
















Comprehensive Species Sampling and Sophisticated Algorithmic Approaches Refute the Monophyly of Arachnida

Jesús A. Ballesteros ^{*,†,1} Carlos E. Santibáñez-López ^{1,2} Caitlin M. Baker ^{1,3} Ligia R. Benavides ³
Tauana J. Cunha ^{3,4} Guilherme Gainett ¹ Andrew Z. Ontano ¹ Emily V.W. Setton ¹
Claudia P. Arango ⁵ Efrat Gavish-Regev ⁶ Mark S. Harvey ^{7,8} Ward C. Wheeler ⁹
Gustavo Hormiga ¹⁰ Gonzalo Giribet ³ and Prashant P. Sharma ^{*,1}

¹Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, USA

²Department of Biology, Western Connecticut State University, Danbury, CT, USA

³Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA

⁴Smithsonian Tropical Research Institute, Panama City, Panama

⁵Office for Research, Griffith University, Nathan, QLD, Australia

⁶National Natural History Collections, The Hebrew University of Jerusalem, Jerusalem, Israel

⁷Collections & Research, Western Australian Museum, Welshpool, WA, Australia

⁸School of Biological Sciences, University of Western Australia, Crawley, WA, Australia

⁹Division of Invertebrate Zoology, American Museum of Natural History, New York, NY, USA

¹⁰Department of Biological Sciences, George Washington University, Washington, DC, USA

[†]Present address: Department of Biology, Kean University, Union, NJ, USA

***Corresponding authors:** E-mails: jeballes@kean.edu; prashant.sharma@wisc.edu.

Associate editor: Emma Teeling

Abstract

Deciphering the evolutionary relationships of Chelicerata (arachnids, horseshoe crabs, and allied taxa) has proven notoriously difficult, due to their ancient rapid radiation and the incidence of elevated evolutionary rates in several lineages. Although conflicting hypotheses prevail in morphological and molecular data sets alike, the monophyly of Arachnida is nearly universally accepted, despite historical lack of support in molecular data sets. Some phylotranscriptomic analyses have recovered arachnid monophyly, but these did not sample all living orders, whereas analyses including all orders have failed to recover Arachnida. To understand this conflict, we assembled a data set of 506 high-quality genomes and transcriptomes, sampling all living orders of Chelicerata with high occupancy and rigorous approaches to orthology inference. Our analyses consistently recovered the nested placement of horseshoe crabs within a paraphyletic Arachnida. This result was insensitive to variation in evolutionary rates of genes, complexity of the substitution models, and alternative algorithmic approaches to species tree inference. Investigation of sources of systematic bias showed that genes and sites that recover arachnid monophyly are enriched in noise and exhibit low information content. To test the impact of morphological data, we generated a 514-taxon morphological data matrix of extant and fossil Chelicerata, analyzed in tandem with the molecular matrix. Combined analyses recovered the clade Merostomata (the marine orders Xiphosura, Eurypterida, and Chasmataspidea), but merostomates appeared nested within Arachnida. Our results suggest that morphological convergence resulting from adaptations to life in terrestrial habitats has driven the historical perception of arachnid monophyly, paralleling the history of numerous other invertebrate terrestrial groups.

Key words: Chelicerata, orthologs, supermatrix, phylogenomics, total evidence.

Introduction

Chelicerates are a diverse group of arthropods that have played a major role as predators in ancient and recent ecosystems. United by the eponymous pincer-like appendages (the chelicerae/chelifores), chelicerates comprise the sister group to the remaining Arthropoda. The most familiar chelicerate orders are members of Arachnida, an assemblage of 12

terrestrial orders (e.g., spiders, scorpions, mites). Chelicerates also include two wholly marine clades—the sea spiders (Pycnogonida) and the horseshoe crabs (Xiphosura)—as well as considerable diversity of derived aquatic lineages within mites (Wheeler and Hayashi 1998; Giribet et al. 2001; Dabert et al. 2016; Giribet and Edgecombe 2019). The fossil record of chelicerates attests to a broader aquatic diversity

that includes freshwater horseshoe crabs, sea scorpions (Eurypterida), and chasmataspids (Dunlop et al. 2007; Dunlop 2010; Lamsdell 2016).

Whereas most higher-level phylogenetic relationships of arthropods have been resolved by the advent of phylogenomic approaches (e.g., Giribet and Edgecombe 2019; Edgecombe 2020), the internal phylogeny of chelicerates has remained elusive. The traditional paradigm of chelicerate evolution postulates a single colonization of land by the common ancestor of a monophyletic Arachnida. In this scenario, extinct lineages such as the chasmataspids and sea scorpions are thought to represent stepping-stones between horseshoe crabs and the origin of arachnids. Phylogenomic studies have recovered weak support for this scenario, with a suite of analyses supporting a nested placement of Xiphosura as derived arachnids (Sharma et al. 2014; Ballesteros and Sharma 2019; Ballesteros et al. 2019; Ontano et al. 2021), a result also recovered in many Sanger-based molecular analyses (Wheeler and Hayashi 1998; Colgan et al. 1998; Edgecombe et al. 2000; Giribet et al. 2001, 2002; Mallatt et al. 2004; Mallatt and Giribet 2006; Masta et al. 2009; Pepato et al. 2010; Arabi et al. 2012; Noah et al. 2020).

A handful of phylogenomic matrices has recovered arachnid monophyly, attributing this result to 1) the use of slowly evolving (i.e., less saturated) genes that are less prone to long-branch attraction (LBA) artifacts or 2) expanded taxonomic sampling (Sharma et al. 2014; Lozano-Fernández et al. 2019; Howard et al. 2020). However, the matrices of these works have been shown to be highly sensitive to model choice and algorithmic approach, and lacked representation of all extant chelicerate orders. Upon addition of libraries representing the missing arachnid orders to these same data sets, support for arachnid monophyly collapsed (Ballesteros et al. 2019; Ontano et al. 2021).

Beyond arachnid monophyly, internal relationships within Chelicerata are unstable across phylogenomic analyses, which is in part attributable to the incidence of multiple fast-evolving lineages that incur LBA artifacts, such as Acariformes, Parasitiformes (treated as single orders in this study), and Pseudoscorpiones. Well-resolved parts of the chelicerate phylogeny include the reciprocal monophyly of Pycnogonida and Euchelicerata (the remaining chelicerate orders), the monophyly and internal relationships of Tetrapulmonata (spiders and three other orders that plesiomorphically bear four book lungs), and the monophyly of each chelicerate order. More recently, phylogenomic analyses emphasizing dense taxonomic sampling, together with rare genomic changes, have supported the clade Panscorpiones (Scorpiones + Pseudoscorpiones), in turn sister group to Tetrapulmonata (forming the clade Arachnopulmonata) (Sharma et al. 2014; Schwager et al. 2017; Gainett and Sharma 2020; Nolan et al. 2020; Gainett et al. 2021; Harper et al. 2021; Ontano et al. 2021) (fig. 1). Nevertheless, data quality and quantity remain limited for several groups in phylotranscriptomic data sets, which may underlie the basal instability within Euchelicerata.

Toward a comprehensive chelicerate phylogeny that can inform the debate on arachnid monophyly, we assembled a

506-taxon phylogenomic data set representing the major lineages of all extant chelicerate orders and densely representing species-rich groups, such as Araneae, Scorpiones, Pseudoscorpiones, and Opiliones. Our analyses examined sophisticated strategies to mitigate LBA, such as subsampling loci to minimize saturation, the use of infinite mixture site-heterogeneous models (CAT-GTR), and recently proposed recoding strategies in tandem with site-heterogeneous models applied to partitioned model analyses.

A common feature of phylogenomic studies is the omission of morphological data in an analytical framework, a practice that has been argued to be detrimental to phylogenetic reconstruction (Mongiardino Koch and Parry 2020). Combined analyses of phylogenomics and morphology have been proposed as a means to improve resolution and evaluate congruence among data classes (Mongiardino Koch and Thompson 2021; Neumann et al. 2021). Although morphological data sets focusing on relationships among fossil taxa typically recover arachnid monophyly (Lamsdell 2016; Wolfe 2017; but see Aria and Caron 2019; reviewed by Nolan et al. 2020; Sharma et al. 2021), most of these matrices have historically suffered from minimal sampling of extant arachnid diversity, exhibit marked character conflict, and fail to recover the few relationships that are consistently supported by molecular phylogenies and genomics (e.g., Tetrapulmonata [Wolfe 2017]; Arachnopulmonata [Lamsdell 2016; Wolfe 2017; Aria and Caron 2019]; Euchelicerata [Garwood and Dunlop 2014]). Therefore, toward assessing the impact of fossil taxa and morphological characters on phylogenomic analyses, we assembled a 514-taxon morphological data set for Chelicerata to complement the phylogenomic data set. The morphological data set included extinct taxa (e.g., Chasmataspida, Eurypterida, Haptopoda, Phalangiotarbita, Synziphosurina, Trigonotarbita, and Uruaraneida) as well as key fossils of extant orders.

Here, we show that analyses of molecular data sets alone, as well as combined analyses of morphology and molecules, consistently recover horseshoe crabs as nested within Arachnida. Interrogation of phylogenetic signal across loci showed that genes and sites supporting arachnid monophyly are more prone to systematic error than the remaining loci, suggesting that arachnid monophyly in molecular phylogenies reflects an analytical artifact.

Results

Partitioned Analyses of Phylogenomic Data Sets

We compiled 506 high-quality transcriptomes or genomes (>95% of libraries generated by us; 75 transcriptomes newly sequenced for this study focused on improving representation of scorpions, palpigrades, and opilioacariforms), sampling 24 outgroup and 482 chelicerate taxa (supplementary tables S1 and S2, Supplementary Material online). Phylogenetically informed inference of orthologs leveraged a recent de novo computation of orthologous genes for Chelicerata (3534 loci identified previously by Ballesteros and Sharma 2019) using the Unrooted Phylogenetic

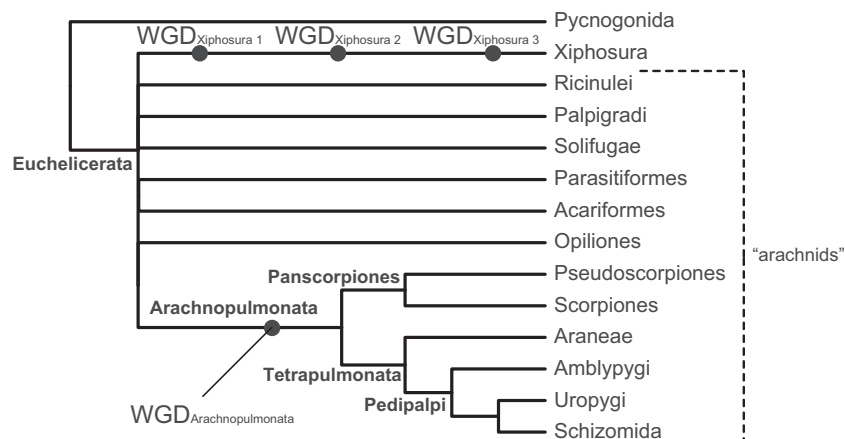


FIG. 1. Higher-level phylogeny of Chelicerata showing well-resolved groups (boldface text adjacent to nodes), based on [Ontano et al. \(2021\)](#). Circles indicate whole-genome duplication events (WGD) subtending specific taxa. Branch lengths are not to scale.

Orthology (UPhO) pipeline ([Ballesteros and Hormiga 2016](#)). As a separate, independent approach to orthology inference, orthologs were drawn from the Benchmarking Universal Single Copy Orthologs loci set for arthropods (BUSCO-Ar) ([Simão et al. 2015](#); [Waterhouse et al. 2018](#)). Initial sets of orthologs were filtered based on maximal taxon decisiveness ([Steel and Sanderson 2010](#)); we retained only loci that had at least one terminal for all the following clades: Araneae, Pedipalpi (=Uropygi + Schizomida + Amblypygi), Scorpiones, Ricinulei, Xiphosura, Solifugae, Opiliones, Palpigradi, Parasitiformes, Acariformes, Pseudoscorpiones, Pycnogonida, Pancrustacea, Myriapoda, and Onychophora. Applying this criterion, we reduced the UPhO ortholog set to 676 loci (Matrix 1) and the BUSCO set to 399 loci (Matrix 2). Thus, every major lineage (i.e., orders or closely related orders [e.g., Pedipalpi; Acariformes; Parasitiformes]) of chelicerates was represented by at least one terminal for every locus, in all analyses. For both matrices, we implemented 1) the site heterogeneous PMSF model for maximum likelihood ([Wang et al. 2019](#)) (LG+C20+F+ Γ_4), 2) traditional partitioned-model maximum likelihood, and 3) gene tree summary (ASTRAL) approaches.

In all six analyses, we recovered the nested placement of Xiphosura within a paraphyletic Arachnida, with support (bootstrap frequency [BS] > 95%; posterior probability > 0.95) and with significance in tests of monophyly ([fig. 2](#); [supplementary table S3, Supplementary Material](#) online). Although relationships of apulmonate arachnid orders varied across topologies, all analyses invariably recovered the monophyly of Tetrapulmonata, Pedipalpi, Euchelicerata, and each chelicerate order. Scorpiones were consistently recovered as the sister group of Tetrapulmonata, whereas Pseudoscorpiones grouped with other long-branched orders (Acariformes and Parasitiformes).

Analyses of Slowly Evolving Matrices

In the case of Pseudoscorpiones, an external and independent phylogenetic data class informs the placement of this long-branched order. Specifically, a shared whole genome

duplication unites the clade Arachnopolmonata, as evidenced by duplications of Hox clusters, systemic paralogy of developmental patterning genes, and enrichment of microRNA families ([Ontano et al. 2021](#)) ([fig. 1](#)). As our analyses of Matrices 1 and 2 did not recover a monophyletic Panscorpiones (with pseudoscorpions clustering with other long-branch orders), we reasoned that these data sets remained exposed to LBA.

Several strategies have been proposed to mitigate LBA in arachnid phylogenetic studies, such as the use of site heterogeneous models, the use of slowly evolving genes, or both (albeit with mixed results across data sets). To mitigate the impact of fast-evolving loci, we generated saturation plots for each locus and isolated a subset of 152 loci with high values for slope (≥ 0.4) and r^2 (≥ 0.95); these loci were concatenated to form Matrix 3 and analyzed using the same approaches as Matrices 1 and 2. Analyses of Matrix 3 with partitioned models, site heterogeneous models, and ASTRAL all recovered the monophyly of Arachnopolmonata (sensu [Ontano et al. 2021](#)) with maximal nodal support ([fig. 2b](#)). Maximum likelihood inference under either partitioned or site heterogeneous models also recovered Panscorpiones (BS = 96% and 90%, respectively). All analyses of Matrix 3 rejected arachnid monophyly with support and with significance in tests of monophyly ([fig. 2d](#)).

Bayesian Inference Analysis with CAT-GTR

Some of the most recalcitrant nodes in the tree of life that are impacted by LBA have been argued to be effectively resolved using analyses under the computationally intensive CAT-GTR infinite mixture model, as implemented in PhyloBayes-mpi ([Lartillot et al. 2013](#)). Examples of such nodes include the placement of Chaetognatha, Xenoturbellida, and Porifera ([Marlétaz et al. 2019](#); [Kapli and Telford 2020](#); but see [Whelan and Halanych 2017](#); [Li et al. 2021](#)). The PhyloBayes approach is notoriously difficult to implement for taxon-rich data sets due to the low probability of convergence, and specifically so under the CAT-GTR model, which requires copious amounts of data for estimation of a large number of rate categories and may be prone to model overfitting ([Li](#)

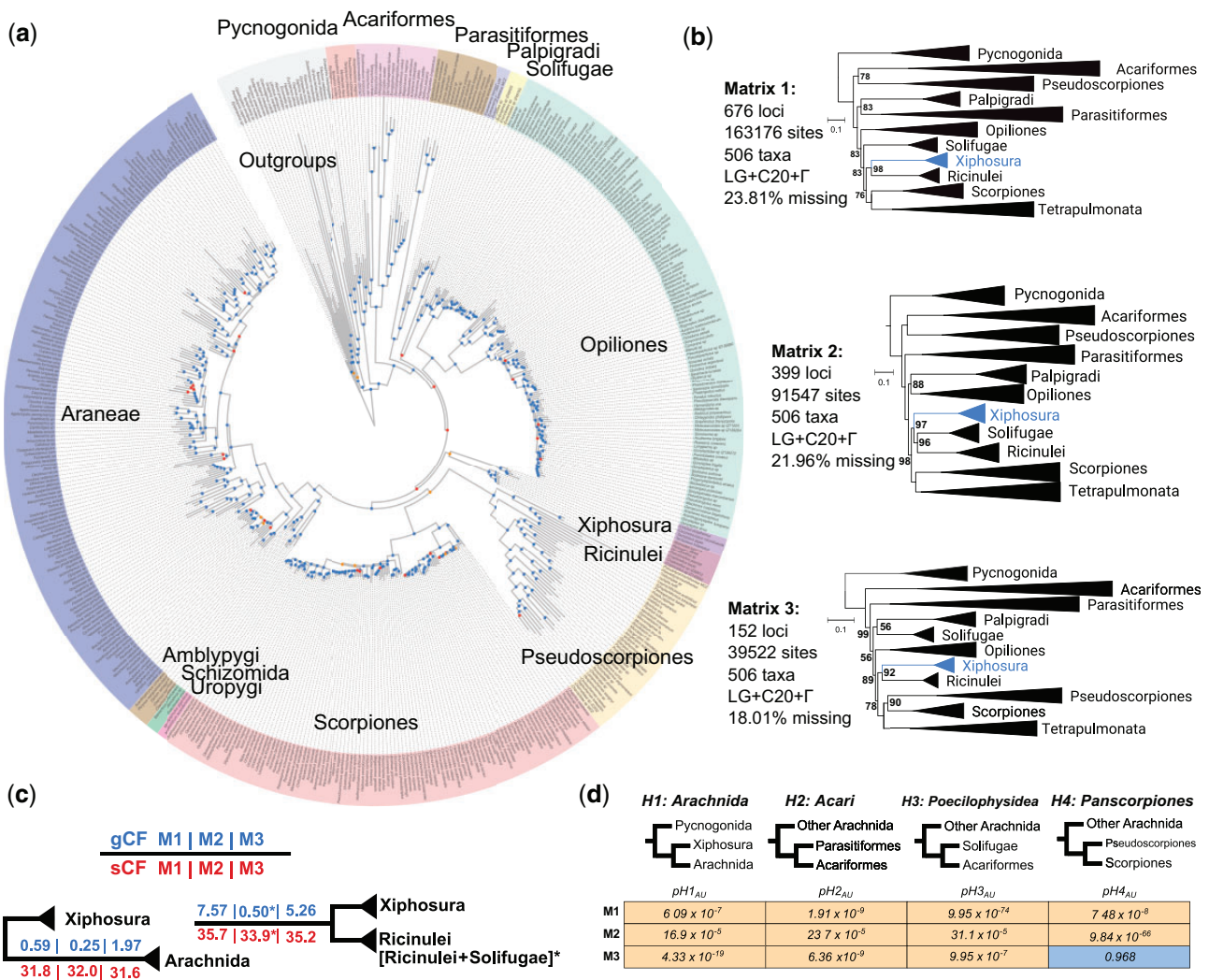


FIG. 2. Comprehensive phylogenomic sampling of all extant chelicerate orders recovers horseshoe crabs as derived arachnids. (a) Phylogenomic relationships of 506 chelicerate data sets based on maximum likelihood analysis of slowly evolving loci (Matrix 3) and site heterogeneous evolutionary models. Colors correspond to orders; note that Acariformes and Parasitiformes are each treated as orders in this study. Dots on nodes indicate high (>95% bootstrap; blue color), medium (90–95% bootstrap; orange color), or low (<90% bootstrap; red color) support. (b) Summary of relationships inferred under site heterogeneous models by three matrices. Numbers on nodes correspond to bootstrap resampling frequencies below 100%; all unlabeled nodes are maximally supported. (c) Gene (gCF) and site (sCF) concordance factors exhibit higher support for the derived placement of Xiphosura under all three 506-taxon matrices. Asterisks indicate tree topologies wherein Xiphosura was recovered as sister group to Ricinulei + Solifugae. (d) Tests of monophyly consistently rejected the monophyly of Arachnida and Acari over the unconstrained topology for Matrices 1–3. Nonsignificant result for Matrix 3 results from the unconstrained recovery of Panscorpiones in this analysis.

et al. 2021). We therefore selected 56 representative terminals from the slow-evolving data set (Matrix 3) such that major taxonomic groups (defined in [supplementary table S4, Supplementary Material](#) online) were each represented by three to five terminals, major basal splits were represented in each lineage, and the selected taxa exhibited the highest possible data completeness. This data set was further filtered with BMGE v 1.12 (Criscuolo and Grigaldo 2010) to remove heteroprecipitous sites (i.e., sites that alter substitution patterns over time), which violate the assumptions of the CAT model (Simion et al. 2017). The resulting matrix (Matrix 4) comprised 14,753 sites. Bayesian inference analysis was run on eight independent chains for >20,000 cycles. To assess the impact of the starting tree on the analysis, two chains (C1 and

C2) used the maximum likelihood tree computed for Matrix 4 as the starting point (which recovered horseshoe crabs in a derived position). Another two chains (C3 and C4) were started on a maximum likelihood tree for Matrix 4 that was constrained to recover arachnid monophyly. Four chains (C5–C8) used random starting trees.

Examination of ESS values and a posteriori tree distribution across all eight chains showed that summary statistics broadly exhibited convergence ([supplementary tables S5 and S6, Supplementary Material](#) online). A high value of the maximum split difference (*maxdiff*) was driven by a soft polytomy at the base of Euchelicerata. We examined estimates both from combined chains and summary topologies resulting from each starting tree type ([supplementary figs. S1 and S2,](#)

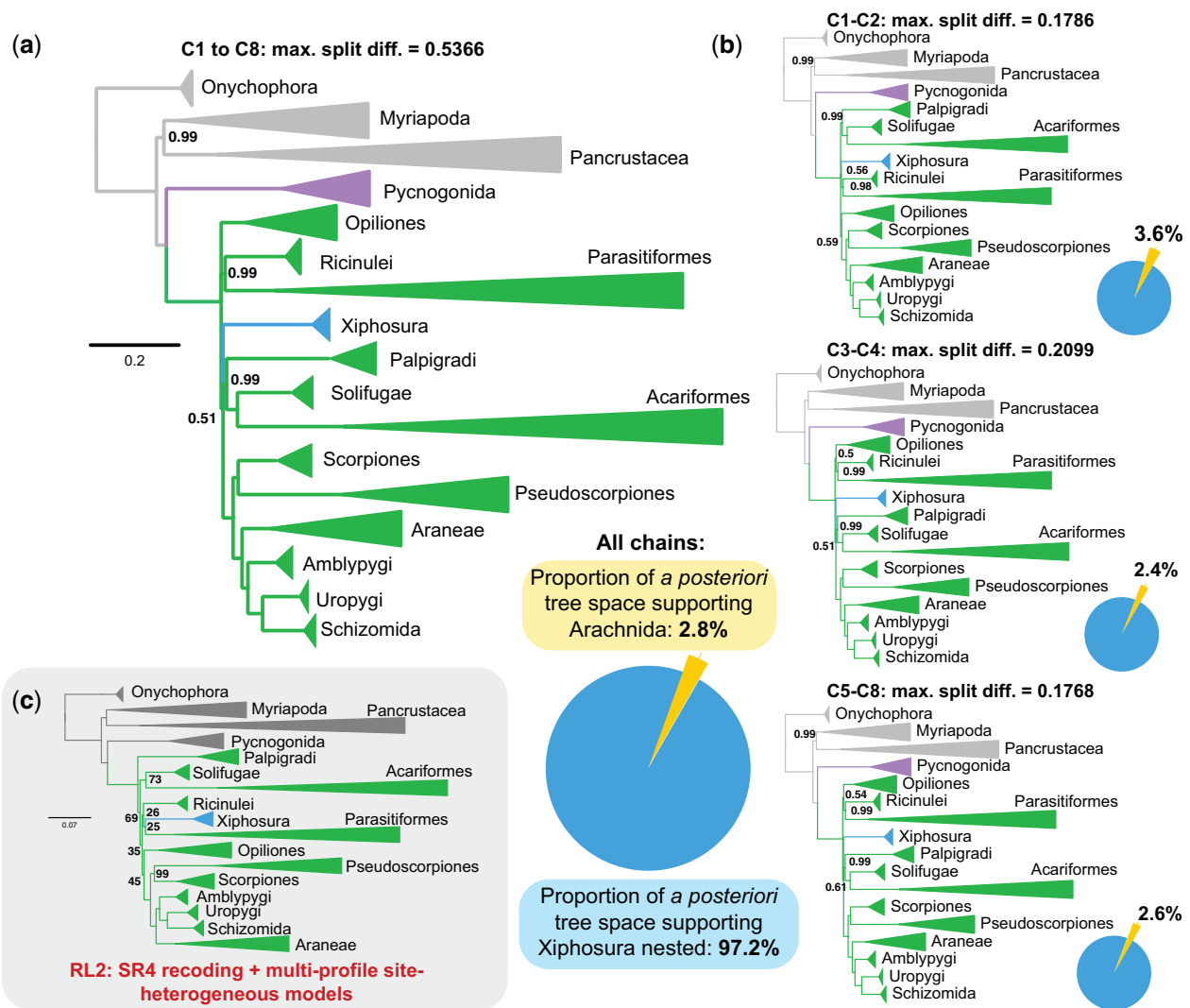


FIG. 3. Site heterogeneous model-based approaches using CAT+GTR+ Γ and SR4 recoding refute the monophyly of Arachnida. (a) Summary tree of eight chains from PhyloBayes-mpi analysis of Matrix 4. Numbers on nodes correspond to posterior probabilities below 1.00; all unlabeled nodes are maximally supported. Lower right: Distribution of support across a posteriori trees for arachnid monophyly (yellow) versus nested placement of Xiphosura (blue). (b) Summary trees from PhyloBayes-mpi analysis separated by starting tree topology. Top: Chains started on maximum likelihood tree topology for Matrix 4 (Xiphosura nested). Middle: Chains started on maximum likelihood tree topology for Matrix 4 with a constraint for arachnid monophyly. Bottom: Chains started on random tree topologies. Nodal support values and pie charts for each summary tree reflect the conventions for (a). (c) Maximum likelihood tree topology based on SR4 recoding and multi-profile tiered site heterogeneous models (RL2 approach). Numbers on nodes correspond to bootstrap resampling frequencies below 100%; all unlabeled nodes are maximally supported.

Supplementary Material online). None of the topologies in the 95% highest posterior density interval of the a posteriori distribution supported the monophyly of Arachnida (PP = 0.02814) (fig. 3a and b). Notably, Bayesian analysis using CAT-GTR rejected the monophyly of Acari (Acariformes + Parasitiformes) in favor of Poecilophysidea (Acariformes + Solifugae; PP = 1.00) and Cephalosomata (Palpigradi + Poecilophysidea; PP = 0.99) (supplementary table S3, Supplementary Material online). Moreover, Acari monophyly was supported in 0% of a posteriori tree space across the eight chains. These results suggest that Acari reflects another LBA artifact. PhyloBayes-mpi was able to recover both the monophyly of Arachnoplumonata (PP = 1.00) and Panscorpiones (PP \geq 0.99), regardless of the starting tree topology.

Partitioned Analysis with Mixture Models and Recoding

A recently proposed method for reconciling divergent results in partitioned versus mixture model studies of recalcitrant nodes makes use of a tiered approach to introduce site-heterogeneous models in tandem with SR4 recoding (RL2, *sensu* Redmond and McLysaght 2021). This approach has been shown to recover consistently the traditional placements of groups such as Porifera in empirical data sets.

Upon applying the RL2 strategy to Matrix 3, we recovered yet another tree topology with a nested placement of Xiphosura, as well as Poecilophysidea (BS = 73%), Panscorpiones (BS = 99%), and Arachnoplumonata (BS = 100%) (fig. 3c). The backbone of Euchelicerata exhibited

negligible support, a result attributable to the loss of information via reducing the peptide alphabet to four states in SR4 recoding. Paralleling this result, previous applications of Dayhoff 6-state recoding to chelicerate data sets have rendered a basal polytomy at the root of Euchelicerata (Lozano-Fernández et al. 2019). These results are consistent with recent critiques of recoding strategies as solutions to saturation and compositional heterogeneity (Hernandez and Ryan 2021; Li et al. 2021).

Tests of Monophyly and Concordance Factors

Tests of monophyly were performed using the approximately unbiased (AU) test (Shimodaira 2002). The different topologies obtained from Matrices 1–3 were constrained to assess support for the monophyly of Arachnida, Acari, Poecilophysidea (Solifugae + Acariformes), and Panscorpiones (Pseudoscorpiones + Scorpiones). AU tests consistently rejected the monophyly of arachnids over the hypothesis of a derived Xiphosura (fig. 2d, supplementary table S3, Supplementary Material online).

Traditional statistical measures of nodal support are prone to inflation in phylogenomic data sets. Gene and site concordance factors (gCF and sCF) have been shown to measure phylogenetic signal irrespective of data set size (Minh et al. 2020). We therefore computed values of gCF and sCF both for unconstrained topologies under Matrices 1–3, as well as their counterparts when constrained to recover the monophyly of Arachnida. gCF and sCF values were consistently lower for Arachnida when compared with the hypothesis of a derived Xiphosura (fig. 2c).

Interrogation of Phylogenetic Signal and Systematic Bias

To examine whether the derived placement of Xiphosura stemmed from a systematic artifact, we explored phylogenetic signal and properties of genes and sites as a function of support for competing tree topologies (Shen et al. 2017). We found that loci favoring arachnid monophyly were consistently in the minority (39–41%) of genes across our data sets, irrespective of orthology criterion (fig. 4a). Proportions of genes supporting arachnid monophyly are comparable to those supporting archaic groupings that have been refuted by phylogenomics and rare genomic changes, such as Dromopoda (=Scorpiones + Opiliones + Solifugae + Pseudoscorpiones; 34–36%) (supplementary fig. S3 and table S3, Supplementary Material online) (Schwager et al. 2017; Nolan et al. 2020; Gainett et al. 2021; Harper et al. 2021; Ontano et al. 2021). Across all matrices, genes exhibited similar distributions of saturation, evolutionary rate and missing data, regardless of support for a monophyletic Arachnida or for Xiphosura nested in Arachnida (fig. 4b).

Furthermore, we discovered that genes supporting arachnid monophyly were shorter and exhibited fewer parsimony informative sites than genes supporting the unconstrained topology, across all matrices (fig. 4b). Short genes with low informativeness have been linked to systematic error across an array of phylogenomic data sets (Shen et al. 2016; Mongiardino Koch 2021), suggesting that arachnid

monophyly may reflect noise rather than true phylogenetic signal. Consistent with this interpretation, we found that sites supporting arachnid monophyly exhibited higher Shannon entropy than sites supporting a nested Xiphosura (fig. 4c). Sites supporting arachnid monophyly were fewer in number and had higher Shannon entropy even when compared with sites supporting a rejected grouping (Dromopoda) that has been falsified by rare genomic changes (supplementary fig. S4, Supplementary Material online).

Combined Analyses of Morphology and Molecules

To assess the impact of morphological data, we began with the character matrix of Huang et al. (2018), the most comprehensively coded morphological matrix of extant chelicerates to date, including recently discovered arachnid fossils that have impacted reconstruction of ancestral states. To this matrix, we added the sea spider *Flagellopantopus blocki* and two exemplars of the extinct order Phalangiotarbida from codings in the literature, as well as all extant chelicerates in the molecular matrix. Errors previously entered in the character coding were corrected (discussed by Sharma et al. 2021). We added new characters from the recent literature pertaining to the neuroanatomy of Xiphosura and several arachnid orders, as well as previously overlooked character systems.

To overcome artifacts stemming from missing and inapplicable character partitions, non-chelicerate outgroup taxa (Onychophora, Mandibulata) were removed from this analysis. For the same reason, we excluded putative chelicerate stem groups of questionable and controversial placement for which molecular sequence data are inapplicable. Pycnogonida was used to root Euchelicerata.

When analyzed by itself under equal weights parsimony, the morphological data set yielded little basal resolution (fig. 5a). A strict consensus of 1,000 equally parsimonious trees recovered a basal polytomy of Euchelicerata. Various interordinal relationships received negligible nodal support, although they accorded closely with recent morphological analyses viz. the recovery of Tetrapulmonata (including Trigonotarbita and Haptopoda) and Acaromorpha (Ricinulei + Acari). Under a Bayesian inference approach (fig. 5b), the morphological data set recovered the monophyly of Panscorpiones, Acaromorpha, and Acari, albeit without support (PP < 0.95), and Tetrapulmonata (including Trigonotarbita and Haptopoda) with support (PP = 0.99). Both approaches recovered the monophyly of Merostomata (a grouping of the marine taxa Xiphosura, Synziphosurina, Eurypterida, and Chasmataspidida; BS = 90%; PP = 1.00).

Maximum likelihood analysis of the combined matrix (fig. 5c) also recovered monophyly of Merostomata, which in turn was recovered as the sister group to Ricinulei. We recovered Trigonotarbita as part of the tetrapulmonates, consistent with the presence of two pairs of book lungs in these groups. Phalangiotarbida was recovered as the sister group of Opiliones. Key fossil taxa were recovered in expected placements, such as the harvestman suborder Tetraophthalmi, and the orders Uraraneida and Haptopoda. Palpigradi was recovered as the sister group of the remaining Euchelicerata with moderate support, paralleling the result of

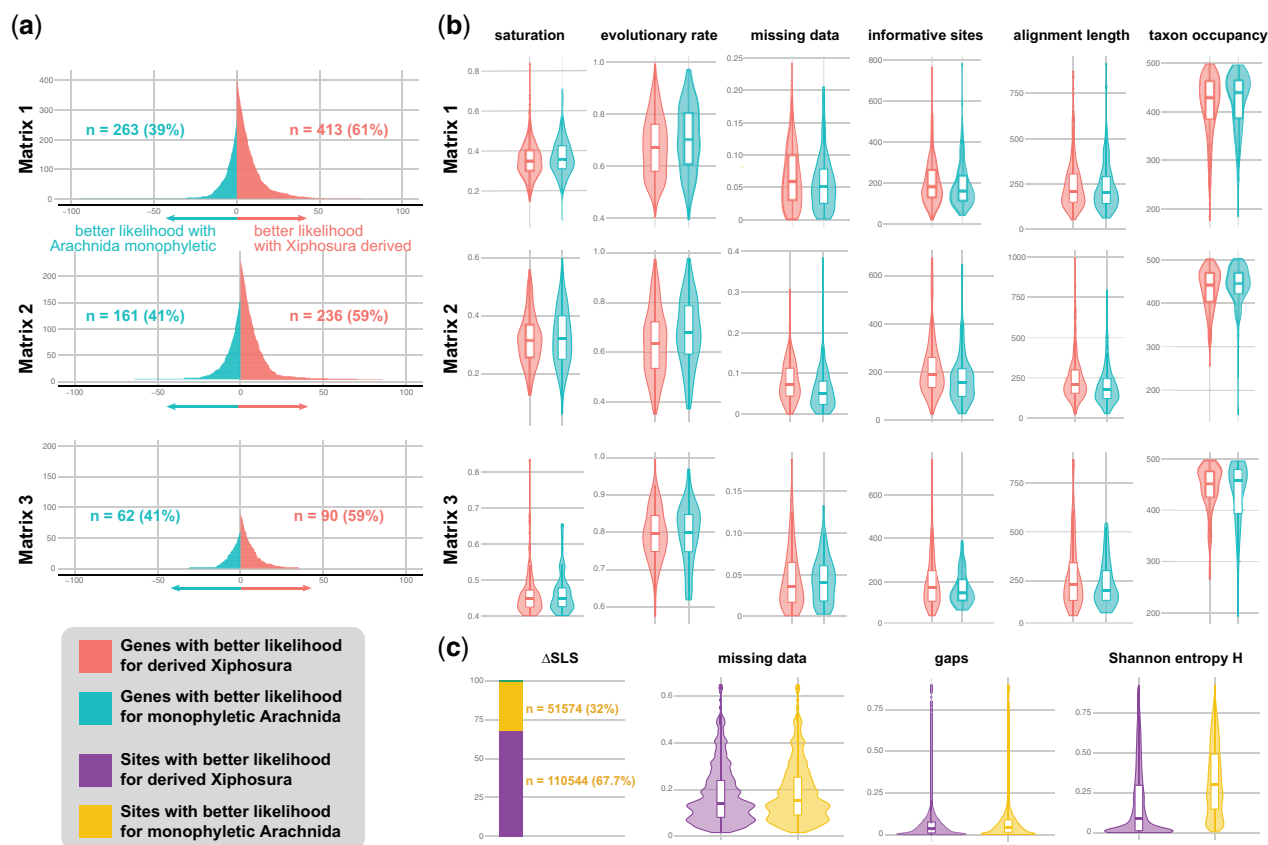


FIG. 4. Dissection of phylogenetic signal shows that a minority of genes supports arachnid monophyly. (a) Δ GLS distributions mapping phylogenetic support for competing hypotheses reveal that a minority of genes (39–41%) supports arachnid monophyly, regardless of orthology criterion (Matrices 1 and 2) and filtering of fast-evolving genes (Matrix 3). These proportions are similar to the proportions of genes supporting falsified groupings (supplementary fig. S3, Supplementary Material online). (b) Genes supporting the derived placement of Xiphosura exhibit comparable or better metrics of systematic bias (e.g., saturation, evolutionary rate, missing data) than genes supporting Arachnida. (c) Δ SLS distributions reveal that the majority of sites (68%) support a derived placement of Xiphosura. Whereas the two categories of sites are similar with respect to missing data, sites supporting arachnid monophyly exhibit high levels of Shannon entropy (exceeding entropy values for falsified groupings; supplementary fig. S4, Supplementary Material online).

the RL2 recoding strategy. Compared with molecular analyses, support values were lowered by the inclusion of morphological data in the combined analysis, a result attributable to the instability incurred by data-poor fossil taxa.

Discussion

Arachnid Monophyly Is Not Supported by Phylogenomic Approaches

Molecular results that recover nontraditional groupings are often labeled as artifacts, especially when morphological patterns and long-held evolutionary scenarios come under question. Like the basal topology of groups like Metazoa, Aves, and Angiospermae, the basal topology of Euchelicerata has long defied stability in molecular data sets. Proposals to “correct” the tree and recover arachnid monophyly using molecular data sets have included restricting analyses to slowly evolving genes (or less saturated genes, a correlate of evolutionary rate) (Sharma et al. 2014; Lozano-Fernández et al. 2019; but see Ballesteros and Sharma 2019), expansion of taxonomic sampling (Lozano-Fernández et al. 2019; but see Ballesteros et al.

2019), the use of site heterogeneous models (Lozano-Fernández et al. 2019; Howard et al. 2020; but see Sharma et al. 2014; Ballesteros and Sharma 2019; Ontano et al. 2021), or some combination thereof. As our analyses show, the derived placement of Xiphosura (possibly with the other merostomate orders) is consistently recovered despite concomitant application of all these putative corrections.

Why have some recent molecular data sets been able to recover arachnid monophyly (albeit with incomplete sampling of arachnid orders)? As previously shown, the matrices of Lozano-Fernández et al. (2019) and Howard et al. (2020) exhibit a number of bioinformatic and analytical artifacts (Ballesteros et al. 2019; Ontano et al. 2021). Upon further reexamining those data sets, we found an unexpectedly high number of outliers in root-to-tip distances across gene trees. Using an annotation strategy based on the *Drosophila melanogaster* proteome, we discovered that the cause of this noise was the widescale inclusion of paralogs in these data sets. Specifically, 29% (68/233) of loci in the Lozano-Fernández et al. Matrix A (constructed using an assemblage of preselected genes, assembled by a BLAST-based strategy), and 41%

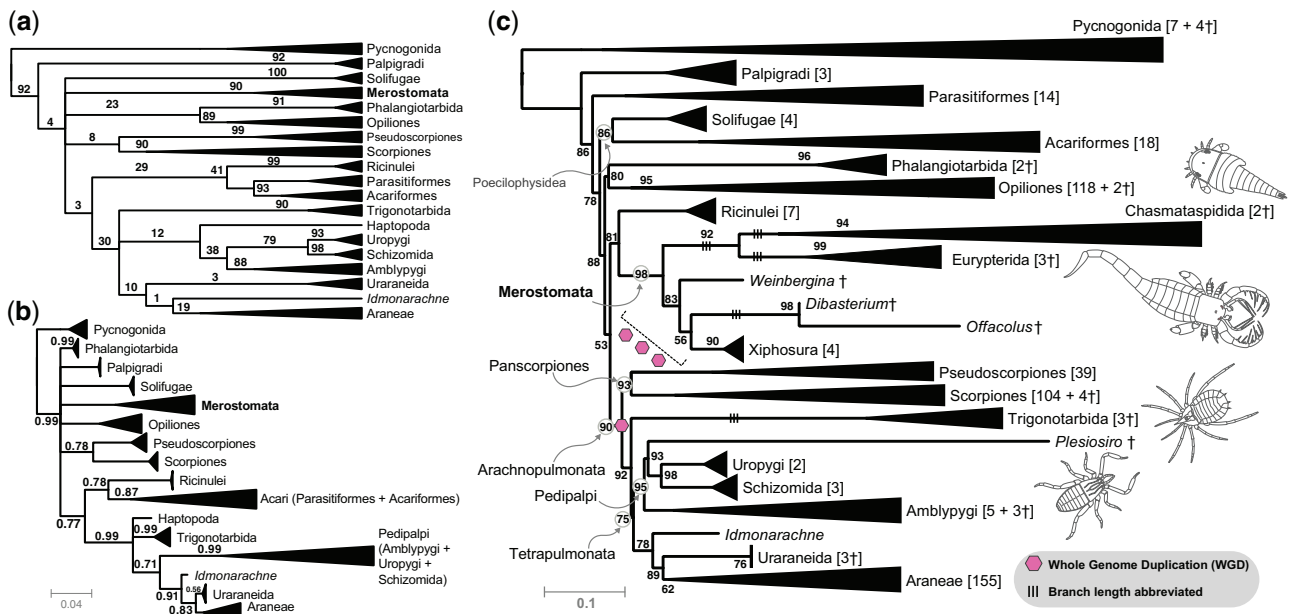


FIG. 5. Inclusion of morphology does not rescue arachnid monophyly in total evidence analyses. (a) Strict consensus of 1,000 equally parsimonious trees inferred for a morphological matrix of 259 morphological characters (482 extant and 32 fossil taxa). Merostomata comprises extinct groups Eurypterida (sea scorpions), Chasmataspidida, and Synziphosurina, as well as Xiphosurida, including the extant Limulidae. Numbers on nodes indicate bootstrap resampling frequencies. (b) Summary tree from Bayesian inference analysis of the morphological matrix. Numbers on nodes indicate posterior probabilities below 1.00; unlabeled nodes are maximally supported. (c) Maximum likelihood total evidence topology based on 152 slowly evolving genes and morphological characters. Numbers on nodes represent bootstrap resampling frequencies; unlabeled nodes are maximally supported. Note that the timing of WGD events in Xiphosura cannot be pinpointed on the branches subtending this group.

(82/200) of loci in the Howard *et al.* 200-locus matrix (constructed using a distance algorithm-based orthology inference strategy) included paralogs, often from distantly related multigene families (supplementary tables S7 and S8, Supplementary Material online). One potential solution to the pitfall of paralogy detection is the use of phylogenetically informed orthology inference criteria (e.g., PhyloTreePruner; UPhO), which has been shown to outperform distance-based prediction strategies (Kocot *et al.* 2013; Ballesteros and Hormiga 2016).

Could the properties of orthologs that are able to recover arachnid monophyly inform the selection of “better” loci for chelicerate phylogenomics? To address this, we examined the distribution of phylogenetic signal in our data sets for genes and sites supporting arachnid monophyly, versus the unconstrained topology, using Δ GLS and Δ SLS approaches (Shen *et al.* 2017). Genes supporting the nested placement of Xiphosura exhibited no evidence of systematic biases compared with the minority of genes supporting arachnid monophyly (39–41%). Instead, we discovered the opposite trend: genes supporting arachnid monophyly tended to have shorter alignment lengths and fewer informative sites than genes supporting a nested Xiphosura, properties that are associated with phylogenetic error (Shen *et al.* 2016; Aharon *et al.* 2019; Mongiardino Koch 2021). Consistent with this interpretation, sites supporting arachnid monophyly exhibited higher Shannon entropy and low structure (i.e., greater randomness). For context, the proportions of genes supporting a grouping that has been clearly refuted by genome

architecture (i.e., Dromopoda, which historically united two arachnophiles orders with two apulmonate orders) are nearly identical to those supporting arachnid monophyly (supplementary fig. S3, Supplementary Material online). Moreover, the number of sites supporting this refuted grouping is higher than those supporting arachnid monophyly (supplementary fig. S4, Supplementary Material online).

These analyses suggest that support for arachnid monophyly does not reflect hidden signal, so much as noise and error in the data sets that have putatively supported this grouping. We submit that the sum of our analyses, however counterintuitive, may reflect a phylogenetically accurate relationship—Xiphosura (and possibly the other merostomates) may simply constitute derived arachnids.

Slowly Evolving Genes and Site Heterogeneous Models Overcome LBA Artifacts in Chelicerate Phylogeny

As anticipated, several groups in our phylogeny reflected long root-to-tip distances, constituting lineages prone to LBA artifacts. The inclusion of Opilioacariformes, a slowly evolving group of Parasitiformes (Klomp *et al.* 2007; Pepato *et al.* 2010), was recently shown to break up the grouping of Acariformes and Parasitiformes, suggesting that Acari is a long-branch artifact (Ontano *et al.* 2021; Ontano *et al.* 2022). In this study, we sampled Opilioacariformes with two libraries, and concordantly, never obtained the monophyly of Acari, particularly when pursuing approaches best suited to mitigating LBA (fig. 3). This outcome suggests that the correspondences of acariform and parasitiform bauplans

(e.g., mouthparts; patterns of tagmosis) represent a case of morphological convergence in chelicerates. Similar convergence of mouthparts occurs in the gnathobasic preoral chambers of Opiliones and Scorpiones, which were previously grouped by a subset of morphological analyses (Shultz 1990, 2007); and the presence of a hexapod larva, another supposed character in support of Acari, which is also shared with Ricinulei.

Phylogenomic subsampling for slowly evolving genes did recover Panscorpiones within Arachnoplumonata (fig. 2b), a result that is attributable to a marked shift in the proportion of genes supporting this group as a function of evolutionary rate (supplementary fig. S5, Supplementary Material online). However, even in maximum likelihood analyses that prioritized slowly evolving genes, we recovered Acariformes and Parasitiformes clustered near the base of the euchelicerate tree, placements that we regarded as possible LBA artifacts. Upon analyzing the slowly evolving matrix with site heterogeneous models in a Bayesian framework (CAT-GTR in PhyloBayes), not only were Panscorpiones and Arachnoplumonata recovered, but this approach also resolved Acariformes as the sister group of Solifugae (=Poecilophysidea), with Poecilophysidea in turn sister group to Palpigradi (=Cephalosomata), with support (fig. 3a). Four-state recoding in tandem with site heterogeneous models eroded all support from the base of Euchelicerata, but this analysis did recover Poecilophysidea as well (fig. 3c).

Intriguingly, Poecilophysidea and Cephalosomata were previously supported by a minority of Sanger-based phylogenetic analyses and were proposed on the basis of patterns of anterior sclerotization in these orders (Alberti and Peretti 2002; Pepato et al. 2010; Dunlop et al. 2012), potentially validating a subset of morphological character systems in chelicerate higher-level phylogeny. A proximate relationship of Palpigradi and Solifugae is also supported by the anatomy of the coxal gland of these orders (Ballesteros et al. 2019). Given the species richness of both Acariformes and Parasitiformes, future efforts to clarify the relative placements of these groups must focus on increasing the representation of basal nodes, a strategy that has been shown to outperform algorithmic and data trimming solutions to resolving the placement of pseudoscorpions (Ontano et al. 2021).

Morphology May Be Confounded by Convergence in Chelicerate Phylogeny

Unlike in other animal clades, the addition of morphological data to molecular partitions does not ameliorate the discordance with the traditional phylogeny of chelicerates; we found that combining morphological and molecular data sets using model-based approaches recovers Merostomata (the marine group that includes horseshoe crabs) as nested within Arachnida. The notion that morphological synapomorphies of Arachnida can outweigh the dissonance found in molecular data found no support in this study. Furthermore, only in combination with molecular data was morphology able to recover clades supported by rare genomic characters (Panscorpiones within Arachnoplumonata); by itself, morphology has never recovered this arrangement

of Arachnoplumonata, either in our analysis or in any previous efforts.

One caveat of our combined analysis is that outgroups like putative stem-groups of Chelicerata (e.g., megacheirans) were not included, as their phylogenetic position is controversial even in morphological data sets (Wolfe 2017; Siveter et al. 2017; Giribet and Edgecombe 2019). The exclusion of these groups may prevent character states from being optimized correctly, such as biramous appendages (the presence of exopods), faceted eyes, and gnathobasic mouthparts in marine groups. To assess this possibility, we trialed fusing our molecular data set (Matrix 3) to two morphological matrices from the literature with widely different taxon sets: a recent, densely sampled matrix of marine crown-group Chelicerata (Bicknell et al. 2019); and a broadly sampled matrix of Panarthropoda (Siveter et al. 2017). Nonoverlapping terminals with molecular data only were removed from these analyses to reduce missing data. Upon fusion with Matrix 3, these supplementary data sets featured minimal sampling of extant arachnid fauna (typically, one exemplar per order), as well as greater proportions of missing data in comparison to our combined matrix. We found that combining data classes destabilized the traditional relationships previously predicted by those studies, either incurring the non-monophyly of Euchelicerata (supplementary fig. S6a and b, Supplementary Material online) or of Chelicerata (supplementary fig. S6c and d, Supplementary Material online). Within Euchelicerata, data sets that broadly represented panarthropod diversity (fossil and extant; Siveter et al. 2017) recovered a nested placement of Merostomata within the arachnids when combined with molecular data (supplementary fig. S6c and d, Supplementary Material online), closely paralleling our results.

These outcomes suggest that morphological data partitions seeking to capture deep chelicerate relationships may feature far less robust phylogenetic signal than commonly portrayed, specifically in a total evidence framework. Concordantly, a recent paleontological study failed to recover even Tetrapulmonata (Wolfe 2017), the only higher-level group that is consistently recovered by most morphological and molecular data sets. Another recent paleontological study that recovered arachnid non-monophyly took the step of constraining Arachnida a priori to ensure the recovery of the traditional topology (Aria and Caron 2019). Tellingly, the morphological data sets that have historically exhibited the greatest congruence with molecular results (viz. Panscorpiones, Poecilophysidea, and Cephalosomata) are those that densely sampled fossil and extant exemplars of Arachnida (e.g., Garwood and Dunlop 2014; Huang et al. 2018). This trend suggests that undersampling extant arachnid diversity and character states in morphological matrices may underlie some of the dissonance between paleontological works and molecular topologies.

Admittedly, the scenario of a nested Xiphosura invites entrenched skepticism, particularly from adherents of paleontology. In addition to an extensive fossil record, horseshoe crabs exhibit an array of putatively plesiomorphic traits that are suggestive of a basally branching placement. The fossil record of merostomates is thought to represent a stepwise

colonization of land via internalization of the book gill of these marine groups (for this reason, the position of scorpions at the base of the Arachnida was a central tenet of this evolutionary transformation series). Recent arguments in favor of arachnid monophyly (Howard et al. 2020) have thus focused on the merostomate faceted eye, which is thought to reflect the ancestral condition; the gnathobasic (enditic) mouthparts of merostomates; the biramous condition of merostomate appendages; and the anatomy of the xiphosuran book gill, which shares correspondences with the book lung of large-bodied arachnids (e.g., scorpions; basally branching spiders). Moreover, arachnid monophyly has historically been defended on the basis of a series of characters, with some of these stemming from the musculoskeletal system (Shultz 2001; but see Sharma et al. 2021).

However, a comparison with the history of mandibulate arthropod phylogeny offers compelling reasons to doubt the linearity of morphological evolutionary scenarios. Within Myriapoda, only one order of centipedes (Scutigermorpha) has retained the faceted compound eye commonly found in extant pancrustaceans and mandibulate fossil outgroups, whereas all other myriapod groups bear ocelli (in some cases, as “semi-compound” aggregations) or are blind (Oakley 2003). This character state distribution suggests that faceted eyes are prone to discretization and loss in terrestrial habitats in a group that colonized land at least by the Devonian (Giribet and Edgecombe 2019). Paralleling this trend, various fossil arachnid groups (e.g., fossil scorpions, Trigonotarbidia, and fossil Ricinulei) exhibit “semi-compound” eyes (comparable to those found in pleurostigmophoran centipedes and many millipedes) in head regions positionally homologous to the faceted eyes of Xiphosura and Eurypterida (Dunlop 2010; Howard et al. 2020). The faceted eyes of merostomates may reflect a plesiomorphic condition retained deep in the euchelicerate tree, like the faceted eye of scutigermorphs within myriapods.

Similarly, discussions of the gnathobasic mouthparts of merostomates echo historical debates over the nature of the gnathobasic mandible of terrestrial mandibulates, as well as other correspondences of head appendages. It was previously thought that Hexapoda and Myriapoda constituted sister groups (the clade Tracheata), a relationship supported by their putatively shared gnathobasic mandible, appendage-free intercalary segment, uniramous appendages, and arrangement of the respiratory organs (tubular tracheae, typically opening as paired spiracles on pleural territories of trunk segments). The gradual overturning of this relationship by molecular phylogenies in favor of the Pancrustacea hypothesis revealed that striking morphological convergences could occur in distantly related taxa as a result of common selection pressures in terrestrial environments (Giribet et al. 2001; Mallatt et al. 2004; Mallatt and Giribet 2006; Giribet and Edgecombe 2019; Edgecombe 2020). In this light, the reduction of gnathobasic mouthparts in terrestrial chelicerate orders could also reflect parallel losses as adaptations to life on land, as evidenced by the uniramous, gnathobasic mandibular architecture of hexapods and myriapods. Parallel

losses of secondary rami and simplification of appendages are also broadly observed in terrestrial arthropods, such as arachnids, myriapods, hexapods, and terrestrial malacostracans (e.g., Isopoda, Amphipoda). We submit that the morphology of merostomate appendages is closely tied to evolution in marine habitats and may reflect retention of plesiomorphies; the absence of these structures in terrestrial arthropod groups does not offer compelling evidence uniting Arachnida.

Convergent evolution of tracheal tubules in other terrestrial groups, such as Onychophora, Hexapoda, and Myriapoda, falsifies the interpretation that a lung-like organ is a necessary stepping-stone to the acquisition of tracheal tubules in chelicerates. The conventional and simplistic evolutionary transformation series of book gill to book lung to tracheal tubule is deeply undermined by the complexity of respiratory organ evolution in Chelicerata. This point is underscored by the recent discovery of a eurypterid with trabeculate respiratory organs well after the appearance of arachnids in the fossil record (340 Ma; Lamsdell et al. 2020), secondarily marine scorpions with lamellate gills (Waeringoscorpio; Dunlop 2010; Howard et al. 2020), and the diversity of modern aquatic mites (Dabert et al. 2016). The recent recovery of Pseudoscorpiones as a derived member of Arachnopumonata, as well as investigations of respiratory structures across spiders, reveals that book lungs have been frequently lost and repeatedly transformed into tracheal tubules, with loss of book lungs observed in multiple miniaturized arachnopulmonate groups (e.g., the posterior book lung pair of Schizomida and most araneomorph spiders; complete loss of book lungs in miniaturized spiders and pseudoscorpions) (Ontano et al. 2021; Ramírez et al. 2021; Lopardo et al. 2021). There is no compelling evidence that evolutionary transitions of respiratory organs have followed a simple, linear series at the base of Arachnida, nor that water-to-land (or the reverse) transitions are rare or irreversible in the arthropod fossil record.

As for the putative musculoskeletal synapomorphies established for Arachnida (Shultz 2001), we submit that the evolution of this entire character system may be closely tied to the selective pressures of a terrestrial lifestyle. Arthropod appendages are highly adaptive structures, and biomechanical demands on locomotory appendages differ greatly between aquatic and terrestrial organisms (Boxshall 2004). As with the correspondences of insect and myriapod musculoskeletal anatomy, there is no evident reason why the musculoskeletal system would constitute a homoplasy-free data source for arachnids.

Taken together, morphological character systems that putatively support arachnid monophyly tend to exhibit high levels of homoplasy upon closer investigation, especially when examining their counterparts in Mandibulata. Given the remarkable morphological convergences exhibited by Hexapoda and Myriapoda, we postulate that parallel evolution in terrestrial arthropod groups may confound inferences of homology in morphological data sets. Although no morphological characters overtly support a closer relationship of Xiphosura to any subset of arachnid orders (but see Lehmann

and Melzer 2013, 2019a, 2019b), the absence of morphological support for numerous, robustly recovered molecular clades is a common feature of ancient invertebrate relationships, as exemplified by the modern higher-level phylogeny of groups like Annelida, Nematoda, and Mollusca (Struck et al. 2015; Smythe et al. 2019; Kocot et al. 2020). Indeed, the discovery of a particular well-supported relationship in molecular data sets typically serves as a catalyst for revitalized morphological study and reinterpretation of previous homology statements, as in the case of Pancrustacea, Ecdysozoa, Arachnospulmonata, and Ambulacraria (Dohle 2001; Richter 2002; Kristensen 2003; Mayer 2006; Lehmann and Melzer 2019a, 2019b; Kapli et al. 2021; Ontano et al. 2021). Given the recovery of Poecilophysidea and Cephalosomata in site heterogenous model-based analyses (fig. 3), reexamination of previously overlooked interordinal groupings may provide a better understanding of hidden phylogenetic signal in specific chelicerate morphological character systems (Pepato et al. 2010; Dunlop et al. 2012). A derived placement of merostomates as a group more proximal to Arachnospulmonata could also reconcile the morphology of extinct marine groups like eurypterids with the unambiguously nested position of Scorpiones, a hypothesis that could be tested through functional genetic approaches to understand the developmental basis for respiratory organ patterning in horseshoe crabs, arachnospulmonates, and apulmonate arachnids.

The nested placement of Xiphosura, together with the reconstruction of multiple terrestrialization events across a grade of arachnid diversity, must be treated as a valid competing hypothesis. Future efforts to integrate new phylogenetic data classes and rare genomic characters (e.g., Schwager et al. 2017; Gainett et al. 2021; Ontano et al. 2021) may offer clearer resolution of relationships among the apulmonate arachnid orders and concision between discordant data sets.

Conclusion

Analyses of molecular data and total evidence phylogenetic approaches do not support arachnid monophyly. The concept of Arachnida may reflect the intuitive idea that terrestrialization is rare or costly in evolutionary history. As revealed by the history of groups like mandibulate arthropods, nematodes, and pulmonate gastropods, terrestrialization has not only evolved many times independently within such taxa, but is also the cause of remarkable and misleading cases of morphological convergence. Even on relatively shallow timescales, repeated colonization of terrestrial environments has been substantiated in arthropod groups like terrestrial Decapoda (at least ten water-to-land transitions; Watson-Zink 2021) and terrestrial Isopoda (at least two transitions; Dimitriou et al. 2019) through the lens of molecular phylogenetics.

The strongest evidence that morphological data sets of Chelicerata may be prone to misinterpretation of homologies is provided by the positions of scorpions and pseudoscorpions, which are united with the other arachnospulmonates by a rare genomic change (an ancient whole-genome duplication event). Morphological data sets, including the data set we generated, have consistently

failed to recover this grouping (with or without the miniaturized Pseudoscorpiones). If morphological data sets can falter in the recovery of the only higher-level chelicerate groups robustly resolved by molecular data and genomic architecture, it stands to reason that phylogenetic signal in morphological data sets may not be sufficiently robust to adjudicate other nodes in chelicerate interordinal phylogeny. Given the history of erstwhile morphological groupings like Tracheata, Uniramia, Articulata, Polychaeta, Pulmonata, Opisthobranchia, Aschelminthes, and numerous others, we postulate that phylogenomic approaches to deep metazoan relationships should treat morphological interpretations with measured skepticism *prima facie*, especially in the context of selective pressures like terrestri- alization that promote morphological convergence and thereby confound inferences of homology.

Future efforts to integrate morphology into the new phylogeny of Chelicerata may be aided by parametric tests for phylogenetic signal across anatomical character systems, with the goal of quantifying informativeness and assessing noise in anatomical partitions (e.g., Bieler et al. 2014; King 2019). Exploration of signal within both morphological and molecular data sets, in tandem with alternative recoding strategies, may be key to identifying congruence between dissonant data classes (e.g., Kulkarni et al. 2021; Lopardo et al. 2021; Redmond and McLysaght 2021). More generally, a multidimensional, modern view of morphological evolution should emphasize implementation of comparative genetic techniques for testing the shared developmental basis of putative homologies (e.g., Smith et al. 2016; Nakamura et al. 2017; Bruce and Patel 2020; Clark-Hachtel and Tomoyasu 2020), especially as it pertains to body plan diversification and the evolution of anatomical disparity. In tandem with such approaches, investigations of chelicerate phylogeny should emphasize the expansion of whole-genome data sets, with the aim of leveraging rare genomic changes, such as genome duplications, as potential arbiters of competing phylogenetic hypotheses (Schwager et al. 2017). This strategy has shown strong promise as a source of benchmarks for assessing performance of molecular data sets in the face of LBA, as well as for realigning interpretations of morphological data, toward reconciliation of chelicerate relationships (Ontano et al. 2021; Ontano et al. 2022).

Materials and Methods

Details of the methods below are provided in the [Supplementary Material](#) online.

Taxon Sampling and Orthology Inference

Taxon selection consisted of 24 outgroup and 482 ingroup terminals; these 506 transcriptomes and genomes ([supplementary table S1, Supplementary Material](#) online) sampled all extant chelicerate orders with multiple terminals. Seventy five new libraries were generated following previously published protocols. Proteomes and peptide sequences were used as inputs. Phylogenetically-informed inference of orthologs leveraged a recent *de novo* computation of orthologous genes for Chelicerata using UPhO (3534 loci identified

previously [Ballesteros and Sharma 2019]). For validation, these collections of putative orthologous sequences were BLASTed (blastp v. 2.9.0+ [Camacho et al. 2009]) against the *D. melanogaster* proteome for annotation using the best hit. Sequences not matching the most common annotation were discarded. Separately, orthologs benchmarked using the BUSCO-Arthropoda v. 3 database were analyzed independently. The set of complete, single-copy BUSCOs > 100 amino acids in length was retained from each library.

Matrix Construction

Initial sets of orthologs were filtered based on taxon decisiveness (Steel and Sanderson 2010). We retained only loci that had at least one terminal for all the following clades: Araneae, Pedipalpi (Uropygi + Schizomida + Amblypygi), Scorpiones, Ricinulei, Xiphosura, Solifugae, Opiliones, Palpigradi, Parasitiformes (treated here as the order uniting Holothyrura, Ixodida, Mesostigmata, and Opilioacariformes), Acariformes (treated here as the order uniting Sarcopitiformes, Trombidiformes, and the non-monophyletic assemblage “Endeostigmata”; Klimov et al. 2018), Pseudoscorpiones, Pycnogonida, Pancrustacea, Myriapoda, and Onychophora. Applying this criterion, we reduced the UPhO ortholog set to 676 loci (Matrix 1) and the BUSCO set to 399 loci (Matrix 2).

To assess the impact of saturation, we generated saturation plots for each locus and isolated a subset of 152 loci with slope ≥ 0.4 and $r^2 \geq 0.95$; these loci were concatenated to form Matrix 3. To operate PhyloBayes-mpi v. 1.8 with the computationally demanding CAT+GTR+ Γ model, we selected 56 representative terminals from the slow-evolving data set (Matrix 2) such that major taxonomic groups (supplementary table S4, Supplementary Material online) were each represented by three to five terminals, major basal splits were represented in each lineage, and the selected taxa exhibited high data completeness. This data set was further filtered with BMGE v 1.12 (Criscuolo and Grigaldo 2010) to remove heteroprecipitous sites to form Matrix 4 for analysis with PhyloBayes-mpi.

Partitioned Analyses

Gene trees were inferred using IQ-TREE v. 1.6.10 (Nguyen et al. 2015) with model-fitting using ModelFinder (Kalyaanamoorthy et al. 2017) and nodal support estimation using the ultrafast bootstrap (Hoang et al. 2018) as follows: *iqtree -mset LG, WAG, JTT, Dayhoff, JTTDCMut, DCMut, PMB -m MFP -bb 1000*. Maximum likelihood analyses of concatenated data sets (Matrices 1–3) were run using a gene partitioning strategy implementing the best substitution models identified during gene tree reconstruction. Tree topologies were inferred using IQ-TREE, with nodal support estimated using ultrafast bootstrapping. Summary coalescent estimates of the species phylogenies were estimated from the individual gene trees using ASTRAL v 3.14.2 (Zhang et al. 2018).

Mixture Model Analyses

We computed maximum likelihood analyses with the posterior mean site frequency model (Wang et al. 2019) for Matrices 1–3, using the LG+C20+F+ Γ implementation.

The use of more site categories (e.g., C60) proved prohibitive for a data set of this size, with the C20 model demanding 1.1 Tb of RAM to compute site-specific model parameters (see also Ontano et al. 2021 for better performance of the C20 model in chelicerate data sets, using rare genomic changes as benchmarks). Analyses were computed using IQ-TREE v. 1.6.10. For all matrices, a simple ML tree search (*-m LG+F+G -bb 1000*) was estimated as the starting tree, followed by estimation of site frequencies and ML searches under the mixture model. Nodal support was estimated using ultrafast bootstrapping. Bayesian inference analysis was performed using PhyloBayes-mpi v. 1.8 (Lartillot et al. 2013) and the CAT+GTR+ Γ model on Matrix 4, which was optimized for this purpose. Bayesian inference analysis was run on eight independent chains for > 20,000 cycles (supplementary tables S5 and S6, Supplementary Material online). Convergence of parameters and topologies was assessed using Tracer 1.7.1 (Rambaut et al. 2018) and native PhyloBayes-mpi summary programs. Summary statistics and chain lengths are provided in the Supplementary Material. Trace files of parameters and tree files from each run are provided in FigShare. Convergence parameters exhibited differences as a function of combining different chains. Different combinations of chains produced varying maximum split differences. Examination of ESS values and a posteriori tree distribution across all eight chains showed that summary statistics broadly exhibited convergence; the high value of the maximum split difference (*maxdiff*) is driven by a soft polytomy at the base of Euchelicerata.

Recoded Mixture Model Analyses

We implemented a site heterogeneous mixture model approach to partitioned phylogenomics using the recently proposed RL2 strategy (Redmond and McLysaght 2021), which implements four-state recoding of amino acid data (SR4; Susko and Roger 2007). Analyses were performed in IQ-TREE v. 1.6.10, following the original implementation (Redmond and McLysaght 2021).

Analyses of Phylogenetic Signal

Δ GLS and Δ SLS approaches were implemented as originally described by Shen et al. (2017) and following our previous approaches (Ballesteros and Sharma 2019). Comparisons of properties of genes and sites were performed in R and visualized using ggplot2 (Wickham 2016).

Morphological Analysis

We developed a morphological matrix of 259 characters coded for 482 extant and 32 fossil chelicerates. Given the unambiguous recovery of Pycnogonida as the sister group to the remaining chelicerates, Pycnogonida were used to root the tree. Character codings were drawn from previous higher-level analyses of sea spiders (Arango 2002), harvestmen (Giribet et al. 2002; Garwood et al. 2014), scorpions (Sharma et al. 2015), and arachnids (Giribet et al. 2002; Shultz 2007; Garwood and Dunlop 2014; Huang et al. 2018). Errors and discrepancies with previous character codings were modified (Sharma et al. 2021) and we additionally coded

new characters informed by recent investigations (Lehmann and Melzer 2013, 2019a, 2019b). Fossil taxa were coded using original descriptions from the literature.

Model-fitting for the morphological data set was performed in IQ-TREE v. 1.6.10 from the data set initially partitioned based on the number of character states. Bayesian analyses using the same partitioning scheme were performed in MrBayes v 3.2.7a (Ronquist et al. 2012) using the Mk1 model with unlinked rate and state frequency parameters per partition. The analyses consisted of four independent runs of 50 M cycles. Equal weights and implied weights parsimony analyses were performed using TNT v. 1.5 (Goloboff and Catalano 2016) (*mxfam* 4000; *nstates* 32; *xinact*; *hold* 10000; *mult*=level 7). Implied weighting was performed using a family of values of concavity.

We additionally performed total evidence analyses using two recently published morphological matrices with differing representations of stem-group chelicerate taxa, (Bicknell et al. 2019; Siveter et al. 2017), complemented by molecular data from Matrix 3. In cases of nonoverlapping taxa, a chimeric terminal was constructed using the closest related species to a given terminal in the morphological data sets. For each chimeric terminal, character codings were checked to ensure their applicability to their morphological counterpart; no coding changes were required for chimeras (supplementary table S9, Supplementary Material online). Due to the degree of missing data in these matrices, analyses were only performed using parsimony (equal and implied weights); model-based analyses consistently failed to converge for these supplementary data sets.

Acknowledgments

Discussions with Jason A. Dunlop, Russell J. Garwood, and Hans Klompen refined some of the ideas presented. Comments from the Editor-in-Chief, Associate Editor, and three anonymous reviewers improved a previous draft of the work. Sequencing was performed at the Bauer Core Facility (Harvard University) and the BioTechnology Center (UW-Madison). High-throughput computational analyses were performed through the Center for High Throughput Computing (CHTC) and the Bioinformatics Resource Center (BRC) of UW-Madison. J.A.B. was supported by a Guyer postdoctoral fellowship. C.E.S.L. was supported by a CONACYT postdoctoral fellowship (reg. 207146/454834). This material is based on work supported by the National Science Foundation grant nos. IOS-1552610 and IOS-2016141 to P.P.S.; DEB-1457539 to G.Giribet; DEB-1457300 and DEB-1754289 to G.H.; and the US-Israel Binational Science Foundation grant no. BSF-2019216 to P.P.S. and E.G.R.

Author Contributions

J.A.B. and P.P.S. designed the project. J.A.B., C.E.S.L., G. Giribet, and P.P.S. designed the taxon sampling strategy. J.A.B., C.E.S.L., C.M.B., L.R.B., T.J.C., G. Gainett, A.Z.O., E.V.W.S., E.G.R., M.S.H., G. Giribet, and P.P.S. performed fieldwork. J.A.B., C.E.S.L., C.M.B., L.R.B., T.J.C., G. Gainett, A.Z.O., E.V.W.S., and P.P.S. performed sequencing. J.A.B., C.E.S.L., and P.P.S. coded

morphological data, with input from C.P.A., M.S.H., W.C.W., G.H., and G. Giribet. J.A.B. performed bioinformatics and phylogenomic analyses. J.A.B. and P.P.S. wrote the manuscript. P.P.S. supervised the work. All authors edited the manuscript.

Data Availability

Raw sequence data that support the findings of this study have been deposited in NCBI Sequence Read Archive. The morphological data matrix and the character list are provided in the supplementary files S1 and S2, Supplementary Material online. All transcriptomic assemblies, gene alignments, gene trees, supermatrices, phylogenomic trees, PhyloBayes-mpi trace files, and scripts have been deposited in FigShare (<https://figshare.com/s/e3056003b19f2c814b83>, <https://figshare.com/s/288283c1cad2d5743a27>, <https://figshare.com/s/3f1e03e3fff477deece2>, <https://figshare.com/s/3c4ff6f6c53a-b4758ea4>, <https://figshare.com/s/1769b233364be501e520>, <https://figshare.com/s/69b2f5a3542e71ed7a4c>).

References

- Aharon S, Ballesteros JA, Crawford AR, Friske K, Gainett G, Langford B, Santibáñez-López CE, Ya'aran S, Gavish-Regev E, Sharma PP. 2019. The anatomy of an unstable node: a Levantine relict precipitates phylogenomic dissolution of higher-level relationships of the armoured harvestmen (Arachnida: Opiliones: Laniatores). *Invertebr Syst.* 33:697–717.
- Alberti G, Peretti AV. 2002. Fine structure of male genital system and sperm in Solifugae does not support a sister-group relationship with Pseudoscorpiones (Arachnida). *J Arachnol.* 30(2):268–274.
- Arabi J, Judson MLI, Deharveng L, Lourenço WR, Cruaud C, Hassani A. 2012. Nucleotide Composition of CO1 sequences in Chelicerata (Arthropoda): detecting new mitogenomic rearrangements. *J Mol Evol.* 74(1–2):81–95.
- Arango CP. 2002. Morphological phylogenetics of the sea spiders (Arthropoda: Pycnogonida). *Org Divers Evol.* 2(2):107–125.
- Aria C, Caron J-B. 2019. A middle Cambrian arthropod with chelicerae and proto-book gills. *Nature* 573(7775):586–589.
- Ballesteros JA, Hormiga G. 2016. A new orthology assessment method for phylogenomic data: unrooted phylogenetic orthology. *Mol Biol Evol.* 33(8):2117–2134.
- Ballesteros JA, Santibáñez-López CE, Kováč L, Gavish-Regev E, Sharma PP. 2019. Ordered phylogenomic subsampling enables diagnosis of systematic errors in the placement of the enigmatic arachnid order Palpigradi. *Proc Biol Sci.* 286(1917):20192426.
- Ballesteros JA, Sharma PP. 2019. A critical appraisal of the placement of Xiphosura (Chelicerata) with account of known sources of phylogenetic error. *Syst Biol.* 68(6):896–917.
- Bicknell RDC, Lustri L, Brougham T. 2019. Revision of “*Bellinurus*” carteri (Chelicerata: Xiphosura) from the Late Devonian of Pennsylvania, USA. *Comptes Rendus Palevol.* 18(8):967–976.
- Bieler R, Mikkelsen PM, Collins TM, Glover EA, González VL, Graf DL, Harper EM, Healy J, Kawauchi GY, Sharma PP, et al. 2014. Investigating the Bivalve Tree of Life—an exemplar-based approach combining molecular and novel morphological characters. *Invert Systematics.* 28(1):32–115.
- Boxshall GA. 2004. The evolution of arthropod limbs. *Biol Rev Camb Philos Soc.* 79(2):253–300.
- Bruce HS, Patel NH. 2020. Knockout of crustacean leg patterning genes suggests that insect wings and body walls evolved from ancient leg segments. *Nat Ecol Evol.* 4(12):1703–1712.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden T. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.

- Clark-Hachtel CM, Tomoyasu Y. 2020. Two sets of candidate crustacean wing homologues and their implication for the origin of insect wings. *Nat Ecol Evol.* 4(12):1694–1702.
- Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, Macaranas J, Cassis G, Gray MR. 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Aust J Zool.* 46(5):419–437.
- Crisuolo A, Gribaldo S. 2010. BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol.* 10:210.
- Dabert M, Proctor H, Dabert J. 2016. Higher-level molecular phylogeny of the water mites (Acariformes: Prostigmata: Parasitengonina: Hydrachnidia). *Mol Phylogenet Evol.* 101:75–90.
- Dimitriou AC, Taiti S, Sfenthourakis S. 2019. Genetic evidence against monophyly of Oniscidea implies a need to revise scenarios for the origin of terrestrial isopods. *Sci Rep.* 9(1):18508.
- Dohle W. 2001. Are the insects terrestrial crustaceans? A discussion of some new facts and arguments and the proposal of the proper name 'Tetraconata' for the monophyletic unit Crustacea + Hexapoda. *Ann Soc Entomol Fr.* 37:85–103.
- Dunlop JA. 2010. Geological history and phylogeny of Chelicerata. *Arthropod Struct Dev.* 39(2–3):124–142.
- Dunlop JA, Kamenz C, Scholtz G. 2007. Reinterpreting the morphology of the Jurassic scorpion Liassoscorpionides. *Arthropod Struct Dev.* 36(2):245–252.
- Dunlop JA, Krüger J, Alberti G. 2012. The sejugal furrow in camel spiders and acariform mites. *Arach Mitteil.* 43:29–36.
- Edgecombe GD. 2020. Arthropod origins: integrating paleontological and molecular evidence. *Annu Rev Ecol Syst.* 51(1):1–25.
- Edgecombe GD, Wilson GDF, Colgan DJ, Gray MR, Cassis G. 2000. Arthropod cladistics: combined analysis of histone H3 and U2 snRNA sequences and morphology. *Cladistics* 16(2):155–203.
- Gainett G, González VL, Ballesteros JA, Setton EVW, Baker CM, Barolo Gargiulo L, Santibáñez-López CE, Coddington JA, Sharma PP. 2021. The genome of a daddy-long-legs (Opiliones) illuminates the evolution of arachnid appendages. *Proc Biol Sci.* 288(1956):20211168.
- Gainett G, Sharma PP. 2020. Genomic resources and toolkits for developmental study of whip spiders (Amblypygi) provide insights into arachnid genome evolution and antenniform leg patterning. *EvoDevo* 11:18.
- Garwood RJ, Dunlop JA. 2014. Three-dimensional reconstruction and the phylogeny of extinct chelicerate orders. *PeerJ.* 2:e641.
- Garwood RJ, Sharma PP, Dunlop JA, Giribet G. 2014. A Paleozoic stem group to mite harvestmen revealed through integration of phylogenetics and development. *Curr Biol.* 24(9):1017–1023.
- Giribet G, Edgecombe GD. 2019. The phylogeny and evolutionary history of arthropods. *Curr Biol.* 29(12):R592–R602.
- Giribet G, Edgecombe GD, Wheeler WC. 2001. Arthropod phylogeny based on eight molecular loci and morphology. *Nature* 413(6852):157–161.
- Giribet G, Edgecombe GD, Wheeler WC, Babbitt C. 2002. Phylogeny and systematic position of Opiliones: a combined analysis of chelicerate relationships using morphological and molecular data. *Cladistics* 18(1):5–70.
- Goloboff PA, Catalano SA. 2016. TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* 32(3):221–238.
- Harper A, Baudouin González L, Schöner A, Janssen R, Seiter M, Holzem M, Arif S, McGregor AP, Summer-Rooney L. 2021. Widespread retention of ohnologs in key developmental gene families following whole genome duplication in arachnospulmonates. *Genes Genomes Genet.* 11(12):jkab299. doi: 10.1093/g3journal/jkab299.
- Hernandez AM, Ryan JF. 2021. Six-state amino acid recoding is not an effective strategy to offset compositional heterogeneity and saturation in phylogenetic analyses. *Syst Biol.* 70(6):1200–1212.
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LE. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol.* 35(2):518–522.
- Howard RJ, Puttick MN, Edgecombe GD, Lozano-Fernández J. 2020. Arachnid monophyly: morphological, palaeontological and molecular support for a single terrestrialization within Chelicerata. *Arthropod Struct Dev.* 59:100997.
- Huang D, Hormiga G, Cai C, Su Y, Yin Z, Xia F, Giribet G. 2018. Origin of spiders and their spinning organs illuminated by mid-Cretaceous amber fossils. *Nat Ecol Evol.* 2(4):623–627.
- Kalyanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 14(6):587–589.
- Kapli P, Natsidis P, Leite DJ, Fursman M, Jeffrie N, Rahman IA, Philippe H, Copley RR, Telford MJ. 2021. Lack of support for Deuterostomia prompts reinterpretation of the first Bilateria. *Sci Adv.* 7:eabe2741.
- Kapli P, Telford MJ. 2020. Topology-dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha. *Sci Adv.* 6:eabc5162.
- King B. 2019. Which morphological characters are influential in a Bayesian phylogenetic analysis? Examples from the earliest osteichthyans. *Biol Lett.* 15(7):20190288.
- Klimov PB, O'Connor BM, Chetverikov PE, Bolton SJ, Pepato AR, Mortazavi AL, Tolstikov AV, Bauman GR, Ochoa R. 2018. Comprehensive phylogeny of acariform mites (Acariformes) provides insights on the origin of the four-legged mites (Eriophyoidea), a long branch. *Mol Phylogenet Evol.* 119:105–117.
- Klomp H, Lekveishvili M, Black WC. 4th. 2007. Phylogeny of parasitiform mites (Acari) based on rRNA. *Mol Phylogenet Evol.* 43(3):936–951.
- Kocot KM, Citarella MR, Moroz LL, Halanych KM. 2013. PhyloTreePruner: A phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. *Evol Bioinform.* 9:429–435.
- Kristensen RM. 2003. Comparative morphology: do the ultrastructural investigations of Loricifera and Tardigrada support the clade Ecdysozoa? In: Legakis A, Sfenthourakis S, Polymeni R, Thessalou-Legaki M, editors. The new panorama of animal evolution. Proceedings of the 18th Congress of Zoology. Sofia (Bulgaria): Pensoft Publishers, pp. 449–459.
- Kocot KM, Poustka AJ, Stöger I, Halanych KM, Schrödl M. 2020. New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships. *Sci Rep.* 10(1):101.
- Kulkarni S, Kallal RJ, Wood H, Dimitrov D, Giribet G, Hormiga G. 2021. Interrogating genomic-scale data to resolve recalcitrant nodes in the spider tree of life. *Mol Biol Evol.* 38(3):891–903.
- Lamsdell JC. 2016. Horseshoe crab phylogeny and independent colonizations of fresh water: ecological invasion as a driver for morphological innovation. *Palaeontology* 59(2):181–194.
- Lamsdell JC, McCoy VE, Perron-Feller OA, Hopkins MJ. 2020. Air breathing in an exceptionally preserved 340-million-year-old sea scorpion. *Curr Biol.* 30(21):4316–4321.
- Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst Biol.* 62(4):611–615.
- Lehmann T, Melzer RR. 2013. Looking like *Limulus*? – Retinula axons and visual neuropils of the median and lateral eyes of scorpions. *Front Zool.* 10(1):40.
- Lehmann T, Melzer RR. 2019a. The visual system of Thelyphonida (whip scorpions): support for Arachnospulmonata. *Arthropod Struct Dev.* 51:23–31.
- Lehmann T, Melzer RR. 2019b. Also looking like *Limulus*? – Retinula axons and visual neuropils of Amblypygi (whip spiders). *Front Zool.* 15:52.
- Li Y, Shen X-X, Evans B, Dunn CW, Rokas A. 2021. Rooting the animal tree of life. *Mol Biol Evol.* 38(10):4322–4333.
- Lopardo L, Michalik P, Hormiga G. 2021. Take a deep breath... The evolution of the respiratory system of symphytognathoid spiders (Araneae, Araneoidea). *Organ Div Evol.* doi: 10.1007/s13127-021-00524-w.
- Lozano-Fernández J, Tanner AR, Giacomelli M, Carton M, Vinther J, Edgecombe GD, Pisani D. 2019. Increasing species sampling in

- chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. *Nat Commun.* 10(1):2295.
- Mallatt JM, Garey JR, Shultz JW. 2004. Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. *Mol Phylogenet Evol.* 31(1):178–191.
- Mallatt JM, Giribet G. 2006. Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Mol Phylogenet Evol.* 40(3):772–794.
- Marlétaz F, Peijnenburg KTCA, Goto T, Satoh N, Rokhsar DS. 2019. A new spiralian phylogeny places the enigmatic arrow worms among gnathiferans. *Curr Biol.* 29(2):312–318.
- Masta SE, Longhorn SJ, Boore JL. 2009. Arachnid relationships based on mitochondrial genomes: asymmetric nucleotide and amino acid bias affects phylogenetic analyses. *Mol Phylogenet Evol.* 50(1):117–128.
- Mayer G. 2006. Origin and differentiation of nephridia in the Onychophora provide no support for the Articulata. *Zoomorphology* 125(1):1–12.
- Minh BQ, Hahn MW, Lanfear R. 2020. New methods to calculate concordance factors for phylogenomic datasets. *Mol Biol Evol.* 37(9):2727–2733.
- Mongiardino K, Thompson JR. 2021. A total-evidence dated phylogeny of Echinoidea combining phylogenomic and paleontological data. *Syst Biol.* 70(3):421–439.
- Mongiardino Koch N. 2021. Phylogenomic subsampling and the search for phylogenetically reliable loci. *Mol Biol Evol.* 38(9):4025–4038.
- Mongiardino Koch N, Parry LA. 2020. Death is in our side: paleontological data drastically modify phylogenetic hypotheses. *Syst Biol.* 69(6):1052–1067.
- Nakamura T, Gehrke AR, Lemberg J, Szymaszek J, Shubin NH. 2017. Digits and fin rays share common developmental histories. *Nature* 537(7619):225–228.
- Neumann JS, Desalle R, Narechania A, Schierwater B, Tessler M. 2021. Morphological characters can strongly influence early animal relationships inferred from phylogenomic data sets. *Syst Biol.* 70(2):360–375.
- Nguyen L-T, Schmidt HA, Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Noah KE, Hao J, Li L, Sun X, Foley B, Yang Q, Xia X. 2020. Major revisions in arthropod phylogeny through improved supermatrix, with support for two possible waves of land invasion by chelicerates. *Evol Bioinform Online.* 16:1176934320903735.
- Nolan ED, Santibáñez-López CE, Sharma PP. 2020. Developmental gene expression as a phylogenetic data class: support for the monophyly of Arachnospulmonata. *Dev Genes Evol.* 230(2):137–153.
- Oakley TH. 2003. On homology of arthropod compound eyes. *Integr Comp Biol.* 43(4):522–530.
- Ontano AZ, Gainett G, Aharon S, Ballesteros JA, Benavides LR, Corbett KF, Gavish-Regev E, Harvey MS, Monsma S, Santibáñez-López CE, et al. 2021. Taxonomic sampling and rare genomic changes overcome long-branch attraction in the phylogenetic placement of pseudoscorpions. *Mol Biol Evol.* 38(6):2446–2467.
- Ontano AZ, Steiner HG, Sharma PP. 2022. How many long branch orders occur in Chelicerata? Opposing effects of Palpigradi and Opilioacariformes on phylogenetic stability. *Mol Phylogenet Evol.* 168:107378.
- Pepato AR, da Rocha CEF, Dunlop JA. 2010. Phylogenetic position of the acariform mites: sensitivity to homology assessment under total evidence. *BMC Evol Biol.* 10(1):235.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol.* 67(5):901–904.
- Ramírez MJ, Magalhaes ILF, Derkarabetian S, Ledford J, Griswold CE, Wood HM, Hedin M. 2021. Sequence capture phylogenomics of true spiders reveals convergent evolution of respiratory systems. *Syst Biol.* 70(1):14–20.
- Redmond AK, McLysaght A. 2021. Evidence for sponges as sister to all other animals from partitioned phylogenomics with mixture models and recoding. *Nat Commun.* 12(1):1783.
- Richter S. 2002. The Tetraconata concept: hexapod-crustacean relationships and the phylogeny of Crustacea. *Org Div Evol.* 2(3):217–237.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 61(3):539–542.
- Schwager EE, Sharma PP, Clarke T, Leite DJ, Wierschin T, Pechmann M, Akiyama-Oda Y, Esposito L, Bechsgaard J, Bilde T, et al. 2017. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biol.* 15(1):62.
- Sharma PP, Ballesteros JA, Santibáñez-López CE. 2021. What is an “arachnid”? Consensus, concision, and confirmation bias in the phylogenetics of Chelicerata. *Diversity* 13(11):568.
- Sharma PP, Fernández R, Esposito LA, González-Santillán E, Monod L. 2015. Phylogenomic resolution of scorpions reveals multilevel discordance with morphological phylogenetic signal. *Proc Biol Sci.* 282(1804):20142953.
- Sharma PP, Kaluziak ST, Pérez-Porro AR, González VL, Hormiga G, Wheeler WC, Giribet G. 2014. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. *Mol Biol Evol.* 31(11):2963–2984.
- Shen X-X, Hittinger CT, Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nat Ecol Evol.* 1(5):126.
- Shen X-X, Salichos L, Rokas A. 2016. A genome-scale investigation of how sequence, function, and tree-based gene properties influence phylogenetic inference. *Genome Biol Evol.* 8(8):2565–2580.
- Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst Biol.* 51(3):492–508.
- Shultz JW. 1990. Evolutionary morphology and phylogeny of Arachnida. *Cladistics* 6(1):1–38.
- Shultz JW. 2001. Gross muscular anatomy of *Limulus polyphemus* (Chelicerata, Xiphosura) and its bearing on evolution in the Arachnida. *J Arachnol.* 29(3):283–303.
- Shultz JW. 2007. A phylogenetic analysis of the arachnid orders based on morphological characters. *Zool J Linn Soc.* 150(2):221–265.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–3212.
- Simion P, Philippe H, Baurain D, Jager M, Richter DJ, Di Franco A, Roure B, Satoh N, Quéinnec É, Ereskovsky A, et al. 2017. A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Curr Biol.* 27(7):958–967.
- Siveter DJ, Briggs DEG, Sutton MD, Legg DA. 2017. A new crustacean from the Herefordshire (Silurian) Lagerstätte, UK, and its significance in malacostracan evolution. *Proc R Soc Lond B.* 284:20170279.
- Smith FW, Boothby TC, Giovannini I, Rebecchi L, Jockusch EL, Goldstein B. 2016. The compact body plan of tardigrades evolved by the loss of a large body region. *Curr Biol.* 26(2):224–229.
- Smythe AB, Holovachov O, Kocot KM. 2019. Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evol Biol.* 19(1):121.
- Steel M, Sanderson MJ. 2010. Characterizing phylogenetically decisive taxon coverage. *Appl Math Lett.* 23(1):82–86.
- Struck TH, Golombek A, Weigert A, Franke FA, Westheide W, Purschke G, Bleidorn C, Halanich KM. 2015. The evolution of annelids reveals two adaptive routes to the interstitial realm. *Curr Biol.* 25(15):1993–1999.
- Susko E, Roger AJ. 2007. On reduced amino acid alphabets for phylogenetic inference. *Mol Biol Evol.* 24(9):2139–2150.

- Wang H-C, Susko E, Roger AJ. 2019. The relative importance of modeling site pattern heterogeneity versus partition-wise heterotachy in phylogenomic inference. *Syst Biol.* 68(6):1003–1019.
- Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol.* 35(3):543–548.
- Watson-Zink VM. 2021. Making the grade: physiological adaptations to terrestrial environments in decapod crabs. *Arthropod Struct Dev.* 64:101089.
- Wheeler WC, Hayashi CY. 1998. The phylogeny of the extant chelicerate orders. *Cladistics* 14(2):173–192.
- Whelan NV, Halaných KM. 2017. Who let the CAT out of the bag? Accurately dealing with substitutional heterogeneity in phylogenomic analyses. *Syst Biol.* 66(2):232–255.
- Wickham H. 2016. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag. ISBN 978-3-319-24277-4; Available from: <https://ggplot2.tidyverse.org>.
- Wolfe JM. 2017. Metamorphosis is ancestral for crown euarthropods, and evolved in the Cambrian or earlier. *Integr Comp Biol.* 57(3):499–509.
- Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19(Suppl 6):153.