed in a clade with homogeneous respiratory systems. In the above ML and TNT analyses, we obtained a group of Trogloraptoridae + Telemidae, which is at odds with morphological and phylotranscriptomic evidence, which instead suggest that *Trogloraptor* is a close relative of Dysderoidea and Caponiidae (Griswold et al. 2012; Michalik et al. 2019).

The analysis using the multispecies coalescent model implemented in SVDquartets produced a tree essentially equivalent to the ML tree, but with *Trogloraptor* sister to Caponiidae (**Fig. S5**). Since *Trogloraptor* is on a very long branch in the ML trees (**Figs. S2, S3**), we considered the possibility that this result may be an artifact of long-branch attraction and/or GC bias.

Analysis of GC content by taxon (**Fig. S6, S7**) shows that *Trogloraptor* sequences have higher GC content than all other taxa. We focused our explorations of possible long-branch attraction on the grouping of *Trogloraptor* + Telemidae, and also of Filistatidae + Hypochilidae, which has little morphological support (e.g., see Griswold et al. 2005). Removing Telemidae from the concatenated alignment and reanalyzing the data resulted in *Trogloraptor* sister to caponiids, as obtained in Michalik et al. (2019) (**Fig. S8a**). Removing instead *Trogloraptor* from the concatenated alignment leaves telemids in the same position as in the analyses of the complete dataset (**Fig. S8b**). We interpret the above as indicating that *Trogloraptor* is attracted by the long branch of telemids, but the placement of telemids is robust. Removing Filistatidae (in addition to Telemidae) leaves Hypochilidae unaffected as sister to Synspermiata, suggesting that the placement of Hypochilidae is not due to long branch attraction with filistatids (**Fig. S8c**).

We then explored the effect of variation in GC content per-UCE locus, following some authors that recommend removing loci with high GC content (e.g., Romigier et al. 2016). The unpartitioned ML phylogenetic trees resulting from the GC-rich (**Fig. S9a**) and AT-rich loci (**Fig. S9b**) are very similar to each other (83 shared branches, out of 94), and to topologies from the full dataset, including the suspicious placement of *Trogloraptor*. Comparable results were obtained with parsimony analyses (not shown). Our conclusion is that there is no per-locus UCE bias, because GC-rich loci and AT-rich loci resulted in the same suspicious placement of *Trogloraptor* with telemids.

Because *Trogloraptor* is very rich in GC content, we produced a modified dataset with Gs and Cs replaced with “?” for *Trogloraptor*. The resulting ML tree (**Fig. S10**) has *Trogloraptor* sister to Caponiidae (BP = 100), as obtained in previous phylotranscriptomic analyses and is congruent with morphological data, while the rest of the tree is unaffected. The parsimony analysis of this dataset also retrieves *Trogloraptor* sister to Caponiidae with the rest of the tree as in the original dataset (not shown). We thus select this modified dataset with Gs and Cs replaced with “?” for *Trogloraptor* as the basis for Bayesian analyses described below, as well as for character evolution analyses (an exploratory analysis with *Trogloraptor* sequences recoded as RY produced the same tree, but with BP = 86 for *Trogloraptor* + Caponiidae, instead of BS = 100).

The Bayesian inference tree (**Figs. 2, S11**) is very similar to the (GC=? in *Trogloraptor*) ML tree (**Fig. S10)**. To speed up analyses and because previous analyses produced the same results under a variety of models and partitions, we used the GTR + gamma model for the entire unpartitioned dataset. The Bayesian tree only differs from the ML tree of **Figure S10** in the position of two groups. (1) The Segestriidae comes from a basal split of Dysderoidea (switching positions with Oonopidae), which agrees with previous sequence analyses (Wheeler et al. 2017; Chousou-Polydouri et al. 2019) and with morphological analyses (Ramírez 2000). This difference is of no consequence for character evolution analyses, because segestriids and oonopids (as well as all dysderoids) are homogeneous in respiratory system morphology. And (2), *Huttonia* appears as sister to Palpimanidae (although with low support, p=0.92) instead of Mecysmaucheniidae, as found in a previous UCE analysis (Wood et al. 2018). Again, this involves a rearrangement of clades homogeneous in respiratory system morphology. Because we focused our study particularly in the clade Synspermiata, where segestriids belong, we choose the Bayesian tree as our reference tree. The gene- and site- concordance factors calculated over the Bayesian tree are in general low, especially for deep branches (average 25% and 43%, respectively; **Fig. S12**).

A tree summarizing which groups are monophyletic across all analyses is presented in **Figure S13**. Note that all groups with transitions in respiratory system morphology are recovered as monophyletic in all analyses.

*Character Evolution Analyses.*

All R scripts for the analyses described below have been deposited on **Dryad**. The multistate character describing the PRS had a much higher phylogenetic signal on the phylogram than on the ultrametric tree (Delta phylogram = 20.04766, Delta ultrametric = 2.194717). We thus used the phylogram for subsequent analyses of character evolution. The comparison of evolutionary models using the AIC is summarized in **Table 1** for the preferred scoring schemes, and in **Table S4** for alternative scoring schemes that we explored to evaluate the robustness of results. The irreversible model (IRV) was selected by AIC as the preferred model for reconstructing ancestral states of the PRS under multistate coding using the Bayesian tree (AIC weight 0.93, **Table 1**), or all trees of a sample of 1000 trees randomly drawn from the post-burnin posterior sample of the Bayesian analysis (**Figs. S15c, d**). The IRV model was also favored by AIC for the alternative coding schemes (**Table S4**).

PRS character transformations inferred from the alternative evolutionary models are compared in **Figures S14 and S15**. The irreversible model (IRV) infers that book lungs are ancestral and transformed to tracheal structures multiple times independently. The equal rates model (ER) implies that book lungs are lost in Araneomorphae and later re-gained twice, in Hypochilidae and Gradungulidae + *Hickmania*. The symmetric model (SYM) is similar to ER but with less ambiguity. The all rates different model (ARD) implies a single re-gain of book lungs, in Gradungulidae + *Hickmania*.

The multistate coding (‘PRS absent’ scored as a separate state) and the multistate NA coding (‘PRS absent’ scored as ‘NA’ = ‘not applicable’ = ‘missing’) produced the same ancestral state inferences (**Figs. S14 a** and **b**, respectively). An alternative strategy of pruning the terminals without PRS also produced the same inference (**Fig. S16**). The simplified binary coding of book lungs vs. tracheae also produced congruent results, with book lungs ancestral and multiple subsequent losses (**Fig. S17**). Because all results from alternative coding schemes agree in the main results, and because the multistate coding represents the different configurations of PRS and the losses of the PRS in the same graph, we used that scoring to summarize our results in **Figure 3**, but note that the results are robust to alternative coding schemes.

The AIC favored model (IRV) implies 10 steps, while the alternative models imply 9 steps, at the expense of postulating one (ARD) or two (ER, SYM) re-acquisitions of book lungs (**Fig. S15**). Besides the statistical support for the IRV model, we believe that the inference of multiple losses of book lungs is the more plausible scenario. All spider book lungs have essentially the same morphology, and these are complex structures, thus we argue that convergence is unlikely. In contrast, tracheal systems are simpler, and are extremely diverse in location, size, extent and branching pattern. This is partly reflected in our coding scheme, as we consider the conditions in Filistatidae and Austrochilinae as non-homologous to other tracheae, based on ontogenetic studies. All other tracheal systems were conservatively coded in a single state, “tubular tracheae”, so we could test their unique/separate origin, although that state encompasses diverse morphologies. Convergence in tracheal systems would not be surprising, since they are already diverse in morphology and simple in structure. Moreover, it is well accepted that tracheae have been acquired independently in several arthropod lineages (Dunlop et al. 2013). Thus, the morphological evidence also makes a strong argument towards choosing our preferred optimization as the most likely candidate reconstruction of the evolution of the respiratory system in true spiders.

The selection of models for the extent of the respiratory system (limited to the opisthosoma vs. tracheae passing to the prosoma) favored the ER model (**Table S4**). The ER and ARD models imply five independent origins of extensive tracheal systems reaching the prosoma, all derived from systems limited to the opisthosoma (**Figs. 3, S18**). The IRV model inferred one additional convergence event (**Fig. S19b**).

The selection of models for the presence of the ampullate + piriform gland system favored the ARD model (**Table S4**), which inferred the origin of the system in Araneomorphae, with a single loss in *Hexophthalma* (**Fig. S20**). The ER model produced the same result.

Selection of models for the multistate coding scheme of webs (‘substrate’, ‘aerial’, ‘webless’) favored the ER model (**Table 1**); the same model was selected when using a binary coding (‘substrate’, ‘aerial’, ‘NA’, **Table S4**). The character mapping places the origin of the aerial web in Araneomorphae and multiple losses or transformations to substrate webs (**Fig. S21**); all models and coding schemes agree with this result, only differing in the number of transformations or losses in derived clades (not shown).

We detected a significant correlation between the ampullate + piriform system and tracheae using Pagel’s test (*p* = 0.03), but not between aerial webs and tracheae (*p* = 0.42). Similar results were obtained using the threshold model between the ampullate + piriform system and tracheae (highest posterior density of correlation 0.050 to 0.797, mean 0.410, p = 0.925, effective sample size 319) and between webs and tracheae (highest posterior density of correlation -0.440 to 0.386, mean -0.064, p = 0.95, effective sample size 267). Given that the ampullate + piriform system appeared only once in evolution, the correlation found with tracheae may be the result of a “pseudoreplication” effect (Maddison and FitzJohn 2014). However, although multiple origins would make a stronger case, the loss of the ampullate + piriform system, corresponding to a loss of the tracheae in *Hexophthalma*, is a second relevant event. Our interpretation of this correlation is that tracheae may be functionally associated with the ampullate + piriform system, but not specifically with the aerial webs. This silk system was a key evolutionary innovation that allowed the evolution of aerial webs, but while webs were lost or modified multiple times in the phylogeny, the ampullate + piriform system was predominantly retained in araneomorphs and used for draglines and other structures.

**Literature Cited**

Aberer A.J., Kobert K., Stamatakis A. 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. *Mol. Biol. Evol.* 31: 2553-2556.

Álvarez-Padilla F., Hormiga G. 2007. A protocol for digesting internal soft tissues and mounting spiders for scanning electron microscopy. *J. Arachnology* 35: 538-543.

Borges R., Machado J.P., Gomes C., Rocha A.P., Antunes A. 2019. Measuring phylogenetic signal between categorical traits and phylogenies. *Bioinformatics* 35: 1862-1869.

Borowiec M.L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4, e1660.

Bossert S., Murray E.A., Blaimer B.B., Danforth B.N. 2017. The impact of GC bias on phylogenetic accuracy using targeted enrichment phylogenomic data. *Mol. Phyl. Evol.* 111: 149-157.

Bristowe W.S. 1930. XXXIII.—Notes on the biology of spiders.—I. The evolution of spiders' snares. *Annals and Magazine of Natural History* 6(33): 334-342.

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

Chernomor O., von Haeseler A., Minh B.Q.2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst. Biol.* 65: 997-1008.

Chifman J., Kubatko L., 2014. Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30(23): 3317-3324.

Chousou-Polydouri N., Carmichael A., Szűts T., Saucedo A., Gillespie R., Griswold C., Wood H.M. 2019. Giant Goblins above the waves at the southern end of the world: The biogeography of the spider family Orsolobidae (Araneae, Dysderoidea). *J. Biogeography* 46(2): 332-342.

Cusimano N., Renner S.S. 2014. Ultrametric trees or phylograms for ancestral state reconstruction: Does it matter? *Taxon* 63: 721-726.

Derkarabetian S., Starrett J., Tsurusaki N., Ubick D., Castillo S., Hedin M. 2018. A stable phylogenomic classification of Travunioidea (Arachnida, Opiliones, Laniatores) based on sequence capture of ultraconserved elements. *ZooKeys* 760: 1–36.

Dunlop J.A., Scholtz G., Selden P.A. 2013. Water-to-land transitions. *In* Arthropod Biology and Evolution (pp. 417-439). Springer, Berlin, Heidelberg.

Faircloth B.C. 2013. illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. <http://dx.doi.org/10.6079/J9ILL>

Faircloth B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci**.** *Bioinformatics* 32: 786–788.

Faircloth B.C. 2017. Identifying conserved genomic elements and designing universal probe sets to enrich them**.** *Meth. Ecol. Evol.* 8: 1103–1112.

Fergusson I.C. 1972. Natural history of the spider *Hypochilus thorelli* Marx (Hypochilidae). *Psyche* 79: 179-199.

Felsenstein J. 2012. A comparative method for both discrete and continuous characters using the threshold model. *American Naturalist* 179(2): 145-156.

Foelix R.F. 1970. Structure and function of tarsal sensilla in the spider *Araneus diadematus*. *J. Experimental Zoology* 175: 99-123.

Goloboff P.A., Farris J.S., Nixon K.C. 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24: 774-786.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotech.* 29: 644.

Griswold C.E., Coddington J.A., Platnick N.I., Forster R.R. 1999. Towards a phylogeny of entelegyne spiders (Araneae, Araneomorphae, Entelegynae). *J. Arachnology* 27: 53-63.

Griswold C.E., Ramírez M.J., Coddington J.A., Platnick N.I. 2005. Atlas of phylogenetic data for entelegyne spiders (Araneae: Araneomorphae: Entelegynae) with comments on their phylogeny. *Proc. California Acad. Sciences* 56 (Suppl. II): 1-324.

Griswold C.E., Audisio T., Ledford J.M. 2012. An extraordinary new family of spiders from caves in the Pacific Northwest (Araneae, Trogloraptoridae, new family). *ZooKeys* 215: 77-102.

Hedin M., Derkarabetian S., Blair J., Paquin P. 2018. Sequence capture phylogenomics of eyeless *Cicurina* spiders from Texas caves, with emphasis on US federally-endangered species from Bexar County (Araneae, Hahniidae). *ZooKeys* 769: 49.

Hedin M., Derkarabetian S., Alfaro A., Ramírez M.J., Bond J.E. 2019. Phylogenomic analysis and revised classification of atypoid mygalomorph spiders (Araneae, Mygalomorphae), with notes on arachnid ultraconserved element loci. *PeerJ* 7:e6864.

Hoang D.T., Chernomor O., von Haeseler A., Minh B.Q., Vinh L.S.2018. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35: 518–522.

Kalyaanamoorthy S., Minh B.Q., Wong T.K.F., von Haeseler A., Jermiin L.S. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14: 587-589.

Kaston B.J. 1964. The evolution of spider webs. *American Zoologist* 4: 191-207.

Katoh D., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30: 772–780.

Kim J., Sanderson M.J. 2008. Penalized likelihood phylogenetic inference: bridging the parsimony-likelihood gap. *Syst. Biol.* 57: 665–674.

Litsios G., Salamin N. 2012a. Effects of phylogenetic signal on ancestral state reconstruction. *Syst. Biol.* 61: 533-538.

Litsios G., Sims C.A., Wüest R.O., Pearman P.B., Zimmermann N.E., Salamin N. 2012b. Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evolutionary Biology* 12: 212.

Maddison W.P., FitzJohn R.G. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Syst. Biol.* 64: 127-136.

Magalhaes I.L.F., Azevedo G.H.F., Michalik P., Ramírez M.J. 2020. The fossil record of spiders revisited: implications for calibrating trees and evidence for a major faunal turnover since the Mesozoic. *Biol Rev* 95: 184-217. doi:10.1111/brv.12559

McCann J., Schneeweiss G.M., Stuessy T.F., Villasenor J.L., Weiss-Schneeweiss H. 2016. The impact of reconstruction methods, phylogenetic uncertainty and branch lengths on inference of chromosome number evolution in American daisies (Melampodium, Asteraceae). *PloS one* 11(9): e0162299.

Michalik P., Kallal R., Dederichs T.M., Labarque F.M., Hormiga G., Giribet G., Ramírez M.J. 2019. Phylogenomics and genital morphology of cave raptor spiders (Araneae, Trogloraptoridae) reveal an independent origin of a flow‐through female genital system. *J. Zool. Syst. Evol. Res.* 57: 737-747.

Midford P., Maddison W. 2006. Correl package for Mesquite, version 0.1. Website http://mesquiteproject. org.

Moya-Laraño J., Vinković D., De Mas E., Corcobado G., Moreno E. 2008. Morphological evolution of spiders predicted by pendulum mechanics. *PLoS ONE* 3(3): e1841. <https://doi.org/10.1371/journal.pone.0001841>

Nguyen L-T., Schmidt H.A., von Haeseler A., Minh B.Q. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 32: 268-274.

Paradis E. 2013. Molecular dating of phylogenies by likelihood methods: a comparison of models and a new information criterion. *Mol. Phyl. Evol.* 67: 436–444.

Paradis E., Schliep K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526-528.

Purcell W.F. 1909. Development and origin of respiratory organs in Araneae. *Quart. J. Microsc. Sc. (N.S.)* 54:1–110.

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard, M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67: 901-904.

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

Ramírez M.J. 2014. The morphology and phylogeny of dionychan spiders (Araneae: Araneomorphae). *Bull. Amer. Mus. Nat. Hist.* 390: 1–394.

Revell L.J. 2014. Ancestral character estimation under the threshold model from quantitative genetics. *Evolution* 68: 743-759.

Romiguier J., Cameron S.A., Woodard S.H., Fischman B.J., Keller L., Praz C.J. 2016. Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Mol. Biol. Evol.* 33: 670-678.

Shear W.A. 1969. Observations on the predatory behavior of the spider *Hypochilus gertschi* Hoffman (Hypochilidae). *Psyche* 76: 407-417.

Shear W.A. 1986. The evolution of web-building behavior in spiders: a third generation of hypotheses. *in* Spiders: Webs, Behavior and Evolution. Stanford: Stanford Univ. Press.

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

Starrett J., Derkarabetian S., Hedin M., Bryson Jr. R.W., McCormack J.E., Faircloth B.C. 2017. High phylogenetic utility of an Ultraconserved element probe set designed for Arachnida. *Mol. Ecol. Res.* 17: 812-823.

Swofford D.L. 2002. PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Talavera G., Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56: 564–577.

Wheeler W.C., Coddington J.A., Crowley L.M., Dimitrov D., Goloboff P.A., Griswold C.E., Hormiga G., Prendini L., Ramírez M.J., Sierwald P., Almeida‐Silva, L., Alvarez‐Padilla F., Arnedo M.A., Benavides Silva L.R., Benjamin S.P., Bond J.E., Grismado C.J., Hasan E., Hedin M., Izquierdo M.A., Labarque F.M., Ledford J., Lopardo L., Maddison W.P., Miller J.A., Piacentini L.N., Platnick N.I., Polotow D., Silva‐Dávila D., Scharff N., Szűts T., Ubick D., Vink C.J., Wood H.M., Zhang J. 2017. The spider tree of life: phylogeny of Araneae based on target‐gene analyses from an extensive taxon sampling. *Cladistics* 33: pp.574-616.

Wolff J.O., Paterno G.B., Liprandi D., Ramírez M.J., Bosia F., van der Meijden A., Michalik P., Smith H.M., Jones B., Ravelo A.M., Pugno N. 2019. Evolution of aerial spider webs coincided with repeated structural optimization of silk anchorages. *Evolution* 73: 2122–2134.

Wood H.M., González V.L., Lloyd M., Coddington J., Scharff N. 2018. Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea). *Mol. Phyl. Evol.* 127: 907-918.

Zerbino D.R., BirneyE. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Gen. Res.* 18: 821–829.