WILEY

Cladistics

Cladistics (2023) 1–15

doi: 10.1111/cla.12551

Congruence between ultraconserved element-based matrices and phylotranscriptomic datasets in the scorpion Tree of Life

Carlos E. Santibáñez-López*[a], Andrés A. Ojanguren-Affilastro[b], Matthew R. Graham[c] and Prashant P. Sharma[d]

[a]*Department of Biology, Western Connecticut State University, Danbury, CT 06810, USA;* [b]*División Aracnología, Museo Argentino de Ciencias Naturales, Buenos Aires C1405DJR, Argentina;* [c]*Department of Biology, Eastern Connecticut State University, Willimantic, CT 06226, USA;* [d]*Department of Integrative Biology, University of Wisconsin—Madison, Madison, WI 53706, USA*

Accepted 6 June 2023

Abstract

Scorpions are ancient and historically renowned for their potent venom. Traditionally, the systematics of this group of arthropods was supported by morphological characters, until recent phylogenomic analyses (using RNAseq data) revealed most of the higher-level taxa to be non-monophyletic. While these phylogenomic hypotheses are stable for almost all lineages, some nodes have been hard to resolve due to minimal taxonomic sampling (e.g. family Chactidae). In the same line, it has been shown that some nodes in the Arachnid Tree of Life show disagreement between hypotheses generated using transcritptomes and other genomic sources such as the ultraconserved elements (UCEs). Here, we compared the phylogenetic signal of transcriptomes vs. UCEs by retrieving UCEs from new and previously published scorpion transcriptomes and genomes, and reconstructed phylogenies using both datasets independently. We reexamined the monophyly and phylogenetic placement of Chactidae, sampling an additional chactid species using both datasets. Our results showed that both sets of genome-scale datasets recovered highly similar topologies, with Chactidae rendered paraphyletic owing to the placement of *Nullibrotheas allenii*. As a first step toward redressing the systematics of Chactidae, we establish the family Anuroctonidae (new family) to accommodate the genus *Anuroctonus*.

Introduction

Scorpions constitute a charismatic lineage of arthropods that probably originated in the Ordovician (~470 Myr; Santibáñez-López et al., 2022), and today comprise nearly 2750 described species (Rein, 2022). These animals have successfully survived multiple mass extinctions and colonized different tropical, temperate and cold habitats. While a basally branching placement of scorpions within Arachnida (often with Opiliones) was generally recovered by cladistic analyses of morphology (Weygoldt and Paulus, 1979; Wheeler and Hayashi, 1998; Giribet et al., 2002; Shultz, 2007), phylogenomic analyses (Sharma et al., 2014a; Ballesteros and Sharma, 2019; Ballesteros et al., 2022) and developmental data (Sharma et al., 2014b; Nolan et al., 2020; Ontano et al., 2021) support scorpion placement as part of Arachnopulmonata (the sister group of Tetrapulmonata, forming a clade of taxa that ancestrally bore book lungs). More recently, rare genomic changes recovered Scorpiones as the sister group of Pseudoscorpiones (forming the clade Panscorpiones), a result consistent with both phylogenomic analyses based on dense taxonomic sampling and the systemic paralogy of genes and miRNAs resulting from shared genome duplication (Ontano et al., 2021; Ballesteros et al., 2022).

Closely paralleling the placement of scorpions among arachnids, the internal phylogeny of the group has advanced significantly in the past 10 years, but is often at odds with morphological hypotheses. Historically, scorpion classification and phylogenetic

*Corresponding author:
E-mail address: santibanezlopezc@wcsu.edu

relationships among scorpion groups were based primarily on a subset of morphological character systems, namely, trichobothrial patterns, sternum shape and the anatomy of the hemispermatophore (Sissom, 1990; Soleglad and Fet, 2003; Prendini and Wheeler, 2005; Monod et al., 2017). Tests of relationships using Sanger data were limited to analyses at the level of family, genus or regional fauna (e.g. Fet et al., 2003; Prendini et al., 2003; González-Santillán and Prendini, 2015; Santibáñez-López et al., 2017a; Loria et al., 2022; Parmakelis et al., 2022; Štundlová et al., 2022). Two competing hypotheses classified scorpions into either four parvorders (Soleglad and Fet, 2003) or 18 families (i.e. Prendini and Wheeler, 2005), with both systems placing different genera among families based on alternative putative synapomorphies. Nevertheless, basic elements of these classification schemes were largely congruent; as examples, the southeast Asian family Chaerilidae was held to be basally branching within Iurida, and Bothriuridae was understood to be closely related to (or part of) Scorpionoidea, the group that exhibits katoikogenic development (Soleglad and Fet, 2003; Coddington et al., 2004).

The first scorpion phylogenomic analysis, based on RNAseq data for 25 exemplars, revealed discordance with traditional morphological systematics, with Chaerilidae invariably recovered as closely related to Buthidae and Pseudochactidae, and Bothriuridae recovered as distantly related to Scorpionoidea (Sharma et al., 2015). Various superfamilies were recovered as non-monophyletic, suggesting broader discrepancies with the traditional classification. This work emended the higher-level classification of scorpions to comprise two parvorders: Buthida Soleglad and Fet, 2003 (with three superfamilies) and Iurida Soleglad and Fet, 2003 (comprising four superfamilies). In the wake of these outcomes, scorpion systematics witnessed rapid proliferation of genomic data across its taxonomic breadth, with the goal of revising the relationships of the group. Recent phylogenomic analyses have recovered some traditional higher-level relationships with support (i.e. among superfamilies), but others were non-monophyletic (Santibáñez-López et al., 2018, 2019a, 2020, 2022; Sharma et al., 2018; Fig. 1). Ten superfamilies are currently recognized and a robust backbone phylogeny now exists for scorpion relationships (Santibáñez-López et al., 2022, Table 1), with better resolution for families like Vaejovidae, Buthidae and Iuridae (Santibáñez-López et al., 2018, 2022; Parmakelis et al., 2022; Štundlová et al., 2022).

Among these superfamilies, phylogenetic relationships within Chactoidea remain obscure. Chactoidea, as defined by Soleglad and Fet (2003), is comprised of four families: Chactidae, Euscorpiidae, Superstitioniidae and Vaejovidae. Although Prendini and Wheeler (2005) rejected Soleglad and Fet's (2003) classification, the first scorpion phylogenomic analysis revalidated Chactoidea and assigned eight families: Caraboctonidae, Chactidae, Euscorpiidae, Scorpiopidae, Superstitioniidae, Troglotayosicidae, Typhochactidae and Vaejovidae. Recent phylogenomic analyses later restricted Chactoidea to three families: Chactidae, Euscorpiidae and Scorpiopidae, as Vaejovidae was transferred to its resurrected superfamily Vaejovoidea (Santibáñez-López et al., 2019a), followed by the restoration of superfamilies Caraboctonoidea and Hadruroidea to accommodate Caraboctonidae and Hadruridae (Santibáñez-López et al., 2020). However, the phylogenetic position and composition of Chactidae *sensu stricto* has not been tested, as only one exemplar of the family Chactidae (*Brotheas granulatus* Simon 1877) has been included in these analyses. Scorpions of this family, which harbours 205 species, are distributed in South America, with a single species found in North America [*Nullibrotheas allenii* (Wood, 1863) in Baja California, Mexico; Santibáñez-López et al., 2019a].

It has been shown that some nodes in the Tree of Life are difficult to resolve regardless of the amount of input data (e.g. Philippe et al., 2011; Alda et al., 2019). Moreover, the choice of genomic markers for phylogenetic inference, which is influenced by sequencing costs and tissue availability, is well understood to impact phylogenetic outcomes (Karin et al., 2020; Alda et al., 2021). While shotgun sequencing of transcriptomes offers numerous advantages for phylogenomic study (such as the ability to test orthology and design matrices suited to specific phylogenetic investigations), this strategy is constrained by high technical demands for tissue preservation and quality. For field collection of species endemic to challenging environments (e.g. deep caves and deserts), obtaining high-quality RNA may be difficult or unfeasible. A promising workaround is the use of ultraconserved elements (UCEs), which are robust to DNA degradation and can be used with dried or ethanol-preserved specimens (e.g. Blaimer et al., 2016; Derkarabetian et al., 2019). Phylogenomic hypotheses generated using UCEs and exons (transcriptomes) generally agree, but several studies have shown disagreement between these topologies, especially at recalcitrant nodes with low phylogenetic signal (e.g. Bossert et al., 2019; Kulkarni et al., 2020; Alda et al., 2021). Within arachnids, a notable example is the case of Symphytognathoidea, a clade of miniaturized spiders that was proposed on the basis of morphological data. While not recovered by phylotranscriptomic analyses, this group was recovered as monophyletic with strong support by UCE datasets (Kulkarni et al., 2020). These results were interpreted to mean that the phylogenetic signal in UCE datasets may be more congruent with morphological data,

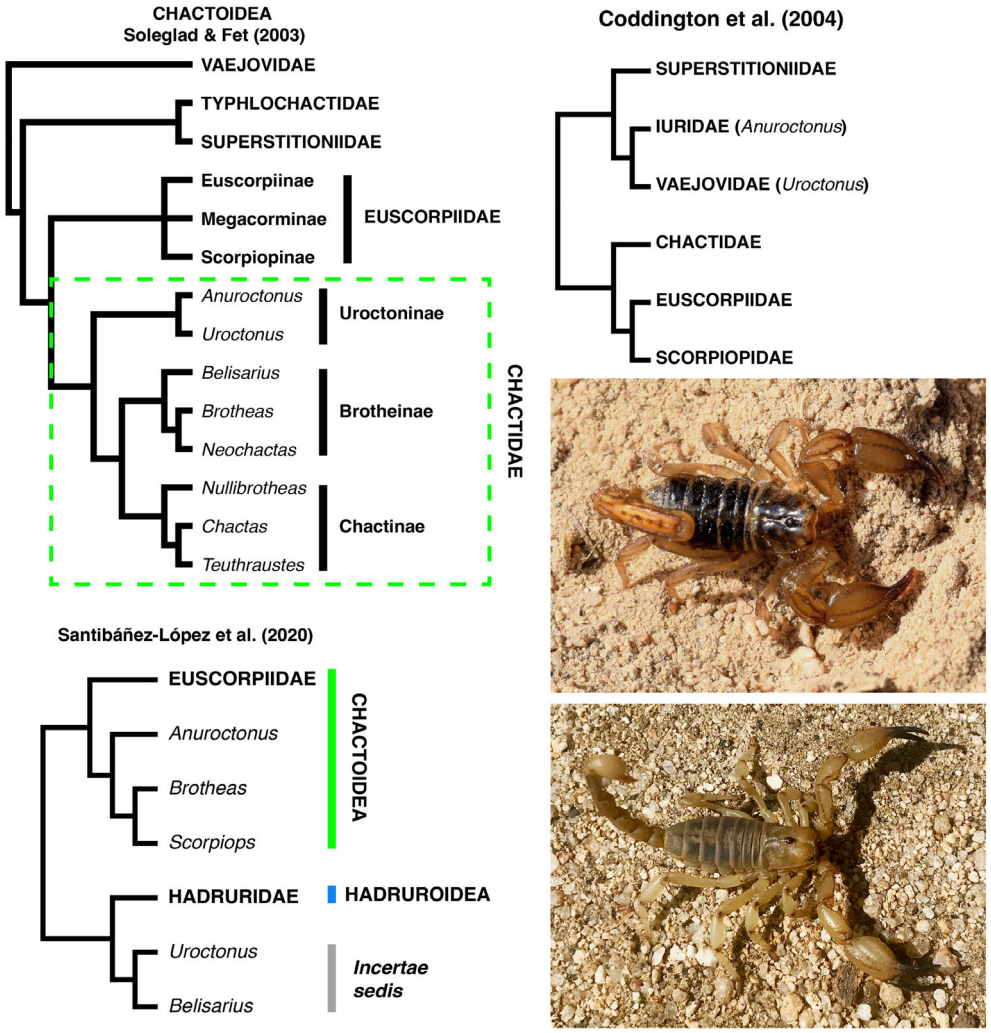

Fig. 1. Historical hypotheses of chactoid scorpion relationships based on morphology (Soleglad and Fet, 2003; Coddington et al., 2004) and genomic datasets (Santibáñez-López et al., 2020). Photographs of live habitus of *Nullibrotheas allenii* (top) and *Anuroctonus pococki bajae* (bottom). Photographs by C. Santibáñez-López.

suggesting superior performance in comparison with phylotranscriptomic matrices.

This inference bears directly on the higher-level systematics of scorpions, whose broad-scale revisions have been based entirely upon mRNA sequencing. It is therefore necessary to interrogate the phylogenetic signal of different genomic datasets to test their power and congruence in resolving scorpion relationships. Compared with other arachnids (i.e. spiders and harvestmen), tests of UCE performance are few within scorpions. Starrett et al. (2017) designed a probe set to explore the utility of these genomic markers and tested these for five scorpion species. Their UCE topology disagreed with the exon-based analysis of Sharma et al. (2015) for three out of four higher-level (i.e. above genus-level) nodes. While limited in scope, these results suggest that scorpion phylogeny may parallel the case of Symphytognathoidea, with different datasets recovering incongruent topologies, an outcome that could heavily impact the recent reclassification of scorpions (Sharma et al., 2015; Santibáñez-López et al., 2018, 2019a, 2020).

To compare the phylogenetic signal of transcriptomes vs. UCEs, we recovered UCEs from published and newly sequenced transcriptomes and genomes (Table 2), and reconstructed phylogenies using both datasets independently. We separately tested the resolution of the traditional Chactoidea and Chactidae using phylogenomic analyses of both data types. The major addition to this analysis was the sampling of the chactid genus *Nullibrotheas* Williams, 1974, which we anticipated to be recovered as the sister group to *Brotheas* C.L. Koch, 1837 and to stabilize this node of the phylogeny. Here, we show that both sets of

Table 1
High-level classification of scorpions proposed by Sharma et al. (2015) and modified by Santibáñez-López et al. (2019a,b, 2020)

Order Scorpiones Koch 1837
Suborder Neoscorpionina Thorell & Lindström 1885
Infraorder Orthosterni Pocock 1911
Parvorder Buthida Soleglad and Fet 2003
Superfamily Buthoidea Koch 1837
Family Buthidae Koch 1837
Superfamily Chaeriloidea Pocock 1893
Family Chaerilidae Pocock 1893
Superfamily Pseudochactoidea Gromov 1998
Family Pseudochactidae Gromov 1998
Parvorder Iurida Soleglad and Fet 2003
Superfamily Bothriuroidea Simon 1880
Family Bothriuroidea Simon 1880
Superfamily Caraboctonoidea Kraepelin 1905
Family Caraboctonidae Kraepelin 1905
Family Superstitioniidae Stanhke 1940
Superfamily Chactoidea Pocock 1893
Family Chactidae Pocock 1893
Family Euscorpiidae Laurie 1896
Family Scorpiopidae Kraepelin 1905
Superfamily Iuroidea Thorell 1876
Family Iuridae Thorell 1876
Superfamily Hadruroidea Stahnke 1974
Family Hadruridae Stahnke 1974
Superfamily Scorpionoidea Latreille 1802
Family Diplocentridae Karsch 1880
Family Hemiscorpiidae Pocock 1893
Family Hormuridae Laurie 1896
Family Rugodentidae Bastawade et al. 2005
Family Scorpionidae Latreille 1802
Family Urodacidae Pocock 1893
Superfamily Vaejovoidea Thorell 1876
Family Vaejovidae Thorell 1876
Incertae sedis
Family Belisariidae Lourenço 1998
Family Heteroscorpionidae Kraepelin 1905
Family Troglotayoscidae Lourenço 1998
Family Typhlochactidae Mitchell 1971

genomic markers recovered highly similar topologies, with all current superfamilies and other major relationships recovered as monophyletic. However, Chactidae was invariably rendered paraphyletic owing to the placement of *Nullibrotheas* in all analyses. Therefore, we undertake taxonomic actions to redress the systematics of Chactidae.

Methods

Taxon sampling

Specimens were collected with the aid of ultraviolet lamps at night from two localities in Baja California Sur (Mexico), one locality in Argentina and one in Chile (Table 2). Scorpions were dissected into RNA*later* solution (Ambion), and their brains, legs and telsons were removed for sequencing. Total RNA was extracted and sequenced, followed by transcriptome assembly, using previously described protocols (e.g. Sharma et al., 2015; Santibáñez-López et al., 2022). Transcriptomes previously published by us were included for outgroup sampling (Sharma et al., 2014a, 2015, 2018; Table 3). New terminals in this analysis consisted of the chactid *Nullibrotheas alleni* and two species of the bothriurid genus *Urophonius* Pocock, 1893 [*U. brachycentrus* (Thorell, 1876) and *U. granulatus* Pocock, 1898].

Matrix assembly and analysis

Orthologous loci were drawn from Markov Cluster Algorithm clustering of 424 loci computed from our previous analysis of scorpions (Santibáñez-López et al., 2022). Untrimmed alignments were used to produce a hidden Markov profile using *hmmerbuild* from hmmer v. 3.2.1 (Mistry et al., 2013). Our newly sequenced transcriptomes then were used as query to search (*hmmersearch*) for matches against the collection of profiles, with an expectation threshold of $e < 10^{-20}$; for cases with more than one hit per locus, the sequence with the best score was preferred. Then, each corresponding sequence was appended to the locus FASTA file aggregating the putative orthologs found in each species. Then, one phylotranscriptomic matrix (Matrix AAm1, 424 partitions) was assembled.

For assembly of UCE matrices, the FASTA files of transcriptomes were converted to a 2-bit format using faToTwoBit (Kent, 2002), and then recovered using PHYLUCE v.1.7 (Faircloth, 2016). The resulting FASTA files were then matched to the sequences from the Spider2Kv1 probe (Kulkarni et al., 2020). Nucleotide sequences from UCEs were assembled, aligned using MAFFT v.7.4 (--auto --anysymbol --quiet; Katoh and Standley, 2013) and trimmed using

Table 2
Localities of newly sequenced scorpions

Species	Locality/region of origin	Latitude	Longitude	Date	Collector	Sequence Read Archive BioProject
Urophonius brachycentrus	10 km south of Viedma, Río Negro Providence, Argentina	−40°53′52.31″	−72°38′33.97″	August 2021	A. Ojanguren-Affilastro, H. Iuri, L. Piacentini	PRJNA922548
Urophonius granulatus	Entrance to Torres del Paine National Park, Magallanes Region, Chile	−51°33′46.34″	−72°38′33.97″	March 2019	A. Ojanguren-Affilastro, J. Pizarro-Araya, F. Alfaro-Kong, J. Calderón, A. Castex	PRJNA922548
Nullibrotheas allenii	Mexico: Near el Pescadero, La Paz, Baja California Sur, Mexico	23°21′57.06″	−110°5′52.65″	August 2019	M. Graham, R. Jones, J. Idjadi, C. Santibáñez-López	PRJNA922548

Table 3
Revised higher level classification of extant scorpions. Taxa of questionable monophyly are indicated with asterisks. Taxa of unknown phylogenetic position based on phylogenomic data are indicated with question marks

Order Scorpiones Koch 1837
Suborder Neoscorpionina Thorell & Lindström 1885
Infraorder Orthosterni Pocock 1911
Parvorder Buthida Soleglad and Fet 2003
Superfamily Buthoidea Koch 1837
Family Buthidae Koch 1837
Family Ananteridae Kraepelin 1908
Superfamily Chaeriloidea Pocock 1893
Family Chaerilidae Pocock 1893
Superfamily Pseudochactoidea Gromov 1998
Family Pseudochactidae Gromov 1998
Parvorder Iurida Soleglad and Fet 2003
Superfamily Bothriuroidea Simon 1880
Family Bothriuroidea Simon 1880
Superfamily Caraboctonoidea Kraepelin 1905
Family Caraboctonidae Kraepelin 1905
Family Superstitioniidae Stahnke 1940
Superfamily Chactoidea Pocock 1893
Family Anuroctonidae Santibáñez-López, Ojanguren-Affilastro, Graham & Sharma new family
Family Chactidae Pocock 1893
Family Euscorpiidae Laurie 1896
Family Scorpiopidae Kraepelin 1905
Superfamily Iuroidea Thorell 1876
Family Iuridae Thorell 1876
Superfamily Hadruroidea Stahnke 1974
Family Hadruridae Stahnke 1974
Superfamily Scorpionoidea Latreille 1802
Family Diplocentridae Karsch 1880
Family Hemiscorpiidae Pocock 1893
Family *Hormuridae Laurie 1896
Family Rugodentidae Bastawade et al. 2005
Family *Scorpionidae Latreille 1802
Family Urodacidae Pocock 1893
Superfamily Vaejovoidea Thorell 1876
Family Vaejovidae Thorell 1876
Incertae sedis
Family *Belisariidae Lourenço 1998
Family? Heteroscorpionidae Kraepelin 1905
Family Troglotayosicidae Lourenço 1998
Family? Typhlochactidae Mitchell 1971

trimAl v.1.2 (-fasta -gappyout; Capella-Gutiérrez et al., 2009) to obtain 1950 selected loci. Three matrices were assembled with minimum taxon occupancy thresholds: Matrix 1 (UCEm1 with at least 57 species per locus), Matrix 2 (UCEm2 with at least 64 species per locus) and Matrix 3 (UCEm3 with at least 77 species per locus). Phylogenetic analysis, model selection and nodal support of each locus were performed using the procedure indicated below. In contrast to the transcriptome-based dataset, UCE concatenated matrices were analysed as one partition using ModelFinder constraining the search to the GTR model only. Parsimony analyses of matrices AAm1 and UCEm3 were conducted using TNT v. 1.6 (Goloboff and Catalano, 2016) with 100 jackknife replicates.

Gene trees were constructed using IQ-TREE v. 2.0.6 (Minh et al., 2020a) and ModelFinder Plus (Kalyaanamoorthy et al., 2017), for automated model fitting, with nodal support estimated using ultrafast bootstrapping (Hoang et al., 2018), and gene and site concordance factors (gCf and sCF; Minh et al., 2020b). Alignments and/or gene trees were visually inspected for chimeric transcripts or paralogous sequences as some mismatches have been observed before (Santibáñez-López et al., 2022). The phylogenetic inference of matrix AAm1 was computed with IQ-TREE, implementing the best-fitting amino acid substitution model per partition, and nodal support using ultrafast bootstrapping. To infer shallow-level relationships in the presence of incomplete lineage sorting, species tree estimation was also performed using the gene trees and the coalescent summary method implemented in ASTRAL III (Mirarab and Warnow, 2015).

Gene properties and tree metrics

To explore information content and identify potential biases in our matrices, we analysed our four datasets using Phykit v. 1.5.0 (Steenwyk et al., 2021) and the R script *genesortR* (Mongiardino, 2021). Metrics compared across loci in our matrices consisted of the number of sites per locus, the number of sites without gaps, the number of parsimony informative sites per locus, the mean long branch score, the average patristic distances, levels of saturation, root-to-tip variance, compositional heterogeneity (for our amino acid set only), Robinson–Foulds similarity and average bootstrap support. To assess topological differences between trees, we used the information metric of Kendall and Colijn (2016) as implemented in the R package TreeSpace (Jombart et al., 2017). To show the variation between our recovered topologies, we projected the metrics of Kendall and Colijn (2016) onto a multidimensional scaling plot using TreeSpace. Since our trees had similar structures, we summarize the content of the UCE analyses into a single 'consensus tree' using the function of *medTree* (TreeSpace). The information content of the individual gene alignments (424 exon loci, and 531 UCE loci) and the four matrices (AAm1, UCEm1, 2 and 3) were evaluated using the four-cluster likelihood mapping (Strimmer and Von Haeseler, 1997) as quartets in IQ-TREE (*–lmap* All). We tested the position of *Nullibrotheas* with respect to three taxa: *Brotheas*, *Scorpiops* and Euscorpiidae.

Results

Phylogenomic analyses

One amino acid matrix (AAm1, with 424 genes and 114 315 amino acids; from Santibáñez-López et al., 2022) and three UCE phylogenomic matrices spanning 149–531 loci (117 649–357 956 nucleotides) were constructed. Assessment of phylogenomic biases included mean long branch scores and the number of parsimony informative sites, as shown in Figs S1–S5.

Maximum likelihood (ML) analysis of the concatenated amino acid dataset recovered, with maximal nodal support, the monophyly of scorpions, the established basal split between the two parvorders (Buthida + Iurida) and the majority of the relationships among and monophyly of the superfamilies (Fig. 2a). Superfamily Caraboctonoidea was not recovered as monophyletic using these loci, as previously shown (Santibáñez-López et al., 2022). Relationships between groups within Buthidae were consistent with those reported previously, with *Lychas variatus* (Thorell, 1876) as the sister group of the remaining buthids. Within the superfamily Bothriuroidea, *Cercophonius*

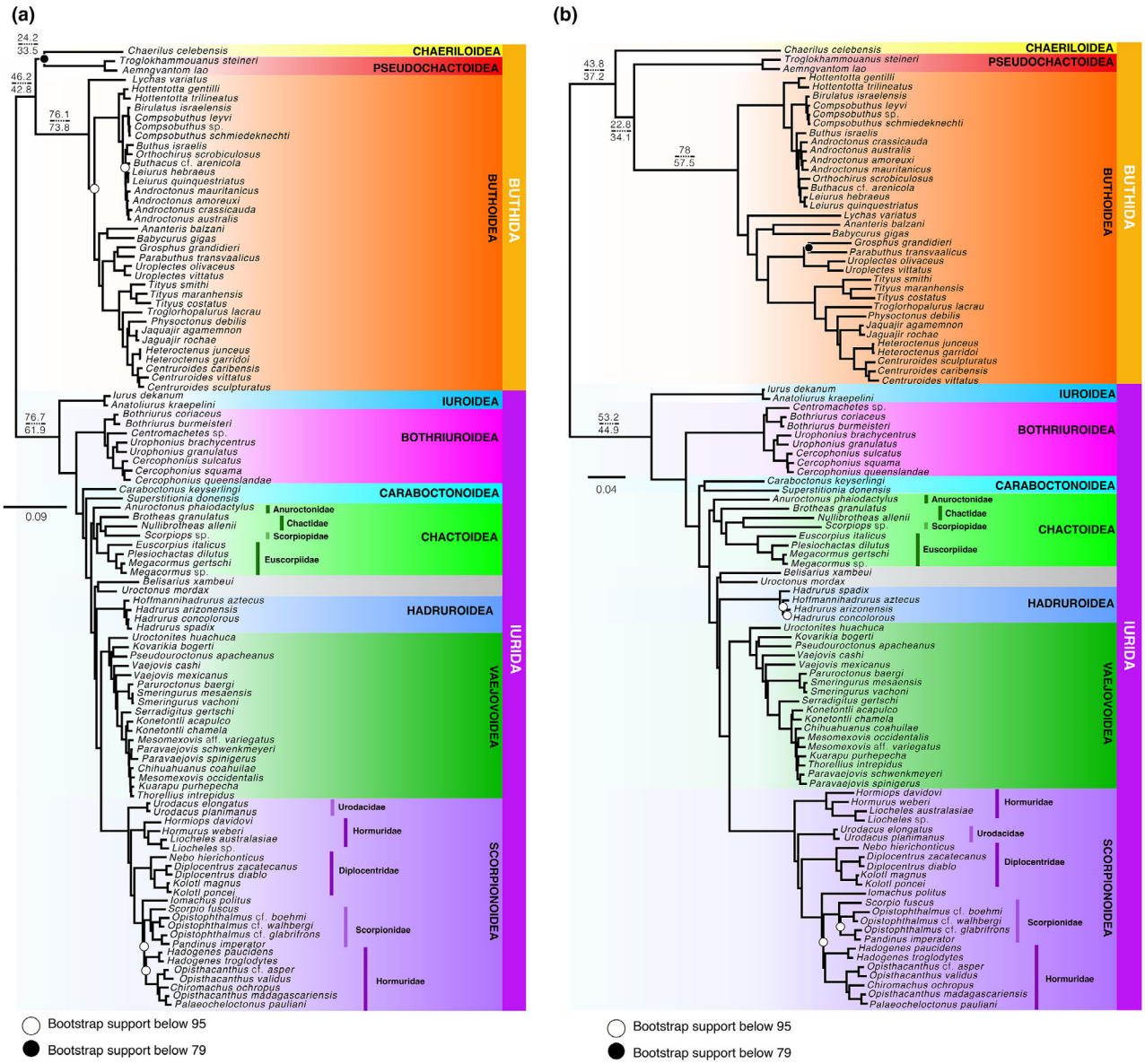

Fig. 2. (a) Maximum likelihood (ML) tree topology recovered from the analysis of 424 amino acid loci (AAm1; log Likelihood = −3 317 472.22). (b) ML tree topology recovered from the analysis of 531 UCE loci (UCEm1; lnL = −5 782 225.20). Site (top) and gene (bottom) concordance factors are indicated near the selected branches on both topologies.

Peters, 1861 [*C. squama* (Gervais 1843), *C. queenslandae* Acosta 1990, *C. sulcatus* Kraepelin 1908] was recovered as the sister taxon to *Urophonius* [*U. brachycentrus* (Thorell 1876) and *U. granulatus* Pocock 1898] and not to *Centromachetes* Lonnberg 1897, as previously inferred (Sharma et al., 2018). Lastly, *N. allenii* was recovered as the sister taxon to *Scorpiops* sp. (Scorpiopidae) across analyses, rendering Chactidae (represented by *B. granulatus*) paraphyletic (Fig. 2a). All other relationships agreed with previous topologies.

Maximum likelihood analyses of the concatenated UCE datasets recovered, with maximal nodal support, the monophyly of the two parvorders. Matrices UCEm1 and UCEm2 (531 and 394 loci respectively) recovered Chaerilidae as a sister taxon to the clade comprising Buthidae and Pseudochactidae (Fig. 2b, Fig. S6). In contrast, the ML analysis of UCEm3 (149 loci) recovered Chaerilidae as a sister taxon to Pseudochactidae, as also seen in our AAm1 topology (Fig. S7). Within Buthidae, all three matrices recovered the monophyly of the "*Buthus*", "*Uroplectes*" and

"*Tityus*" groups, and the paraphyly of the "*Ananteris*" group (owing to the exclusion of *L. variatus*) as shown previously (Santibáñez-López et al., 2022; Štundlová et al., 2022). Unlike in our AA1m topology, the three UCE matrices recovered the monophyly of Caraboctonoidea (*Caraboctonus* Pocock 1893 + *Superstitionia* Stahnke 1940). Further, as in our AAm1 topology, *N. allenii* was recovered as the sister taxon to the genus *Scorpiops* Peters 1861 (*Scorpiops* sp.) in all UCE datasets, rendering Chactidae non–monophyletic. All analyses (AA and UCEs) recovered the polyphyly of Hormuridae (within Scorpionoidea), and the relationship between *Uroctonus mordax* Thorell 1876 and *Belisarius xambeui* Simon 1879 (Fig. 2, Figs S6 and S7). Lastly, parsimony analyses of AAm1 and UCEm3 recovered highly similar topologies to those mentioned above (Fig. 3).

Species tree analyses of the amino acid loci (424) recovered the monophyly of both parvorders, the monophyly of the superfamilies and similar relationships within superfamilies as in our ML analysis of the AAm1 matrix (Fig. S8). *Nullibrotheas allenii* was recovered as a sister taxon to *Scorpiops* sp., but unlike in the ML topologies, *B. granulatus* was recovered as a sister taxon to Euscorpiidae (genera *Euscorpius* Thorell 1876, *Plesiochactas* Pocock 1900 and *Megacormus* Karsch 1881), and not as the sister group to *N. allenii* and *Scorpiops* sp. Similarly, species tree analyses of the UCE loci (149, 394, and 531) recovered similar topologies to those mentioned before with the following exceptions. The Astral trees from the 531 and 394 loci trees recovered *N. allenii* as a sister taxon to *Scorpiops* sp., and *B. granulatus* as the sister taxon to Euscorpiidae (Figs S9 and S10). In contrast, the Astral tree from the 149 loci trees recovered *B. granulatus* as the sister taxa to the clade comprising *N. allenii* and *Scorpiops* sp. All analyses recovered the relationship between *U. mordax* and *B. xambeui*, but the phylogenetic position of this clade changes in each topology (Fig. S11).

While the ML and species coalescent analyses of both datasets (AA and UCEs) showed a large degree of congruence (Fig. 4a–c), the multidimensional scaling of topological tree space of phylogenetic analyses recovered three tree clusters (Fig. 4c). All trees recovered from UCEs (ML and Astral) are clustered together, suggesting that they are more similar to each other whereas the AA Astral topology is the most different of all trees. Since all UCE trees were very similar, they were summarized into a "consensus" tree using *treespace*, and then compared with the AAm1 tree (Fig. 4b,c).

To assess gene overlapping across the data types, we retrieved the longest sequence from each of the 660 gene partitions from Santibáñez-López et al. (2022), created a database and performed translated BLAST searches (tBLASTn) with the longest sequence from each of the 531 UCE loci. Of 531 UCEs, 44 loci retrieved hits with 90–100% identity and *E*-values lower than $1e^{-102}$, suggesting little overlap between these two types of datasets (Fig. 4d).

Quartet likelihood mapping

The quartet likelihood mapping to test the phylogenetic position of *Nullibrotheas* as either sister taxa to *Scorpiops*, *Brotheas* or Euscorpiidae using all concatenated matrices (AAm1, UCEm1-3) recovered the quartet (*Nullibrotheas* + *Scorpiops*), which was consistent with the ML and Astral results, with 100% frequency (Fig. 5a,b,e). Sampling of quartets across 424 individual exons and 531 individual UCE loci supported this quartet with <50% (38% and 20% respectively; Fig. 5c,d,f,g), 38% (AA) and 46% (UCEs) non-informative quartets (Fig. 5).

Discussion

Phylogenetic position of Nullibrotheas allenii *and the status of* Anuroctonus *Pocock, 1893*

Previous scorpion phylogenomic analyses rejected the monophyly of Chactidae based on the position of the Nearctic species *Anuroctonus phaiodactylus* (Wood, 1863) and *U. mordax*, both members of subfamily Uroctoninae (*sensu* Soleglad and Fet, 2003; but see Prendini and Wheeler, 2005). In recent phylogenomic analyses, *U. mordax* has been recovered as a sister taxon to *B. xambeui* as the superfamily *incertae sedis*, and *A. phaiodactylus* has been consistently recovered as the sister taxon to all species within Chactoidea (e.g. Santibáñez-López et al., 2020). These results suggested that Chactidae could be a diphyletic lineage comprising two subfamilies (Brotheinae and Chactinae). Under this scenario, only one representative of the family Chactidae has been included in previous phylogenomic analyses: the Neotropical species *Brotheas granulatus* (a member of the subfamily Brotheinae). Thus, the sampling of additional Chactidae was necessary to revise the Chactidae and assess the placement of *Anuroctonus* Pocock, 1893.

Contrary to our expectations, our topologies consistently recovered *N. allenii* as the sister taxon to *Scorpiops* sp., recapitulating the recurring result that the morphology-based classification of scorpions does not agree with genome-scale phylogenies. Our results refute the inclusion of genus *Nullibrotheas* and *Brotheas* within a monophyletic Chactidae (Fig. 5a–g), as previously suggested (Soleglad and Fet, 2003; Prendini and Wheeler, 2005). This placement suggests that the subfamily Brotheinae probably merits elevation to family rank in future revisions of Chactidae.

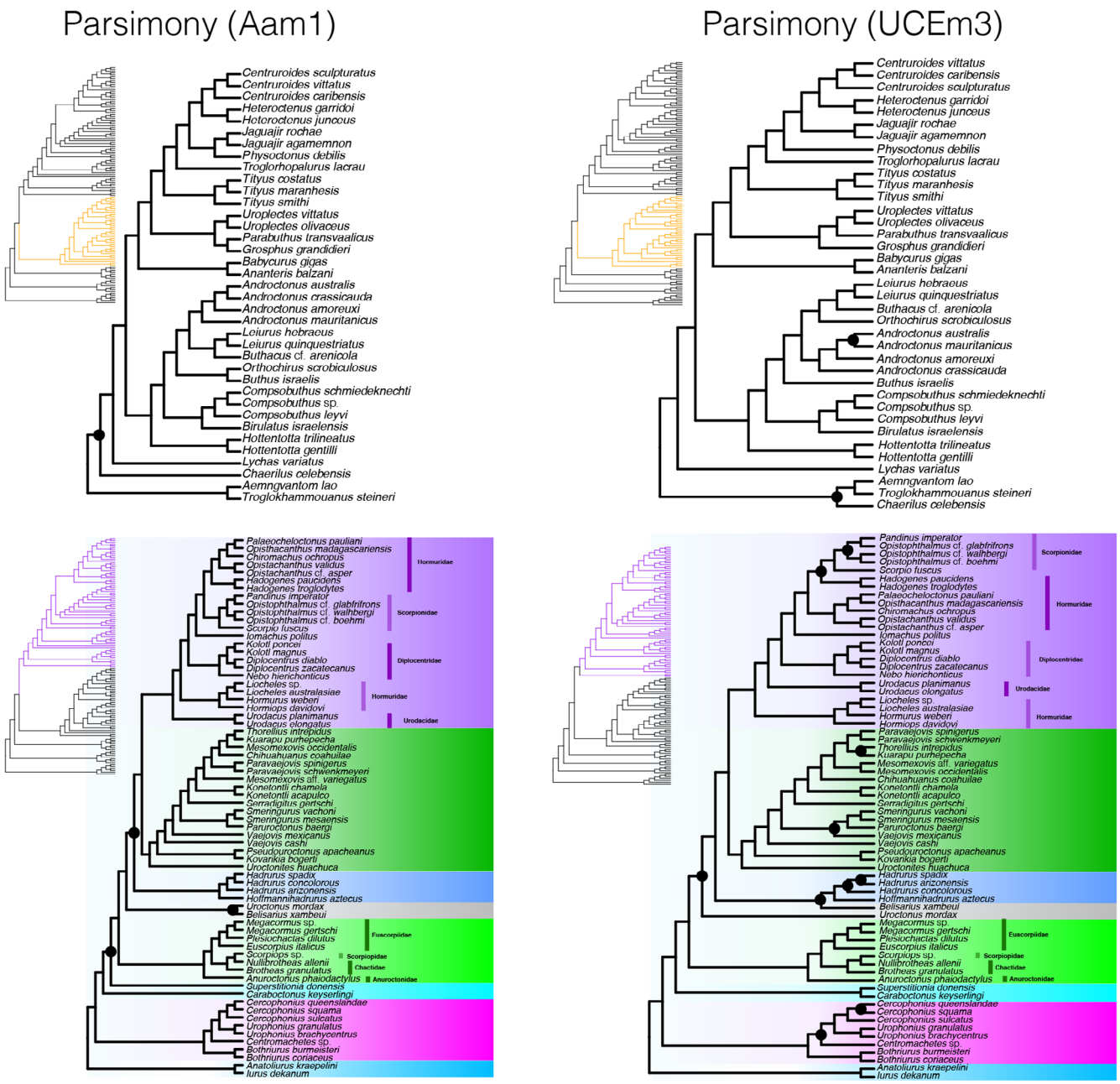

Fig. 3. Most parsimonious tree topologies recovered from the analysis of 424 genes (AAm1; Length = 530 327; Consistency Index = 0.431; Retention Index = 0.744) and 149 UCE loci (UCEm3; L = 465 305; CI = 0.267; RI = 0.670). Topologies are split into the two parvorders: Buthida (in orange) and Iurida (purple). Circles on nodes indicate <70% of jackknife nodal support.

One of the most enigmatic scorpion genera is *Anuroctonus*, as stated by its complicated systematic history. Originally, this species was described as *Centrurus phaiodactylus* by Wood (1863) from the "Utah Territory" USA, and later transferred to *Anuroctonus* and placed in the subfamily Iurini of family Iuridae (Pocock, 1893). Later, this genus was transferred to subfamily Vejovini of family Scorpionidae by Kraepelin (1894), to the subfamily Uroctoninae within Vaejovidae (Bücherl, 1971), to the subfamily Hadrurinae within Vaejovidae (Stahnke, 1974), and as a member of the superfamily Chactoidea (Francke and Soleglad, 1981). In more recent years, *Anuroctonus* was a member of Uroctoninae within family Chactidae (Soleglad and Fet, 2004), and lastly, as member of family Iuridae (Prendini and Wheeler, 2005). While these subfamilies and families have been revised previously (e.g. Sharma et al., 2015; Santibáñez-López et al., 2019a, 2020), the position of

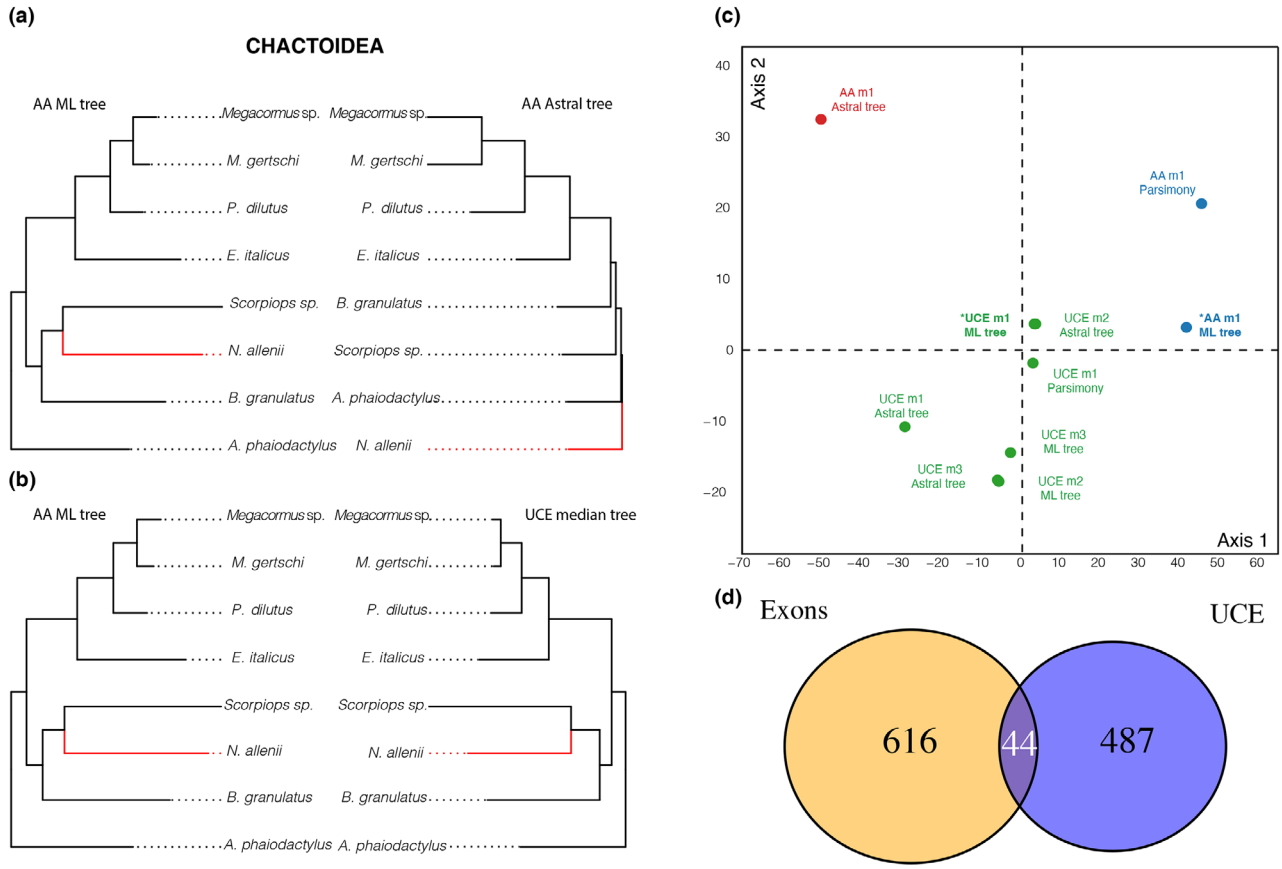

Fig. 4. (a–d) Tree topologies from the analysis of the AAm1 matrix (ML and coalescent species methods) and the median UCE tree differ in the position of multiple taxa within the superfamily Chactoidea (a, b). (c) Multidimensional scaling plot of topological tree space of the different phylogenetic analysis conducted here using the Kendall and Colijn (2016) method. (d) Venn diagram showing the overlap of 44 exons (out of 660) with 44 UCE loci (out of 531). These 44 exons/UCE loci were between 90 and 100% identical with *E*-values lower than $1e^{-102}$.

Anuroctonus remained as *incertae sedis* in Chactoidea. Here, we propose the creation of family Anuroctonidae (a new family, see the section below) to accommodate the genus *Anuroctonus*. This family is diagnosed by small denticle(s) on the base of the ventral edge of the cheliceral movable finger, developed dorsal and ventral pedipalp patellar spurs, pedipalp trichobothrial pattern type "C" (Vachon, 1974), with neobothriotaxy on the patella and chela surfaces (type Ch3 as defined by Soleglad and Fet, 2003), sternum type II (subpentagonal), lateral ocelli type 3A (Loria and Prendini, 2014) and the "twofold" spermatophore type (Monod et al., 2017).

Congruence between phylogenomic datasets reaffirms the modern classification of scorpions

In many ways, the decade-long predominance of higher-level scorpion phylogeny by phylotranscriptomic datasets has fulfilled numerous promises and unlocked new dimensions of scorpion biology. Beyond facilitating the inference of deep phylogenetic relationships, transcriptomic datasets offered exceptional insights into the composition and evolution of toxins in resting venom gland tissues (e.g. He et al., 2013; Sunagar et al., 2013; Santibáñez-López et al., 2016a, 2017b, 2018, 2019b, 2022; Díaz et al., 2023). Transcriptome data have revealed evidence for ancient, shared genome duplication in arachnopulmonates, a vital discovery in the placement of Scorpiones in the chelicerate Tree of Life (Schwager et al., 2017; Nolan et al., 2020; Ontano et al., 2021). Additionally, the data facilitated the developmental genetic study of patterning genes during scorpion embryogenesis (Sharma et al., 2014b; Setton and Sharma, 2018; Sharma, 2018). Despite these insights, the era of phylogenetic investigation using transcriptome data has now passed its zenith. With the backbone phylogeny of the order well resolved, the aforementioned challenges and costs of mRNA sequencing make this strategy unattractive for the purposes of phylogenetic inference alone.

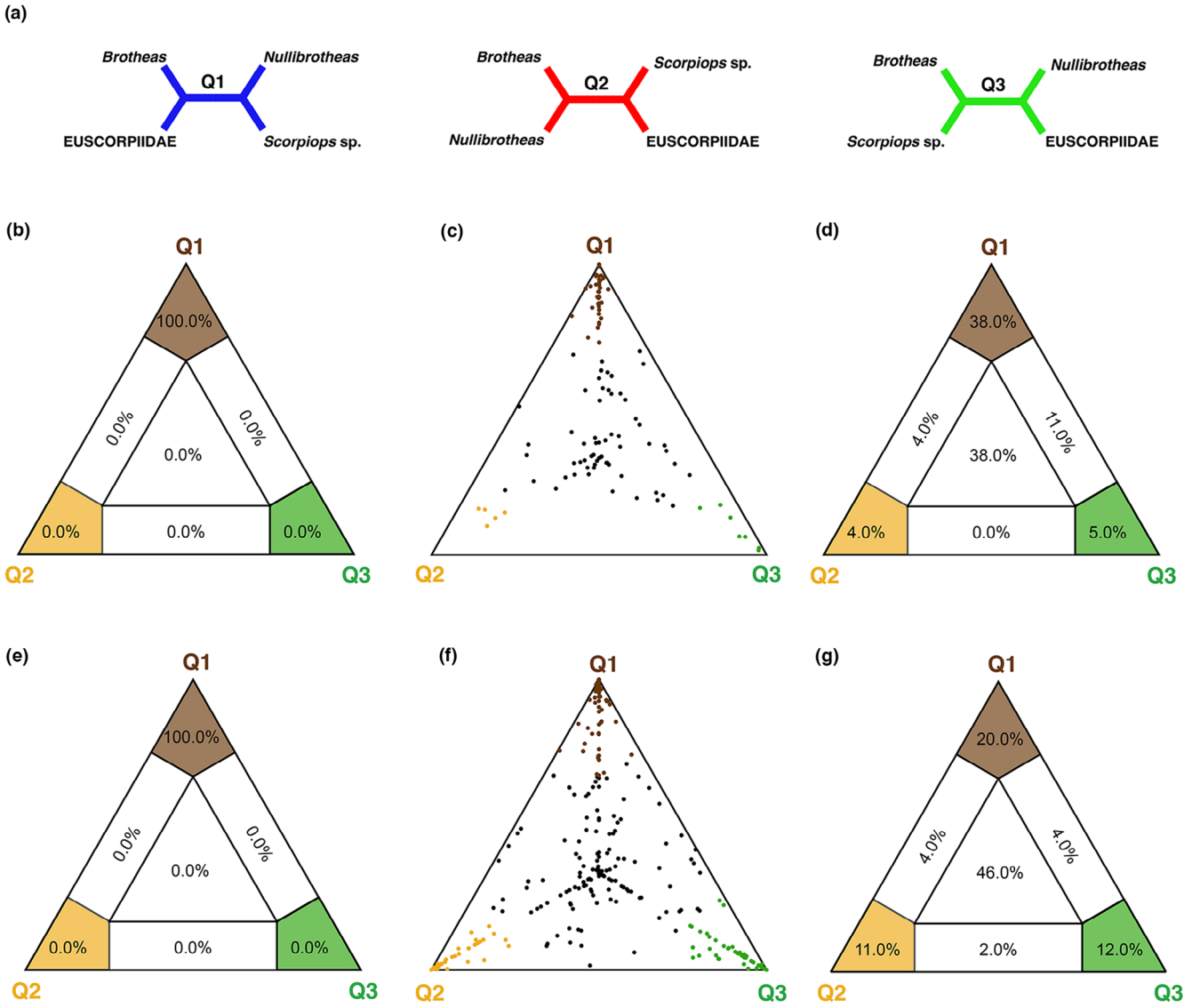

Fig. 5. (a) Quartet likelihood mapping of the three alternative topologies to test the position of *Brotheas* and *Nullibrotheas*. The right column (b and e) shows the results from the concatenated matrix (AAm1 and UCEm1 respectively) with the respective percentage of the informative regions of the map. The centre column (c and f) aggregates the mapping of all quartets analysed in 424 (c, AAm1) and 531 (f, UCE) loci, with the summary distribution of the proportion of the informative areas shown in d (AAm1) and g (UCEm1).

Ultraconserved elements offer a solid workaround for approaching phylogenetic questions, and especially so as a function of their efficacy with degraded tissues and aging collections (Blaimer et al., 2016; Derkarabetian et al., 2019). Assessment of UCE effectiveness varies across arachnid orders. In the case of spiders and harvestmen, results have been promising (Kulkarni et al., 2020; Derkarabetian et al., 2021), whereas tests of UCE datasets with Parasitiformes have yielded unusual outcomes, such as Opiliones and Ricinulei being nested within the parasitiforms (Van Dam et al., 2019). This result strongly conflicts with morphological data, a previous generation of Sanger-based molecular phylogenetic inferences, and phylotranscriptomic and genomic approaches (Giribet et al., 2002; Shultz, 2007; Sharma et al., 2014a; Leite et al., 2018; Gainett et al., 2021; Ballesteros et al., 2022). In the case of scorpions, efforts to leverage UCEs to infer scorpion relationships were limited to a seven-taxon proof-of-concept study by Starrett et al. (2017), with results that were largely incongruent with phylotranscriptomics.

Here, we compared the performance of UCE datasets to phylotranscriptomic counterparts for 126 scorpion taxa. We found that regardless of occupancy threshold, and lack of gene overlap between data

types, higher-level relationships were highly congruent across analyses; the only nodes that exhibited discrepancies between analyses were those splits that were already known to be unstable in phylotranscriptomic matrices (e.g. the placement of *Lychas*; the position of *Uroctonus* with respect to *Hadrurus* and *Hoffmannihadrurus*). These outcomes contrast with the case study of Symphytognathoidea by Kulkarni et al. (2020), who recovered the miniaturized spiders as a clade with support, contrary to exon-based analyses. A possible explanation for this discordance may be related to heterogeneous rates of evolution, with miniaturized taxa often exhibiting accelerated substitution rates in comparison with larger-bodied outgroups (e.g. Roxo et al., 2017; but see also Rainford et al., 2016). Disparities of evolutionary rate and generation times are especially pronounced in spiders, with large-bodied species often taking years to reach sexual maturity (e.g. many mygalomorphs), whereas small-bodied entelegyne groups reproduce several times per year (e.g. Elgar, 1995; Huber, 2005; Mason et al., 2018). For such taxa, nucleotide-based analyses may be more prone to saturation and rate heterogeneity effects than amino acid-based analyses, owing to the smaller alphabet of nucleotide sequences. Consistent with this interpretation, we found no discordance between UCE and exon datasets and their attendant analyses in this study, which probably reflects the long generation times and comparable evolutionary rates across the scorpion Tree of Life.

These results validate recent efforts to reframe and reclassify scorpions, as part of efforts to render constituent taxa monophyletic and reduce reliance upon morphological characters that are uninformative or homoplastic (Sharma et al., 2015; Santibáñez-López et al., 2018, 2019a). Furthermore, our work provides a first set of comprehensive UCE datasets for inference of scorpion relationships, with broad sampling across the order, which we anticipate will prove a valuable resource for elaboration and expansion for study of derived clades.

Outstanding goals for scorpion phylogeny

In the wake of the rapid influx of genome-scale datasets into scorpion systematics over the past decade, several basic questions about the shape of the scorpion Tree of Life have been resolved and appear insensitive to the type of phylogenetic loci analysed. We add the caveat that the UCEs we explored are almost entirely from coding regions of transcriptomes, owing to the dearth of high-quality genomes spanning the order Scorpiones at the time of this writing. Beyond testing new marker types and targeting key missing taxa to enrich our understanding of scorpion relationships, we identity four salient inquiries as high-value targets for scorpion phylogenetic studies that may be facilitated by the establishment of our UCE datasets:

1. *Systematic assessment of Chactidae*. The taxonomic action taken herein only partly resolves the non-monophyly of Chactidae, as we never obtained *Nullibrotheas* and *Brotheas* as sister groups across our analyses. While the taxonomic sampling of chactids is minimal in this study, the strong support for this outcome hints at the likely non-monophyly of Chactidae as presently defined and reflects a recurring pattern in scorpion phylogenomics, wherein a result established by a smaller dataset is robustly recovered by the addition of tips to the tree. Prominent examples include the dissolution of Scorpionoidea *sensu* Sissom (1990) (Bothriuridae + the remaining scorpionoids) and Chactoidea *sensu* Sissom (1990) (chactids and vaejovids forming a clade) (Sharma et al., 2015; Santibáñez-López et al., 2018, 2019a). Future investigations must target *Chactas* and other key Neotropical genera to assess the internal relationships of this putative family.
2. *Biogeography of Bothriuridae*. As in the case for scorpion systematics, only a handful number of families, genera or species have been explored for biogeographic study using molecular datasets (e.g. Bryson et al., 2013; Monod and Prendini, 2015; Graham et al., 2017; Esposito and Prendini, 2019; Borges et al., 2020; Parmakelis et al., 2022). The relationships of Bothriuridae are especially compelling from the perspective of comparative biogeography, because these scorpions exhibit a temperate Gondwanan distribution, occurring in Australia, southern South America and southern Africa (Sharma et al., 2018). Available molecular data for Bothriuridae, including Sanger data, are missing for much of the diversity of the South American fauna, and entirely missing for the African bothriurid genera *Lisposoma* and *Brandbergia*, rendering an incomplete biogeographic reconstruction for this scorpion superfamily. Sampling of the transcriptomes of many of these groups is hindered by their rarity and their restriction to very specific and remote habitats (e.g. *Brandbergia*). We anticipate that the historical biogeography of bothriurids can finally be explored with genome-scale datasets through UCE sequencing of preserved museum specimens.
3. *The root of Buthidae*. Buthidae encompasses half the diversity of extant species and the large majority of medically significant species, making this family enormously significant for biological investigations beyond taxonomy. Recent molecular studies have recovered support for some of the buthid groups delimited by Fet et al. (2005) based on

morphology, such as the *Buthus* group and *Tityus* group; other groups were not supported (Santibáñez-López et al., 2022; Štundlová et al., 2022). The major discrepancy between recent densely sampled, Sanger-based studies (228 exemplars; Štundlová et al., 2022) and more sparsely sampled, transcriptome-based studies (32 exemplars; Santibáñez-López et al., 2022) pertains to the root of the buthid tree. A four-gene Sanger dataset and some exon-based matrices (the densest dataset of Santibáñez-López et al., 2022) both recover the *Buthus* group as the sister group to the remaining buthids, but support for a clade composed of the remaining buthids was mixed in both studies. Future efforts to understand the evolution of venoms within this group must resolve the root of Buthidae with confidence.

4. *Missing families and the evolution of troglobitism*. Testing the phylogenetic validity and position of Heteroscorpionidae, Rugodentidae, Troglotayosicidae and Typhlochactidae remains an outstanding objective for higher-level scorpion phylogeny. To our knowledge, no molecular data are available for these groups. The troglobitic members of Typhlochactidae are especially intriguing from the perspective of morphology, but may be phylogenetically misplaced, owing to convergent patterns of evolution incurred by adaptations to life in darkness. Prendini et al. (2010) suggested that troglomorphism in endogean species might have evolved from obligate troglobitic typhlochactids. Unfortunately, this hypothesis has not been tested using molecular data. UCE data from troglobitic species could potentially uncover the evolutionary history of these morphological traits, along with elucidating the biogeography histories of these animals, as many of these species display disjunct distributions in cave systems (Santibáñez-López et al., 2014).

Taxonomy

Family ANUROCTONIDAE Santibáñez-López, Ojanguren-Affilastro, Graham et Sharma new family

Type genus. *Anuroctonus* Pocock, 1893 by present designation.

Diagnosis. Neobothriotaxy type Ch3 (Soleglad and Fet, 2003), with 18–25 (external surface) and 10–19 (ventral surface) trichobothria on the pedipalp patella. Lateral ocelli type 3A (Loria and Prendini, 2014). Twofold spermatophore type with a capsule with a short protruding sperm duct (Monod et al., 2017).

Composition. This family includes only one genus (*Anuroctonus*) and the species *A. phaiodactylus* and *A. pococki* (with two subspecies: *A. pococki poccocki*, *A. pococki bajae*).

Distribution. This species is restricted to California, Nevada and Utah in the USA, and in northern Baja California, Mexico.

Acknowledgements

We are indebted to Sara Ceccarelli and George Graham for providing support during the field trip in Baja California, Mexico. We thank Ryan Jones, Joshua Idjadi, Hernán Iuri, Luis Piacentini, Fermin Alfaro, Juan Calderon, Alberto Castex and Jaime Pizarro for their help in the field work in Argentina, Chile and Mexico. Access to computing nodes for intensive tasks was provided by the Bioinformatics Resource Center of the University of Wisconsin—Madison. Fieldwork in Argentina in Chile was conducted by Andrés Ojanguren-Affilastro and Jaime Pizarro. Specimens in Mexico were collected under permits issued by SEMARNAT to Sara Ceccarelli. This project was partially funded by the Connecticut State University American Association of University Professors (CSU-AAUP) grant awarded to Carlos Santibáñez (W94112), by the NSF grants DEB-1754030 awarded to Matthew Graham, and IOS-2016141 awarded to Prashant Sharma.

Conflict of interest

None declared.

REFERENCES

Alda, F., Tagliacollo, V.A., Bernt, M.J., Waltz, B.T., Ludt, W.B., Faircloth, B.C., Alfaro, M.E., Albert, J.S. and Chakrabarty, P., 2019. Resolving deep nodes in an ancient radiation of Neotropical fishes in the presence of conflicting signals from incomplete lineage sorting. Syst. Biol. 68, 573–593.

Alda, F., Ludt, W.B., Elías, D.J., McMahan, C.D. and Chakrabarty, P., 2021. Comparing ultraconserved elements and exons for phylogenomic analyses of middle American cichlids: when data agree to disagree. Genome Biol. Evol. 13, evab161.

Bücherl, W., 1971. Classification, biology and venom extraction of scorpions. In: Bücherl, W. and Buckley, E. (Eds.), Venomous Animals and their Venoms, Vol. 3. Academic Press, New York, pp. 317–348.

Ballesteros, J.A. and Sharma, P.P., 2019. A critical appraisal of the placement of Xiphosura (Chelicerata) with account of known sources of phylogenetic error. Syst. Biol. 68, 896–917.

Ballesteros, J.A., Santibáñez-López, C.E., Baker, C.M., Benavides, L.R., Cunha, T.J., Gainett, G., Ontano, A.Z., Setton, E.V., Arango, C.P., Gavish-Regev, E. and Harvey, M.S., 2022. Comprehensive species sampling and sophisticated algorithmic

approaches refute the monophyly of Arachnida. Mol. Biol. Evol. 39, msac021.
Blaimer, B.B., Lloyd, M.W., Guillory, W.X. and Brady, S.G., 2016. Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. PLoS ONE 11, e0161531.
Borges, A., Lomonte, B., Angulo, Y., de Patiño, H.A., Pascale, J.M., Otero, R., Miranda, R.J., De Sousa, L., Graham, M.R., Gómez, A. and Pardal, P.P., 2020. Venom diversity in the Neotropical scorpion genus *Tityus*: implications for antivenom design emerging from molecular and immunochemical analyses across endemic areas of scorpionism. Acta Trop. 204, 105346.
Bossert, S., Murray, E.A., Almeida, E.A., Brady, S.G., Blaimer, B.B. and Danforth, B.N., 2019. Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. Mol. Phylogenet. Evol. 130, 121–131.
Bryson, R.W., Jr., Savary, W.E. and Prendini, L., 2013. Biogeography of scorpions in the *Pseudouroctonus minimus* complex (Vaejovidae) from South-Western North America: Implications of ecological specialization for pre-quaternary diversification. J. Biogeogr. 40, 1850–1860.
Capella-Gutiérrez, S., Silla-Martínez, J.M. and Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.
Coddington, J.A., Giribet, G., Harvey, M.S., Prendini, L. and Walter, D.E., 2004. Arachnida. In: Cracraft, J. and Donoghue, P.C.J. (Eds.), Assembling the Tree of Life. Oxford University Press, New York, NY, pp. 296–318.
Díaz, C., Chang-Castillo, A., Lomonte, B., Bonilla, F., Víquez, C., Alfaro-Chinchilla, A., Triana, F. and Sasa, M., 2023. Venomics of the scorpion *Tityus ocelote* (Scorpiones, Buthidae): Understanding venom evolution in the subgenus *Archaeotityus*. Int. J. Pept. Res. Ther. 29, 1–17.
Derkarabetian, S., Benavides, L.R. and Giribet, G., 2019. Sequence capture phylogenomics of historical ethanol-preserved museum specimens: unlocking the rest of the vault. Mol. Ecol. Resour. 19, 1531–1544.
Derkarabetian, S., Baker, C.M., Hedin, M., Prieto, C.E. and Giribet, G., 2021. Phylogenomic re-evaluation of Triaenonychoidea (Opiliones: Laniatores), and systematics of Triaenonychidae, including new families, genera and species. Invertebr. Syst. 35, 133–157.
Elgar, M.A., 1995. The duration of copulation in spiders: comparative patterns. Rec. Aust. Mus. 52, 1–11.
Esposito, L.A. and Prendini, L., 2019. Island ancestors and New World biogeography: a case study from the scorpions (Buthidae: Centruroidinae). Sci. Rep. 9, 1–11.
Faircloth, B.C., 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. 32, 786–788.
Fet, V., Gantenbein, B., Gromov, A.V., Lowe, G. and Lourenço, W.R., 2003. The first molecular phylogeny of Buthidae (Scorpiones). Euscorpius 4, 1–10.
Fet, V., Michael, S.E. and Lowe, G., 2005. A new trichobothrial character for the high-level systematics of Buthoidea (Scorpiones: Buthida). Euscorpius 23, 1–40.
Francke, O.F. and Soleglad, M.E., 1981. The family Iuridae Thorell (Arachnida, Scorpiones). J. Arachnol. 9, 233–258.
Gainett, G., González, V.L., Ballesteros, J.A., Setton, E.V., Baker, C.M., Barolo Gargiulo, L., Santibáñez-López, C.E., Coddington, J.A. and Sharma, P.P., 2021. The genome of a daddy-long-legs (Opiliones) illuminates the evolution of arachnid appendages. Proc. Biol. Sci. 288, 20211168.
Giribet, G., Edgecombe, G.D., Wheeler, W.C. and Babbitt, C., 2002. Phylogeny and systematic position of Opiliones: a combined analysis of chelicerate relationships using morphological and molecular data. Cladistics 18, 5–7.
Goloboff, P.A. and Catalano, S.A., 2016. TNT version 1.5, including a full implementation of phylogenetic morphometrics. Cladistics 32, 221–238.
González-Santillán, E. and Prendini, L., 2015. Phylogeny of the north American vaejovid scorpion subfamily Syntropinae Kraepelin, 1905, based on morphology, mitochondrial and nuclear DNA. Cladistics 31, 341–405.
Graham, M.R., Wood, D.A., Henault, J.A., Valois, Z.J. and Cushing, P.E., 2017. Ancient lakes, Pleistocene climates and river avulsions structure the phylogeography of a large but little-known rock scorpion from the Mojave and Sonoran deserts. Biol. J. Linn. Soc. 122, 133–146.
He, Y., Zhao, R., Di, Z., Li, Z., Xu, X., Hong, W., Wu, Y., Zhao, H., Li, W. and Cao, Z., 2013. Molecular diversity of Chaerilidae venom peptides reveals the dynamic evolution of scorpion venom components from Buthidae to non-Buthidae. J. Proteomics 89, 1–14.
Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q. and Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.
Huber, B.A., 2005. Sexual selection research on spiders: progress and biases. Biol. Rev. 80, 363–385.
Jombart, T., Kendall, M., Almagro-Garcia, J. and Colijn, C., 2017. Treespace: statistical exploration of landscapes of phylogenetic trees. Mol. Ecol. Resour. 17, 1385–1392.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., Von Haeseler, A. and Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.
Karin, B.R., Gamble, T. and Jackman, T.R., 2020. Optimizing phylogenomics with rapidly evolving long exons: comparison with anchored hybrid enrichment and ultraconserved elements. Mol. Biol. Evol. 37, 904–922.
Katoh, K. and Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Kendall, M. and Colijn, C., 2016. Mapping phylogenetic trees to reveal distinct patterns of evolution. Mol. Biol. Evol. 33, 2735–2743.
Kent, W.J., 2002. BLAT—The BLAST-like alignment tool. Genome Res. 12, 656–664.
Kraepelin, K., 1894. Revision der Scorpione. II. Scorpionidae und Bothriuridae. Beiheft zum Jahrbuch der Hamburgischen Wissenschaftlichen Anstalten 11, 1–248.
Kulkarni, S., Wood, H., Lloyd, M. and Hormiga, G., 2020. Spider-specific probe set for ultraconserved elements offers new perspectives on the evolutionary history of spiders (Arachnida, Araneae). Mol. Ecol. Resour. 20, 185–203.
Leite, D.J., Baudouin-Gonzalez, L., Iwasaki-Yokozawa, S., Lozano-Fernandez, J., Turetzek, N., Akiyama-Oda, Y., Prpic, N.M., Pisani, D., Oda, H., Sharma, P.P. and McGregor, A.P., 2018. Homeobox gene duplication and divergence in arachnids. Mol. Biol. Evol. 35, 2240–2253.
Loria, S.F. and Prendini, L., 2014. Homology of the lateral eyes of Scorpiones: a six-ocellus model. PLoS ONE 9, e112913.
Loria, S.F., Ehrenthal, V.L., Nguyen, A.D. and Prendini, L., 2022. Climate relicts: Asian scorpion family Pseudochactidae survived Miocene aridification in caves of the Annamite Mountains. Insect Syst. Divers. 6, 3.
Mason, L.D., Wardell-Johnson, G. and Main, B.Y., 2018. The longest-lived spider: mygalomorphs dig deep, and persevere. Pac. Conserv. Biol. 24, 203–206.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A. and Lanfear, R., 2020a. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534.
Minh, B.Q., Hahn, M.W. and Lanfear, R., 2020b. New methods to calculate concordance factors for phylogenomic datasets. Mol. Biol. Evol. 73, 2727–2733.
Mirarab, S. and Warnow, T., 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31, i44–i52.
Mistry, J., Finn, R.D., Eddy, S.R., Bateman, A. and Punta, M., 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41, e121.
Mongiardino, K.N., 2021. Phylogenomic subsampling and the search for phylogenetically reliable loci. Mol. Biol. Evol. 38, 4025–4038.

Monod, L. and Prendini, L., 2015. Evidence for Eurogondwana: the roles of dispersal, extinction and vicariance in the evolution and biogeography of Indo-Pacific Hormuridae (Scorpiones: Scorpionoidea). Cladistics 31, 71–111.
Monod, L., Cauwet, L., González-Santillán, E. and Huber, S., 2017. The male sexual apparatus in the order Scorpiones (Arachnida): a comparative study of functional morphology as a tool to define hypotheses of homology. Front. Zool. 14, 1–48.
Nolan, E.D., Santibáñez-López, C.E. and Sharma, P.P., 2020. Developmental gene expression as a phylogenetic data class: support for the monophyly of Arachnopulmonata. Dev. Genes Evol. 230, 137–153.
Ontano, A.Z., Gainett, G., Aharon, S., Ballesteros, J.A., Benavides, L.R., Corbett, K.F., Gavish-Regev, E., Harvey, M.S., Monsma, S., Santibáñez-López, C.E. and Setton, E.V., 2021. Taxonomic sampling and rare genomic changes overcome long-branch attraction in the phylogenetic placement of pseudoscorpions. Mol. Biol. Evol. 38, 2446–2467.
Parmakelis, A., Dimitriadou, D., Gkigkiza, E., Karampatsou, L., Stathi, I., Fet, V., Yağmur, E.A. and Kovařík, F., 2022. The evolutionary history of the relict scorpion family Iuridae of the eastern Mediterranean. Mol. Phylogenet. Evol. 177, 107622.
Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G. and Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9, e1000602.
Pocock, R.I., 1893. Notes on the classification of scorpions, followed by some observations on synonymy, with descriptions of new genera and species. Ann. Mag. Nat. Hist. 6, 303–330.
Prendini, L. and Wheeler, W.C., 2005. Scorpion higher phylogeny and classification, taxonomic anarchy, and standards for peer review in online publishing. Cladistics 21, 446–494.
Prendini, L., Crowe, T.M. and Wheeler, W.C., 2003. Systematics and biogeography of the family Scorpionidae (Chelicerata: Scorpiones), with a discussion on phylogenetic methods. Invertebr. Syst. 17, 185–259.
Prendini, L., Francke, O.F., Vignoli, V., 2010. Troglomorphism, trichobothriotaxy and typhlochactid phylogeny (Scorpiones, Chactoidea): more evidence that troglobitism is not an evolutionary dead-end. Cladistics. 26(2), 117–142.
Rainford, J.L., Hofreiter, M. and Mayhew, P.J., 2016. Phylogenetic analyses suggest that diversification and body size evolution are independent in insects. BMC Evol. Biol. 16, 1–17.
Rein, J.O. (Ed)., 2022. The Scorpion Files (version Jun 2015). In: Roskov, Y., Abucay, L., Orrell, T., Nicolson, D., Flann, C., Bailly, N., Kirk, P., Bourgoin, T., DeWalt, R.E., Decock, W. and De Wever, A. (Eds.), Species 2000 & ITIS Catalogue of Life, 2016 Annual Checklist. Digital resource at www.catalogueoflife.org/annual-checklist/2016. Species 2000: Naturalis, Leiden, the Netherlands. Accessed December 20th, 2022.
Roxo, F.F., Lujan, N.K., Tagliacollo, V.A., Waltz, B.T., Silva, G.S., Oliveira, C. and Albert, J.S., 2017. Shift from slow-to fast-water habitats accelerates lineage and phenotype evolution in a clade of neotropical suckermouth catfishes (Loricariidae: Hypoptopomatinae). PLoS ONE 12, e0178240.
Santibáñez-López, C.E., Francke, O.F. and Prendini, L., 2014. Shining a light into the world's deepest caves: phylogenetic systematics of the troglobiotic scorpion genus *Alacran* Francke, 1982 (Typhlochactidae: Alacraninae). Invertebr. Syst. 28, 643–664.
Santibáñez-López, C.E., Francke, O.F., Ureta, C. and Possani, L.D., 2016a. Scorpions from Mexico: from species diversity to venom complexity. Toxins 8, 2.
Santibáñez-López, C.E., Kriebel, R. and Sharma, P.P., 2017a. *Eadem figura manet*: measuring morphological convergence in diplocentrid scorpions (Arachnida: Scorpiones: Diplocentridae) under a multilocus phylogenetic framework. Invertebr. Syst. 31, 233–248.
Santibáñez-López, C.E., Cid-Uribe, J.I., Zamudio, F.Z., Batista, C.V., Ortiz, E. and Possani, L.D., 2017b. Venom gland transcriptomic and venom proteomic analyses of the scorpion *Megacormus gertschi* Díaz-Najera, 1966 (Scorpiones: Euscorpiidae: Megacorminae). Toxicon 133, 95–109.
Santibáñez-López, C.E., Kriebel, R., Ballesteros, J.A., Rush, N., Witter, Z., Williams, J., Janies, D.A. and Sharma, P.P., 2018. Integration of phylogenomics and molecular modeling reveals lineage-specific diversification of toxins in scorpions. PeerJ 6, e5902.
Santibáñez-López, C.E., González-Santillán, E., Monod, L. and Sharma, P.P., 2019a. Phylogenomics facilitates stable scorpion systematics: reassessing the relationships of Vaejovidae and a new higher-level classification of Scorpiones (Arachnida). Mol. Phylogenet. Evol. 135, 22–30.
Santibáñez-López, C.E., Graham, M.R., Sharma, P.P., Ortiz, E. and Possani, L.D., 2019b. Hadrurid scorpion toxins: evolutionary conservation and selective pressures. Toxins 11, 637.
Santibáñez-López, C.E., Ojanguren-Affilastro, A.A. and Sharma, P.P., 2020. Another one bites the dust: taxonomic sampling of a key genus in phylogenomic datasets reveals more non-monophyletic groups in traditional scorpion classification. Invertebr. Syst. 34, 133–143.
Santibáñez-López, C.E., Aharon, S., Ballesteros, J.A., Gainett, G., Baker, C.M., González-Santillán, E., Harvey, M.S., Hassan, M.K., Abu Almaaty, A.H., Aldeyarbi, S.M. and Monod, L., 2022. Phylogenomics of scorpions reveal contemporaneous diversification of scorpion mammalian predators and mammal-active sodium channel toxins. Syst. Biol. 71, 1281–1289.
Schwager, E.E., Sharma, P.P., Clarke, T., Leite, D.J., Wierschin, T., Pechmann, M., Akiyama-Oda, Y., Esposito, L., Bechsgaard, J., Bilde, T. and Buffry, A.D., 2017. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. BMC Biol. 15, 1–27.
Setton, E.V. and Sharma, P.P., 2018. Cooption of an appendage-patterning gene cassette in the head segmentation of arachnids. Proc. Natl Acad. Sci. USA 115, E3491–E3500.
Sharma, P.P., 2018. Chelicerates. Curr. Biol. 28, R774–R778.
Sharma, P.P., Kaluziak, S.T., Pérez-Porro, A.R., González, V.L., Hormiga, G., Wheeler, W.C. and Giribet, G., 2014a. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. Mol. Biol. Evol. 31, 2963–2984.
Sharma, P.P., Schwager, E.E., Extavour, C.G. and Wheeler, W.C., 2014b. Hox gene duplications correlate with posterior heteronomy in scorpions. Proc. Biol. Sci. 281, 20140661.
Sharma, P.P., Fernández, R., Esposito, L.A., González-Santillán, E. and Monod, L., 2015. Phylogenomic resolution of scorpions reveals multilevel discordance with morphological phylogenetic signal. Proc. Biol. Sci. 282, 20142953.
Sharma, P.P., Baker, C.M., Cosgrove, J.G., Johnson, J.E., Oberski, J.T., Raven, R.J., Harvey, M.S., Boyer, S.L. and Giribet, G., 2018. A revised dated phylogeny of scorpions: phylogenomic support for ancient divergence of the temperate Gondwanan family Bothriuridae. Mol. Phylogenet. Evol. 122, 37–45.
Shultz, J.W., 2007. A phylogenetic analysis of the arachnid orders based on morphological characters. Zool. J. Linn. Soc. 150, 221–265.
Sissom, W.D., 1990. Systematics, biogeography and paleontology. In: Polis, G.A. (Ed.), Biology of Scorpions. Stanford University Press, Stanford, CA, pp. 64–160.
Soleglad, M.E. and Fet, V., 2003. High-level systematics and phylogeny of the extant scorpions (Scorpiones: Orthosterni). Euscorpius 11, 1–56.
Soleglad, M.E. and Fet, V., 2004. The systematics of the scorpion subfamily Uroctoninae (Scorpiones: Chactidae). Rev. Iber. Aracnol. 10, 81–128.
Stahnke, H.L., 1974. Revision and keys to the higher categories of Vejovidae. J. Arachnol. 1, 107–141.
Starrett, J., Derkarabetian, S., Hedin, M., Bryson, R.W., Jr., McCormack, J.E. and Faircloth, B.C., 2017. High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. Mol. Ecol. Resour. 17, 812–823.

Steenwyk, J.L., Buida, T.J., Labella, A.L., Li, Y., Shen, X.X. and Rokas, A., 2021. PhyKIT: a broadly applicable UNIX shell toolkit for processing and analyzing phylogenomic data. Bioinformatics 37, 2325–2331.

Strimmer, K. and Von Haeseler, A., 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc. Natl. Acad. Sci. USA 94, 6815–6819.

Štundlová, J., Šťáhlavský, F., Opatova, V., Stundl, J., Kovařík, F., Dolejš, P. and Šmíd, J., 2022. Molecular data do not support the traditional morphology-based groupings in the scorpion family Buthidae (Arachnida: Scorpiones). Mol. Phylogenet. Evol. 173, 107511.

Sunagar, K., Undheim, E.A., Chan, A.H., Koludarov, I., Muñoz-Gómez, S.A., Antunes, A. and Fry, B.G., 2013. Evolution stings: the origin and diversification of scorpion toxin peptide scaffolds. Toxins 5, 2456–2487.

Vachon, M., 1974. Etude des caracteres utilises pour classer les familles et les genres de Scorpions (Arachnides). 1. La trichobothriotaxie en arachnologie. Sigles trichobothriaux et types de trichobothriotaxie chez les Scorpions. Bull. Mus. Nat. Hist. Nat. Paris 140, 857–958.

Van Dam, M.H., Trautwein, M., Spicer, G.S. and Esposito, L., 2019. Advancing mite phylogenomics: designing ultraconserved elements for Acari phylogeny. Mol. Ecol. Resour. 19, 465–475.

Weygoldt, P. and Paulus, H.F., 1979. Untersuchungen zur Morphologie, Taxonomie und Phylogenie der Chelicerata. Zool. Syst. Evol. 17, 85–116.

Wheeler, W.C. and Hayashi, C.Y., 1998. The phylogeny of the extant chelicerate orders. Cladistics 14, 173–192.

Wood, H.C., 1863. Descriptions of new species of North American Pedipalpi. Proc. Acad. Nat. Sci. Phila. 1863, 107–112.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Modified boxplots summarizing information content in the four matrices.

Fig. S2. Seven gene properties against the order in which the 424 amino acid genes are ranked by their phylogenetic usefulness.

Fig. S3. Six loci properties against the order in which the 531 UCE loci are ranked by their phylogenetic usefulness.

Fig. S4. Six loci properties against the order in which the 394 UCE loci are ranked by their phylogenetic usefulness.

Fig. S5. Six loci properties against the order in which the 149 UCE loci are ranked by their phylogenetic usefulness.

Fig. S6. Maximum likelihood tree topology recovered from the analysis of 394 UCE loci (UCEm2).

Fig. S7. Maximum likelihood tree topology recovered from the analysis of 149 UCE loci (UCEm3).

Fig. S8. Astral tree topology recovered from the analysis of 424 amino acid gene trees (AAm1).

Fig. S9. Astral tree topology recovered from the analysis of 531 UCE loci trees (UCEm1).

Fig. S10. Astral tree topology recovered from the analysis of 394 UCE loci trees (UCEm2).

Fig. S11. Astral tree topology recovered from the analysis of 149 UCE loci trees (UCEm3).