Molecular Phylogenetics and Evolution
May 2023, Volume 182, Pages 107726 (21p.)
https://doi.org/10.1016/j.ympev.2023.107726
https://archimer.ifremer.fr/doc/00820/93201/

Phylogeny of sea spiders (Arthropoda: Pycnogonida) inferred from mitochondrial genome and 18S ribosomal RNA gene sequences

Sabroux Romain [1], Corbari Laure [1], Hassanin Alexandre [1, *]

[1] Institut Systématique Evolution Biodiversité (ISYEB), Sorbonne Université, MNHN, CNRS, EPHE, UA, 57 rue Cuvier, CP 51, 75005 Paris, France

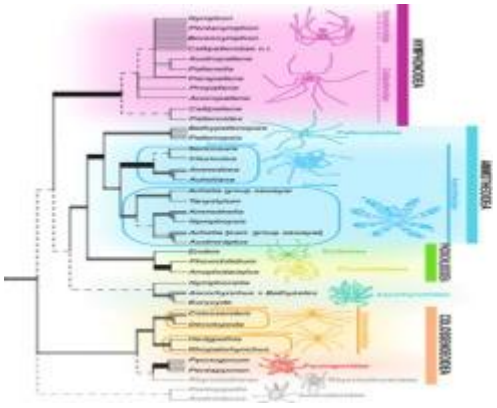
* Corresponding author : Alexandre Hassanin, email address : alexandre.hassanin@mnhn.fr

Abstract :

The phylogeny of sea spiders has been debated for more than a century. Despite several molecular studies in the last twenty years, interfamilial relationships remain uncertain. In the present study, relationships within Pycnogonida are examined in the light of a new dataset composed of 160 mitochondrial genomes (including 152 new sequences) and 130 18S rRNA gene sequences (including 120 new sequences), from 141 sea spider morphospecies representing 26 genera and 9 families. Node congruence between mitochondrial and nuclear markers was analysed to identify the most reliable relationships. We also reanalysed a multilocus dataset previously published and showed that the high percentages of missing data make phylogenetic conclusions difficult and uncertain.

Our results support the monophyly of most families currently accepted, except Callipallenidae and Nymphonidae, the monophyly of the superfamilies Ammotheoidea (Ammotheidae + Pallenopsidae), Nymphonoidea (Nymphonidae + Callipallenidae), Phoxichilidioidea (Phoxichilidiidae + Endeidae) and Colossendeoidea (Colossendeidae + Pycnogonidae + Rhynchothoracidae), and the sister-group relationship between Ammotheoidea and Phoxichilidioidea. We discuss the morphological evolution of sea spiders, identifying homoplastic characters and possible synapomorphies. We also discuss the palaeontological and phylogenetic arguments supporting either a radiation of sea spiders prior to Jurassic or a progressive diversification from Ordovician or Cambrian.

Graphical abstract



Highlights

► 152 mitogenomes and 120 18S ribosomal genes of sea spiders were sequenced. ► Phylogenetic signal extraction is impacted when high levels of missing data are included. ► Strong support for four superfamilies, six families, four subfamilies. ► Most cephalic appendage characters have evolved by convergence in different families. ► Poorly resolved deep relationships may be due to radiation before Jurassic.

Keywords : Pantopoda, molecular systematics, missing data, suprafamilial synapormophies, homoplasy, radiation.

1. INTRODUCTION

Sea spiders (Class Pycnogonida) are inconspicuous, yet fascinating marine arthropods. These animals have a bizarre anatomy and morphology: (i) a proboscis extending beyond the head enabling to suck-up their prey or a piece of it (Dietz et al., 2018), (ii) a reduced body which constraints the digestive guts and gonads into the legs (Frankowski et al., 2022), (iii) specialized ovigerous legs (or ovigers) which enable males of most families to carry their offspring (Brenneis et al., 2017). At least 11 fossil species of Pycnogonida have been discovered from the Silurian (425 Myrs [million years]) to Jurassic (150 Myrs), plus two potential sea spider fossils from Cambrian and Ordovician, although their status is still debated (see a review in Sabroux et al., 2019a). Extant diversity comprises about 1,400 species (Bamber et al., 2022) divided into 11 families and 82 genera that are all included in a single order, Pantopoda. The number of known sea spiders species is limited compared to other arthropods groups, such as Diplopoda (*c.a.,* 12,000), Crustacea (*c.a.*, 67,000), Arachnida (*c.a.,* 95,000) or Hexapoda (*c.a.,* 1,024,000) (Coddington et al., 2004; Golovatch et al., 2009; Stork, 2018), nonetheless pycnogonids are remarkably diversified (fig. 1), inhabiting almost all benthic habitats from littoral to abyssal waters and tropical to polar latitudes (Arnaud and Bamber, 1987) and feeding on preys of a large taxonomic range (algae, biofilms, bryozoans, cnidarians, echinoderms, mollusks, polychaetes; Dietz et al. 2018). Some species are even ectoparasites (*e.g.*, Arnaud, 1978; Tomiyama et al., 2016). They display variable sets of appendages: cephalic appendages (chelifores, palps, ovigers) are independently present or absent depending on the taxon and sometimes on the sex; and the number of walking legs varies from eight to twelve (Arnaud and Bamber, 1987).

Although several authors have suggested that sea spiders progressively lost their appendages during evolution (*e.g.*, Stock, 1994; Munilla, 1999), this assumption did not rely on any solid phylogenetic background, but on subjective interpretation of their morphological patterns (*e.g.,* Bamber, 2007). The first attempt to propose an analysis-based classification of sea spiders dates back to Fry (1978), who

studied 45 morphological characters. Later, Munilla and de Haro (1981) interpreted sea spiders evolution through electrophoretic and immunological study of their protein content. Arango (2002) published the first most-parsimonious tree of sea spiders based on 36 morphological characters and 37 taxa, and then the first molecular phylogeny based on two nuclear markers (18S and 28S rRNA genes [18S and 28S]) sequenced for 15 species (Arango, 2003). Since these two studies were poorly conclusive, Arango and Wheeler (2007) have continued their efforts and performed an analysis combining three nuclear (nu) markers, three mitochondrial (mt) markers, and 78 morphological characters for 63 taxa including four fossil species. In parallel, Nakamura et al. (2007) proposed a 18S phylogeny based on 57 taxa. These two studies suggested almost simultaneously that Ammotheidae and Ascorhynchidae represent two distinct families. However, their conclusions were limited in scope due to DNA contamination, sequencing errors, missing data and inappropriate choice of outgroups, as discussed by Arabi et al. (2010) who published a phylogeny based on 35 taxa and five molecular markers and discussed the impact of mitogenome rearrangements on tree reconstruction. Although interfamilial relationships were weakly supported, the study of Arabi et al. was the first in which all families but Callipallenidae were found monophyletic. Chow et al. (2012) produced another 18S phylogeny of sea spiders based on 25 taxa to determine the position of the genus *Nymphonella,* which was found nested within *Ascorhynchus* (represented by six species in the study) with significant support (posterior probability [PP] $\geq$ 0.95, bootstrap percentages [BP] $\geq$ 50). Focusing on the family Ammotheidae, Sabroux et al. (2017) analysed the 18S and 5' barcode fragment of the mt cytochrome c oxidase subunit 1 gene (CO1) for 159 and 179 taxa, respectively. The results supported the monophyly of Ammotheidae and its division into two subfamilies, Ammotheinae and Achelinae. More recently, Ballesteros et al. (2021) proposed a "phylogenomic resolution of sea spider diversification" based on 89 sea spiders and 84 molecular markers (12 mt genes, 5.8S, 18S and 28S rRNA genes, 20 ultra-conserved elements [UCE] and 49 nu exons). Although interfamilial relationships were found supported by BP comprised between 66 and 100%, their dataset contained up to 75% of missing data per nucleotide position, which may be problematic for phylogenetic reconstruction.

In the present study, interfamilial relationships within Pycnogonida are re-examined in the light of a new dataset composed of 160 sea spider mitogenomes (including 152 new sequences) and 130 nu 18S sequences (including 120 new sequences) from 141 morphospecies representing 27 genera and nine families. The mitogenome provides an important amount of data (typically 15,000 bp in Pycnogonida; Masta et al., 2010) mostly consisting of 13 protein-coding genes, which are known to evolve more rapidly than protein-coding genes of the nu genome (Allio et al., 2017). As a drawback, this fast evolution can lead to high levels of saturation for inferring deep relationships. In addition, the strong bias in base composition observed in mt genes can also impact phylogenetic reconstruction (Hassanin, 2006; Hassanin et al., 2005). Mitochondrial introgression may also have a misleading impact for interpreting shallow phylogenetic relationships and species delimitation (*e.g.*, Audzijonyte and Väinölä, 2006; Petzold and Hassanin, 2020). The topological comparisons with trees inferred from nuDNA data are therefore crucial to characterize the most reliable relationships, *i.e.*, the nodes supported by both mt and nu datasets. Therefore, the 18S gene, which has been widely used for arthropod phylogeny (*e.g.*, Mallatt et al., 2012; Nosenko et al., 2013; Sabroux et al., 2017), was chosen here to test node congruence between mtDNA and nuDNA datasets. The multilocus dataset published in Ballesteros et al. (2021) was also reanalysed, focusing on node repeatability and the impact of missing data on tree reconstruction.

2. MATERIAL AND METHODS

2.1 Mitochondrial genome sequencing, assembly, and annotation

A total of 152 DNA extracts were selected from our pycnogonid DNA bank of *c.a.* 600 samples extracted from specimens of the collections of the *Muséum national d'Histoire naturelle* of Paris (MNHN) and CO1 barcoded as previously detailed (Arabi et al., 2010; Hassanin, 2006; Sabroux et al., 2017, 2019b). The DNA samples were quantified with a Qubit® 2.0 Fluorometer using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). The selection of taxa was made so that diversity was maximized, with at least 141 species/morphospecies and 27 genera, collected from various localities in tropical and Antarctic waters (table 1). Our study includes the holotypes of *Ammothella dirbergi*, *Anoplodactylus madibenthos*, *Ascorhynchus iguanarum*, *Ascorhynchus*

quartogibbus, *Hedgpethia tibialis*, *Nymphon dorlis*, *Nymphon martinicum*, *Pycnogonum cesairei* and *Tanystylum boucheti*, and the paratypes of *Ammothella dirbergi*, *Eurycyde kaiouti* and *Tanystylum ingrallis* (Bamber, 2013; Sabroux et al., 2022; Stock, 1991). The specimens were deposited in the MNHN collections referring through inventory numbers (code MNHN-IU-; see table 1).

Libraries were prepared as indicated in Hassanin et al. (2021) using the TruSeq® Nano DNA Library Prep kit (Illumina, San Diego, CA, USA) after pooling 150 ng of total DNA of 10-12 species belonging to distant taxonomic groups (*i.e.,* different phyla, classes, orders or families). Libraries were sequenced at the "Institut du Cerveau et de la Moelle épinière" (Paris, France) using NextSeq® 500 system with either NextSeq 500 Mid Output Kit v2 (300 cycles) or NextSeq 500 High Output Kit v2 (300 cycles) (Illumina, San Diego, CA, USA).

The mt genomes were assembled using Geneious Prime 2020.0.4 (Kearse et al., 2012) and annotated using as references the seven pycnogonid mt genomes available in GenBank in June 2022: *Achelia bituberculata* AY457170, *Ammothea carolinensis* GU065293, *Ammothea hilgendorfi* GU370075, *Tanystylum orbiculare* GU370074, *Colossendeis megalonyx* HQ450773, *Nymphon gracile* DQ666063, *Nymphon* sp. GU370076 (Carapelli et al., 2013; Dietz et al., 2011; Masta et al., 2010; Park et al., 2007; Podsiadlowski and Braband, 2006).

For each DNA sample, the available CO1 sequence was used as bait with low mismatch (0-2%) to assemble its mt genome using multiple iterations, so that the genome fragment extends progressively on both 5' and 3' extremities. In complement to this approach, the reads were mapped to the mt genome of *Ammothea carolinensis* (GU065293; Carapelli et al., 2013) using high mismatch percentages (from 20 to 50%). Then, mapped reads were *de novo* assembled with a low mismatch (1-2%) and contigs >500 nt with depth >10X were further used as baits, as detailed above for the CO1 gene. An additional mt genome of *Nymphon striatum* was assembled from the draft genome SRR10993134 (Jeong et al., 2020) using the same approach. Because of high genetic distances between the seven pycnogonid mt genomes available in GenBank (see list above) and our new genomes, annotation had to be generally refined by eye to delineate precisely the protein-coding genes, focusing on initiating and stop codons. The 152 new mitogenomes were deposited in GenBank under accession numbers are listed in table 1.

The mt-197 dataset includes 13 protein-coding genes extracted from complete mt genomes of 160 pycnogonids and 37 outgroup taxa representing major lineages of Arthropoda (Euchelicerata, Myriapoda, and Pancrustacea) as well as Onychophora. To avoid long branch attraction artefacts due to convergent inversions of base compositional bias, scorpions and non-Mesothelae spiders were excluded from the analyses (Hassanin et al., 2005; Arabi et al., 2012). To avoid spurious placement of the root (Rota-Stabelli and Telford, 2008), we also excluded Pseudoscorpiones, Mesostigmata and Trombidoformes mites, because of the very long branches of these taxa in our preliminary analyses.

2.2 18S rRNA sequencing

The 18S-157 dataset contains 157 sequences of the 18S rRNA gene, including 120 new sequences, which were amplified and sequenced using the three primer sets described in Arabi et al. (2010). We selected the same specimens already used for mt genomes, except for three conspecifics (from the same locality or geographical zone) (table 1). The taxonomic sampling was completed with 18S sequences available in GenBank, including ten pycnogonids (seven previously published in Sabroux et al., 2017) and 27 outgroup species. The 120 new 18S genes sequences were deposited in GenBank under accession numbers are listed in table 1.

2.3 DNA alignments used for this study

The 13 protein-coding genes (genes of ATP synthase membrane subunits 6 and 8 [ATP6 and ATP8], of cytochrome c oxidase subunits 1, 2 and 3 [CO1, CO2, CO3], and of the NADH-ubiquinone oxidoreductase chain 1, 2, 3, 4, 4L, 5, 6 [ND1, ND2, ND3, ND4, ND4L, ND5, ND6]) were extracted from mt genomes and concatenated in a mt alignment. The 14 gene alignments (13 mt genes and 18S) were performed on MEGA7 (Kumar et al., 2016) using Muscle (Edgar, 2004) and then refined by eyes. Ambiguity was treated as in Sabroux et al. (2017): regions with ambiguous positions for homology were removed from the alignments, but regions providing phylogenetic information at the family levels

were aligned as separate family-per-family shifted blocks. The 13 mt genes were concatenated into a single Nexus file.

We furthermore reanalysed the Matrix 3 published by Ballesteros et al. (2021), a dataset hereafter referred as M3-110 (as it contains 110 taxa, including 89 pycnogonids). This dataset is composed of the four following subdatasets: mt genome data (mt-110), nu exons (OG-110), ultra-conserved elements (UCE-110), and nu ribosomal genes (rib-110).

2.4 Phylogenetic analyses

Phylogenetic analyses were performed on CIPRES platform (Miller et al., 2010) using RAxML 8.2.12 (Stamatakis, 2014) for maximum likelihood (ML) analyses with 1000 fast bootstrap replicates. The ML bootstrap consensus trees were constructed using PAUP*4.0a167 (Swofford and Bell, 2017) from the RAxML bootstrap trees. All datasets but mt-211, 18S-211 and mt+18S-211 were also analysed with Bayesian Inference (BI) using MrBayes 3.2.7 (Ronquist and Huelsenbeck, 2003) running four chains for 10^7 generations and a default 25% burn-in. The DNA alignments were partitioned by genes and codon positions (when appropriate) using a GTR+G+I model for each partition following jModelTest (Posada, 2008). In this study, we focused on nodes supported by ML bootstrap percentage (BP) $\geq$ 50% and/or Bayesian posterior probability (PP) $\geq$ 0.95, although node congruence/repeatability between mt and nu datasets constitutes another important criterion.

Node repeatability between Bayesian trees reconstructed from independent datasets (*e.g.*, mt and 18S) was assessed using SuperTRI v.157 (Ropiquet et al., 2009). The lists of bipartitions obtained from Bayesian analyses were transformed into a weighted binary matrix for supertree construction using SuperTRI v57. Each binary character corresponds to a node, which was weighted according to its frequency of occurrence in one of the lists of bipartitions (*e.g.*, mt and 18S). SuperTRI produces three node support values: repeatability (Rep) is the ratio of the number of datasets supporting the specified node to the total number of datasets; Mean Posterior Probability (MPP) is calculated from the posterior probabilities (PP) obtained in the Bayesian analyses of the different datasets; and SuperTRI bootstrap

percentages (SBP) is obtained from PAUP*4.0a167 (Swofford and Bell, 2017) after 1000 bootstrap replicates of the MRP (Matrix Representation with Parsimony) file reconstructed under SuperTRI v57.

3. RESULTS

3.1 DNA alignments

The mt-197 alignment (Supplementary data A) contains 17,989 bp for 197 taxa. To limit the impact of mutational saturation on phylogenetic analysis, we also used a recoded dataset mt-197-RY, in which all positions were degenerated to purine or pyrimidine. The alignment includes in average 0.58% of missing data per sea spider taxon (table 2). Nine of the 11 families (*sensu* Bamber et al., 2007) are covered. The two remaining families (Austrodecidae and Rhynchothoracidae) were not included in the dataset as all attempts to sequence their mitogenomes led to high levels of missing data. During the review process of this article, Zehnpfennig et al. (2022) have published one mitogenome for *Austrodecus* sp. (GenBank accession number: OK623745) and two mitogenomes for *Rhynchothorax* sp. (OK649914 and OK649915). However, our reanalyses showed that no Austrodecidae and Rhynchothoracidae mitogenomes were in fact included in the study of Zehnpfennig *et al.* (2022). Based on the pictures provided by the authors, we concluded that the specimens identified as *Rhynchothorax* (Rhynchothoracidae) are actually *Achelia* (Ammotheidae); this was corroborated by their phylogenetic tree (Figure 2 in Zehnpfennig *et al.*, 2022) in which the two putative *Rhynchothorax* specimens appeared as the sister-group of *Achelia bituberculata*; this result was confirmed by our Neighbour Joining tree (Appendix A) performed with MEGA 7 (Kumar et al., 2016) in which the two putative *Rhynchothorax* specimens were found nested within *Achelia* (BP = 99). In addition, our BLAST analysis of the CO1 sequence of the putative *Austrodecus* mitogenome published by Zehnpfennig *et al.* (2022) showed 99.94% of nucleotide identity with *Ammothea calmani* (GenBank accession number: OK583907) and *Ammothea clausi* (OK573458) and between 79.13% and 76.36% with the three CO1 sequences available in GenBank for Austrodecidae (DQ390048, MT865028 and MT865049). These results indicate that the OK623745 mitogenome was generated from *Ammothea* rather than *Austrodecus*, thus

suggesting mislabelling or DNA contamination. Because of these multiple errors, we decided not to integrate the genomic data produced by Zehnpfennig *et al.* (2022) in our analyses.

The 18S-157 alignment (Supplementary data B) contains 1,711 bp for 157 taxa. It includes 1.2% of missing data in average (excluding outgroups). A taxonomically reduced mt dataset, named mt-157 (in which 40 mt genomes were excluded) was used to make phylogenetic comparisons with the 18S-157 analysis. The mt-157 dataset includes 0.59% of missing data. The two datasets mt-157 and 18S-157 were also concatenated into a single file, mt+18S-157 dataset (length: 19,700 bp; missing data: 0.2%).

The M3-110 dataset contains 24,142 bp for 110 taxa; it was published in Ballesteros et al. (2021) as "Matrix 3". Because missing positions and gaps were not distinguished by different symbols (*e.g.,* "N" for missing data and "-" for gaps) in the alignments provided by Ballesteros et al., it was only possible to calculate the percentage of missing data + gaps. The M3-110 dataset presents 59.4% of missing data + gaps (table 2). The percentages of missing data in subdatasets mt-110, OG-110, UCE-110, and rib-110 are 52.1%, 62.0%, 67.3% and 62.1%, respectively. To allow phylogenetic comparisons between the four subdatasets, we excluded the 27 taxa for which one, two or three subdatasets were missing. The reduced dataset, named M3-83, includes therefore 72 pycnogonids and 11 outgroup taxa. The dataset M3-83 contains 54.0% of missing data + gaps. The percentages of missing data in subdatasets mt-83, OG-83, UCE-83, and rib-83 are 45.3%, 58.4%, 62.6% and 53.8%, respectively.

In order to discuss the results published in Ballesteros et al. (2021), the mt-157 and 18S-157 datasets were completed with taxa extracted from the M3-110 dataset. Due to high percentages of missing data, we included only the 54 taxa for which mt and 18S sequences do not contain more than 70% of missing data. Since only 18 taxa have less than 50% of missing data in the alignments, using a lower threshold would have resulted in the exclusion of too many taxa to be worthy. The resulting mt-211 and 18S-211 datasets contain 14.6% and 15.5% of missing data, respectively. The mt+18S-211 concatenated dataset (19,700 bp; Supplementary data C) presents 14.7% of missing data.

3.2 Phylogenetic analyses based on mt datasets

The mt trees are presented in figure 2 for mt-197 dataset and in Supplementary data D for mt-197-RY and mt-157 datasets. The results provided maximal support (PP = 1, BP = 100) for the monophyly of Pycnogonida, the grouping of Callipallenidae and Nymphonidae, the sister-group relationship between Ammotheidae and Pallenopsidae, and the monophyly of the families Ammotheidae, Colossendeidae, Endeidae, Pallenopsidae, Phoxichilidiidae and Pycnogonidae, the subfamilies Colossendeinae (*Colossendeis*, *Decolopoda*), Hedgpethiinae (*Hedgpethia*, *Rhopalorhynchus*) and Ammotheinae *sensu* Sabroux et al. (2017) (*i.e.*, including the genera *Ammothea*, *Acheliana*, *Cilunculus*, and *Sericosura*), the genera *Cilunculus, Hedgpethia*, *Nymphopsis, Sericosura*, and *Tanystylum*, as well as the "*Achelia* group *sawayai*" *sensu* Sabroux et al. (2017) (*i.e.*, including *Achelia sawayai*, *Achelia assimilis*, and *Achelia gracilis*). Within Ammotheinae, the clade grouping *Cilunculus* and *Sericosura* was highly supported in all analyses (PP = 1, BP = 98-100). All mt analyses also provided support for a sister-group relationship between Colossendeidae and Pycnogonidae (PP = 1, BP = 70-96), and the clade uniting Ammotheidae, Ascorhynchidae, Endeidae, Pallenopsidae and Phoxichilidiidae (PP = 1, BP = 42-75).

Several genera were found paraphyletic with maximal support values: *Ammothea* due to the inclusion of *Acheliana*, *Colossendeis* due to the inclusion of *Decolopoda australis*, and *Pentapycnon* and *Pycnogonum* due to the sister-group relationship between *Pentapycnon geayi* and *Pycnogonum cesairei*, and between *Pycnogonum* sp. MNHN-IU-2007-296 and *Pentapycnon* cf. *bouvieri*. The genus *Ammothella* was found paraphyletic due to the grouping of *Nymphopsis* with *Ammothella exornata* in the mt-197 and mt-157 trees (PP = 1, BP = 78-84), or the grouping of *Nymphopsis* with *Ammothella exornata*, and *Ammothella dirbergi* in the mt-197-RY tree (PP = 1, BP = 80). The genus *Ascorhynchus* was found paraphyletic due to the inclusion of *Bathyzetes* sp. (PP = 1, BP = 87-98) and *Eurycyde* sp. MNHN-IU-2012-869 (not included in mt-157 dataset) (PP = 1, BP = 87-91). Reciprocally, analyses supported *Eurycyde* paraphyly (PP = 1, BP = 87-95). Finally, *Achelia* was found polyphyletic with maximal support as the "*Achelia* group *sawayai*" was grouped with *Tanystylum,* while *Achelia bituberculata* and *Achelia* sp. (or *Achelia* sp. alone in mt-157) were grouped to Ammotheidae gen. sp. (PP = 1, BP = 93-100).

A few less supported nodes were found in most mt analyses. The group composed of Ammotheidae, Ascorhynchidae, Colossendeidae, Endeidae, Pallenopsidae, Phoxichilidiidae, and Pycnogonidae was well supported in the mt-197 and mt-157 trees (PP = 1, BP = 59-68), but less supported by the mt-197-RY analyses (PP = 0.92, not recovered in the ML bootstrap consensus). The subfamily Achelinae (*i.e.*, including the genera *Achelia*, *Ammothella*, *Nymphopsis*, *Tanystylum*) was supported in mt-197 and mt-157 trees (PP = 1, BP = 68-72), but not in the mt-197-RY tree (no robust alternative hypothesis).

Some relationships were found to be conflicting between mt datasets or between BI and ML bootstrap methods. The monophyly of Ascorhynchidae was supported by the mt-197-RY dataset (PP = 1, BP = 63) and bootstrap analyses of mt-157 and mt-197 datasets (BP = 65-73), while BI analyses of the latter two datasets rather supported the paraphyly of Ascorhynchidae (PP = 0.70-1). In addition, the families Ammotheidae, Ascorhynchidae, and Pallenopsidae were found grouped together with the mt-197-RY dataset (PP = 0.97, BP = 51), whereas ML bootstrap analyses of mt-197 and mt-157 datasets rather supported the clade uniting Ammotheidae, Endeidae, Pallenopsidae and Phoxichilidiidae (BP = 95-96; not found in the Bayesian tree).

3.3 Phylogenetic analyses based on 18S-157 and mt+18S-157 datasets

Nodes recovered in both mt-157 and 18S-157 trees represent 55.8% (87/156) of the nodes of the mt+18S-157 tree (fig. 3). In agreement with the mt-157 and mt+18S-157 trees (Rep = 1), the 18S-157 tree (Appendix B) provided strong support (PP ≥ 0.95, BP ≥ 50) for the monophyly of Pycnogonida, the families Endeidae, Pallenopsidae, Phoxichilidiidae, Pycnogonidae, the subfamilies Achelinae, Ammotheinae and Colossendeinae, and the genera *Cilunculus*, *Nymphopsis*, *Sericosura,* and "*Achelia* group *sawayai*". It also supported the polyphyly of *Achelia* and the paraphyly of *Ammothea*, *Pycnogonum* and *Pentapycnon*. Still in agreement with mt-157 and mt+18S-157, but with lower support (PP < 0.95 and/or BP < 50), the 18S-157 tree recovered the monophyly of Colossendeidae, Hedgpethiinae, *Eurycyde*, and *Tanystylum*, the sister-group relationships between Ammotheidae and Pallenopsidae, Endeidae and Phoxichilidiidae, the grouping of Callipallenidae and Nymphonidae, as well as the paraphyly of *Ascorhynchus* and *Colossendeis*.

In disagreement with the mt-157 and mt+18S-157 trees, the 18S-157 tree showed the paraphyly of *Ascorhynchus sensu lato* (*i.e.*, *Ascorhynchus* + *Bathyzetes*) due to the sister-group relationship between *Ascorhynchus seticauda* and the genus *Eurycyde* (PP = 98, BP = 83). The families Ammotheidae, Endeidae, Pallenopsidae, and Phoxichilidiidae were found closely related with the 18S-157 dataset (PP = 1, BP = 80), and with the ML bootstrap analyses of mt-157 and mt+18S-157 datasets (BP = 95/98). Other topologies found with BI were not highly supported (PP < 0.95).

3.4 Analyses of the M3-110 and M3-83 datasets

The M3-110 tree (Appendix C) was found very similar to that published by Ballesteros et al. (2021: fig. 3). Among the 109 nodes of the tree, 53 (48.6%) were recovered with the mt-110 subdataset and with at least one of the three nu subdatasets (OG-110, UCE-110 and rib-110) (see results in Supplementary data D). These nodes include the clades Callipallenidae + Nymphonidae (PP = 1, BP = 100, Rep = 0.75, MPP = 0.62, SBP = 100), Colossendeidae + Pycnogonidae + Rhynchothoracidae (PP = 1, BP = 99, Rep = 0.5, MPP = 0.42, SBP = 99.4), Pycnogonidae + Rhynchothoracidae (PP = 0.54, BP = 100, Rep = 0.5, MPP = 0.47, not found in SBP tree), *Achelia* + *Austroraptus* (PP = 1, BP = 99, Rep = 1, MPP = 1, SBP = 70.7), the families Endeidae (PP = 1, BP = 100, Rep = 0.75, MPP = 0.83, SBP = 100), Nymphonidae (PP = 1, BP = 61, Rep = 0.5, MPP = 0.35, not found in SBP tree), Pallenopsidae (PP = 1, BP = 100, Rep = 0.75, MPP = 0.8, SBP = 100), Phoxichilidiidae (PP = 1, BP = 100, Rep = 0.5, MPP = 0.48, SBP = 100), Pycnogonidae (PP = 1, BP = 100, Rep = 1, MPP = 0.93, SBP = 100), the subfamily Colossendeinae (PP = 1, BP = 100, Rep = 0.75, MPP = 0.68, not found in SBP tree) and the genera *Ammothea* (PP = 1, BP = 100, Rep = 1, MPP = 0.97, SBP = 100), *Austroraptus* (PP = 1, BP = 100, Rep = 0.5, MPP = 0.5, SBP = 100), *Boreonymphon* (PP = 1, BP = 100, Rep = 0.75, MP = 0.8, SBP = 100), *Callipallene* (PP = 1, BP = 100, Rep = 0.75, MPP = 0.75, SBP = 100), *Pallenella* (PP = 1, BP = 100, Rep = 0.75, MPP = 0.77, SBP = 100) and *Sericosura* (all support values maximal). Several taxa were found paraphyletic, including the family Callipallenidae due to the inclusion of Nymphonidae (PP = 1, BP = 89, Rep = 0.5, MPP = 0.35, not found in SBP tree), the genus *Achelia* due to the inclusion of *Austroraptus* (PP = 0.99-1, BP = 100, Rep = 0.75, MPP = 0.75-0.77, SBP = 100), the genus *Nymphon*

due to the inclusion of *Pentanymphon* (PP = 1, BP = 100, Rep = 0.75, MPP = 0.66, SBP = 100) and *Boreonymphon* (PP = 1, BP = 100, Rep = 0.75, MPP = 0.8, SBP = 100), and the genus *Colossendeis* due to the inclusion of *Decolopoda* (PP = 1, BP = 100, Rep = 0.75, MPP = 0.68, not found in SBP tree).

Some nodes that were found not supported by the mt subdataset were however recovered with two or three nu subdatasets: Pycnogonida (PP = 1, BP = 100, Rep = 0.75, MPP = 0.75, SBP = 92.3), *Achelia* + *Ammothea* + *Austroraptus* + *Sericosura* + *Tanystylum* (PP = 1, BP = 82, Rep = 0.5, MPP = 0.36, SBP = 70.7) (which is here regarded as Ammotheidae, to the difference of Ammotheidae *sensu* Bamber (2007), which includes *Paranymphon*), Austrodecidae (PP = 1, BP = 100, Rep = 0.67, MPP = 0.55, SBP = 92.3), Colossendeidae (PP = 1, BP = 99, Rep = 0.5, MPP = 0.25, not found in SBP tree) and Ammotheinae (PP = 1, BP = 98, Rep = 0.75, MPP = 0.68, SBP = 1).

The number of nodes supported by both mt and nu subdatasets was higher for the M3-83 dataset (fig. 4) than for the M3-110 dataset: 61% (50/82 nodes) *versus* 48.6% (53/109 nodes). First of all, all nodes recovered with mt and nu M3-110 subdatasets were also found with mt and nu M3-83 subdatasets, to the exceptions of the clade including Pycnogonidae and Rhynchothoracidae (PP = 1, BP = 100, Rep = 0.25, MPP = 0.46, SBP = 100) and the family Nymphonidae (PP = 1, BP = 0.62, Rep = 0.25, MPP = 0.23, not found in SBP tree). In addition, several additional nodes were recovered by both mt and nu subdatasets, such as the subfamily Ammotheinae (PP = 1, BP = 96, Rep = 0.75, MPP = 0.15, SBP = 100), the family Ammotheidae (PP = 1, BP = 100, Rep = 0.75, MPP = 0.7, SBP = 100) and its sister-group relationship with Pallenopsidae (PP = 1, BP = 100, Rep = 0.5, MPP = 0.5, SBP = 100).

3.5 Phylogenetic analyses based on mt+18S-211 dataset

The tree reconstructed from the mt+18S-211 dataset (fig. 5, Appendix D) presents 85/210 nodes supported by both mt-211 and 18S-211 subdatasets (40.5% of total number of nodes). They are in agreement with the mt+18S-157 tree, and include the monophyly of Pycnogonida, of the families Ammotheidae (BP = 100), Colossendeidae (BP = 100), Endeidae (BP = 100), Pallenopsidae (BP = 100), Phoxichilidiidae (BP = 100), and Pycnogonidae (BP = 100), of the clades Nymphonidae +

Callipallenidae (BP = 100), Ammotheidae + Pallenopsidae (BP = 100), Endeidae + Phoxichilidiidae (BP = 83), Colossendeidae + Pycnogonidae (BP = 98), Ammotheidae + Endeidae + Pallenopsidae + Phoxichilidiidae (BP = 96), *Ascorhynchus* + *Eurycyde* (BP = 63); the subfamilies Colossendeinae and Hedgpethiinae (BP = 100); the genera *Austroraptus* (BP = 100), *Cilunculus* (BP = 100), *Eurycyde* (BP = 100), *Nymphopsis* (BP = 100), *Parapallene* (BP = 100), *Sericosura* (BP = 100). Like in the mt+18S-157 tree (and unlike M3-83), the group uniting Ammotheidae, Ascorhynchidae, Endeidae, Pallenopsidae and Phoxichilidiidae was also recovered monophyletic (BP = 95). In addition to mt+18S-157 supported nodes, mt+18S-211 recovered the clade uniting Ammotheidae, Ascorhynchidae, Callipallenidae, Endeidae, Nymphonidae, Pallenopsidae and Phoxichilidiidae (BP = 56).

Unlike in M3-83 and M3-110 trees (figs 4, 5 and Appendix C), the genus *Tanystylum* was found paraphyletic due to the inclusion of *Tanystylum californicum* within *Ammothea* (BP = 93) and the family Ascorhynchidae *sensu lato* was found monophyletic since *Nymphonella* were found as sister-group to *Ascorhynchus* + *Eurycyde* (BP = 93).

Finally, the mt+18S-211 dataset enables to find additional relationships among groups included only either in mt+18S-157 or M3-110 datasets: *Achelia* sp. MNHN-IU-2013-18597 was found as sister-group to the clade including *Achelia transfugoides*, *Achelia spicata* and *Austroraptus* (BP = 100), and *Callipallene* clustered with *Pallenoides* (BP = 100).

4. DISCUSSION

4.1 Effect of missing data on the phylogenetic signal

To study node congruence between two independent markers (mitogenome and 18S nu gene), we constructed the mt+18S-157 dataset by maximizing the taxonomic sampling and limiting the percentage of missing data. The selection of taxa was made from a bank of about 600 DNA extracts fed by the multiple recent MNHN expeditions in various tropical, temperate, and Antarctic localities (table 1). We retained only DNA samples with low levels of missing data for both mitogenome and 18S gene (maximum percentage of missing data per sequence: 14.8% in mt-197 and 30.3% in 18S-157; table 2).

Since the mitogenome and 18S nu gene are independent phylogenetic markers, all nodes supported by both of them can be considered as reliable (*e.g.*, Arabi et al., 2010; Sabroux et al., 2017; Xue et al., 2017). Because the mitogenome evolves more rapidly than the nuclear genome (Allio et al., 2017), mutational saturation and long branch attraction artefacts due to convergence in asymmetric base composition bias can be highly misleading for inferring deep relationships (Hassanin et al., 2005; Hassanin, 2006). These issues can be however addressed in part by degenerating nucleotide sequences into a binary purine/pyrimidine coding (Hassanin, 2006; Simmons, 2017), so that a deep node recovered in both 18S-157 and mt-197-RY trees can be also considered as reliable.

Ballesteros et al. (2021) performed targeted capture for a set of markers belonging to the four subdatasets mt-110, OG-110, rib-110 and UCE-110. The method consists in using DNA probes from one or several taxa to capture by DNA hybridization homologous sequences found in a DNA extract. The capture yields are good if the targeted loci are well-conserved among studied taxa, generating large datasets with relatively low levels of missing data (*e.g.,* 1,500 loci and 13% of missing data in McCormack et al., 2013; and about 1,500 loci for 9% of missing data in Hugall et al., 2016). However, the M3-110 dataset of Ballesteros et al. (2021) contains very high levels of missing data (table 2). Our analyses show that even a small reduction of missing data between M3-110 and M3-83 datasets (59.4% to 54%) results in a significant increase in the percentage of reliable nodes, *i.e.*, repeated with both mt and nu subdatasets (from 53/109 = 48.6% in the Bayesian M3-110 tree to 50/82 = 61% in the M3-83 Bayesian tree; fig. 4 and Appendix C). Similarly, the impact of missing data on node congruence can be also addressed by comparing mt+18S-211 and mt+18S-157 datasets. The level of missing data is much higher in the mt+18S-211 dataset (14.7%) than in the mt+18S-157 dataset (0.2%). As expected, we observe that the percentage of reliable nodes is much lower in the mt+18S-211 tree (85/210 reliable nodes = 40.5%; Appendix D) than in the mt+18S-157 tree (87/156 reliable nodes = 55.8%; fig. 3). These results confirm previous studies indicating that the noise induced by missing data can have a strong negative effect on the extraction of phylogenetic signal (Roure et al., 2013, Philippe et al. 2017, Smith et al., 2020).

4.2 Subdivision of Ammotheidae into Ammotheinae and Achelinae

Based on an analysis of CO1 and 18S markers, Sabroux et al. (2017) already proposed to divide the family Ammotheidae into two subfamilies: Ammotheinae (including the genera *Ammothea*, *Acheliana*, *Sericosura* and *Cilunculus*) and Achelinae (including *Achelia*, *Tanystylum*, *Nymphopsis*, *Ammothella*). In agreement with previous studies (Arabi et al., 2010; Sabroux et al., 2017; Ballesteros et al., 2021), we found a strong support for the monophyly of Ammotheinae (the node was found in both mt+18S-157 and M3-83 trees). The monophyly of Achelinae was found supported in mt-157, mt-197, mt-197-RY, mt+18S-157, 18S-157 BI and ML, and 18S-211 ML analyses. However, the subfamily was found paraphyletic with the M3-83 dataset, as *T. californicum* and *T. orbiculare* appeared more closely related to Ammotheinae (PP = 1, BP = 86). However, this node cannot be considered as reliable as it was found only by one of the four subdatasets and without support (OG-83; PP = 0.55).

In both mt-211 and mt+18S-211 datasets, *Tanystylum orbiculare* grouped with other *Tanystylum* representatives, whereas *Tanystylum californicum* appeared as nested within *Ammothea* (Appendix D and Supplementary data D). *Tanystylum californicum* also shows 15 unambiguous mitochondrial synapomorphies shared with nine *Ammothea* (including *Acheliana*) (Appendix E). By contrast, the single unambiguous 18S synapomorphy of *Ammothea* species was not found in the *Tanystylum californicum* 18S sequence. This suggests that the sample of *T. californicum* provided by Ballesteros et al. (2021; GenBank accession numbers: MT864817, MT864903, MT864961, MT865014, MT865089, MT865152, MT865299, MT865306, MT865327, MT865390) has been contaminated (at least partially) by *Ammothea* DNA.

Furthermore, mt-197, mt-157 and 18S-157, mt+18S-157 and mt+18S-211 datasets all support the genus *Achelia* to be polyphyletic, as the "*Achelia* group *sawayai*" (Sabroux et al., 2017) was sister-group to *Tanystylum*. Other *Achelia* representatives were grouped with *Austroraptus* when it was included in the dataset (*i.e.*, in M3-110, M3-83 and mt+18S-211). The taxonomic status of *Achelia* needs therefore to be revised. The type species of the genus, *Achelia echinata* (Child, 1998) was not included in the analyses, but its morphology suggests that it belongs to the "*Achelia* group *sawayai*". Further studies

addressing the *Achelia* phylogeny should therefore include *A. echinata*, as well as more species of *Achelia* in order to determine if a new genus has to be described.

4.3 Monophyly of Ascorhynchidae

According to Bamber et al. (2022), the family Ascorhynchidae is represented by eight genera: *Ascorhynchus*, *Bathyzetes*, *Boehmia*, *Calypsopycnon*, *Eurycyde*, *Heterofragilia*, *Nymphonella* and *Pycnofragilia*. Four of them were included in our study: *Ascorhynchus*, *Bathyzetes*, *Eurycyde* and *Nymphonella*. *Bathyzetes* was unambiguously recovered within *Ascorhynchus* in all analyses, which calls for a taxonomic reassessment of the former. The grouping of *Eurycyde* with *Ascorhynchus* + *Bathyzetes* is supported by the mt+18S-211, 18S-157 and mt-197-RY datasets.

The position of *Nymphonella* is more difficult to address. This genus exhibits a very unusual morphology, including unmatched palp and leg articulation (*e.g.,* Guille and Soyer, 1967). Before adult stage, *Nymphonella tapetis* lives as an ectoparasite infesting the gills and mantle of different bivalve species including the economically important clam *Ruditapes philippinarum* (Yoshinaga et al., 2011). Though parasitism among sea spider larvae is common (Brenneis et al., 2017), larval ectoparasitism on bivalves seems to be fairly rare among sea spiders (though possibly not unique to *Nymphonella*; see Tharme et al., 1996) which may explain the very divergent morphology of *Nymphonella* (Tomiyama et al., 2015). Since parasitism may result in higher substitution rates (Hassanin, 2006), it may explain difficulties to determine its phylogenetic position. In their study dedicated to this bizarre genus, Chow et al. (2012) recovered *Nymphonella tapetis* as nested within *Ascorhynchus* with strong support relying on 18S gene (PP = 1, BP = 99). By contrast, Ballesteros et al. (2021) recovered *Nymphonella* as sister-group of the clade Callipallenidae + Nymphonidae, in agreement with the classification of Bamber (2007), though this topology is not found in M3-83 tree and is supported in M3-110 tree only in BI analyses (PP = 1, BP < 50). It is neither found with subdatasets, except with rib-110 and rib-83 (but with low support; PP = 0.66). By contrast, *Nymphonella* was found related to the clade uniting *Ascorhynchus*, *Bathyzetes* and *Eurycyde* in the mt+18S-211 tree with strong support (BP = 93). However, in absence of support from the mt subdataset mt-211 (Rep = 0.5), it is not yet possible to

consider this topology as reliable. The lack of support may be linked to the high percentage of missing data for *Nymphonella tapetis* (54.7%) in the M3-110 dataset.

4.4 Interfamilial relationships within Pycnogonida

Analysis of mt/nu congruence in mt+18S-157 and M3-83 datasets showed that several interfamilial nodes can be regarded as reliable (fig. 5). The group uniting Ammotheidae and Pallenopsidae, hereafter referred as superfamily Ammotheoidea Dohrn, 1881, was found monophyletic with both nu and mt markers in both mt+18S-157 and M3-83 datasets. It was already proposed by Arabi et al. (2010) (PP = 1, not found in ML), although with low support, and in Ballesteros et al. (2021) ($94 \leq BP \leq 100$). Similarly, the group composed of Callipallenidae and Nymphonidae, hereafter referred as superfamily Nymphonoidea Wilson, 1878 (differing from Bamber, 2007 definition by excluding Pallenopsidae) was supported by both mt+18S-157 and M3-83 datasets and recovered with mt and nu subdatasets. This superfamily was found monophyletic in most previous studies (Arango and Wheeler, 2007; Nakamura et al., 2007; Arabi et al., 2010; Ballesteros et al., 2021). The superfamily uniting Endeidae and Phoxichilidiidae, already described by Bamber (2007) as Phoxichilidioidea Sars, 1891, was found monophyletic with both mt and nu markers of the mt+18S-157 dataset. It was already suggested by Arabi et al. (2010; PP = 0.86, not found in the bootstrap ML analysis), and in some analyses of Ballesteros et al. (2021; $60 \leq BP \leq 87$). Finally, the clade including Colossendeidae, Pycnogonidae, and Rhynchothoracidae, hereafter referred as superfamily Colossendeoidea Jarzynsky, 1870 was supported by both mt and nu markers of the M3-83 dataset, and the clade grouping Colossendeidae with Pycnogonidae was also found in the mt+18S-157 tree. Colossendeoidea was previously recovered by Nakamura et al. (2007) (PP = 0.88, $BP < 50$), and by Ballesteros et al. (2021) with maximal BP value.

In addition, the grouping of Pycnogonidae and Rhynchothoracidae was recovered with two M3-110 subdatasets, one mt and one nu (mt-110 and UCE-110). The same pattern was obtained with strong support in Nakamura et al. (2007; PP = 1, BP = 86) and Arabi et al. (2010; PP = 1, BP = 89) based on 18S analyses, and in Ballesteros et al. (2021). However, this node was only recovered with the UCE-83 subdataset (BP = 97; Rep = 0.25). The M3-83 and M3-110 datasets include only one

Rhynchothoracidae with 83.3% of missing data. Since there is no alternative hypothesis supported by other subdatasets, we therefore regard the node Pycnogonidae + Rhynchothoracidae as the most likely. Additional, more complete DNA data, including the full mitogenome, should be sequenced for Rhynchothoracidae to confirm their sister-group relationship with Pycnogonidae.

Ammotheoidea and Phoxichilidioidea were found grouped together with M3-83 and M3-110 datasets (PP = 1, BP = 82/51), similarly to Nakamura et al. (2007) (PP = 1, BP = 95), Arabi et al. (2010) (PP = 0.93, BP = 73 based on 18S) and some analyses of Ballesteros et al. (2021; 62 < BP ≤ 89). In our analyses, this result was only found by the UCE-83 subdataset (PP = 0.63); it was also supported by 18S-157 (PP = 1, BP = 80), while the mt-197-RY dataset rather supported the grouping of Ammotheoidea with Ascorhynchidae (PP = 0.97, BP = 51). Strikingly, there is a strong incongruence between ML and BI analyses with mt-197, mt-157 and mt+18S-157 datasets: ML bootstrap consensus analyses provided a strong support for Ammotheoidea + Phoxichilidioidea (BP = 95-98), whereas BI analyses supporting the grouping of Ammotheoidea with Phoxichilidiidae and Ascorhynchidae (mt-197; PP = 1). Therefore, repeatability of the node Ammotheoidea + Phoxichilidioidea was supported at least in ML bootstrap consensus analyses, and was also confirmed in mt-211, 18S-211 and mt+18S-211 analyses (BP = 24-96). Support from alternative hypothesis by mt-197-RY dataset was instead relatively low and could be linked with the loss of signal while degenerating to purine/pyrimidine. Consequently, we regard the node Ammotheoidea + Phoxichilidioidea as likelier than alternative hypotheses.

Other, interfamilial relationships were insufficiently supported over the different analyses: Ascorhynchidae were recovered as sister clade to Ammotheoidea in mt-197-RY (PP = 0.97, BP = 51) but this result was not strongly supported in mt-157 and mt+18S-157 analyses (see above) and is not recovered in the 18S-157 tree. And while 18S-157 and M3-110 analyses supported a clade uniting Ammotheoidea, Ascorhynchidae, Nymphonoidea, and Phoxichilidioidea (PP = 0.87-1, BP = 40-64), mt-197 and M3-83 rather supported the clustering of Ammotheoidea with Ascorhynchidae, Austrodecidae, Colossendeoidea and Phoxichilidioidea.

We suspect that difficulties to recover interfamilial relationships are linked to the long branch of Nymphonoidea, as also recovered by Arabi et al. (2010), Sabroux et al. (2017) and Ballesteros et al.

(2021); and to the long branch of Pycnogonida within the Arthropoda tree. The consequence is that interfamilial relationships are expected to be impacted by long branch attraction toward the root. Such artefactual attraction could be observed in the 18S tree published by Nakamura et al. (2007) in which Pycnogonida root was misplaced within the genus *Ascorhynchus* (Arabi et al., 2010).

Another possible explanation to the weak support to interfamilial relationships lies in the hypothesis that sea spiders underwent a radiation event. Indeed, the sea spiders fossil record provides evidence for a transition between a mostly or totally non-Pantopoda diversity during Palaeozoic to a Pantopoda-only diversity in Mesozoic (Bergström et al., 1980; Poschmann and Dunlop, 2006; Kühl et al., 2013, Charbonnier et al., 2007; Sabroux et al., 2019a). This may signify that an extinction event wiped-out most of sea spider diversity, and that extant sea spiders derived from a relict clade (Charbonnier et al., 2007; Arabi et al., 2010; Sabroux et al., 2019a). This clade, designated by Hedgpeth (1954, 1978) as order Pantopoda, is well characterized among Pycnogonida total group by the reduction of abdomen as an unsegmented terminal tagma. Difficulties to recover interfamilial relationships with molecular data may therefore result from a rapid diversification (*i.e.*, radiation) of Pantopoda. For the moment, the Palaeozoic and Mesozoic fossil records are separated by a hiatus of 250 million years, which impedes to determine how rapid was the observed fauna transition, nor when it occurred.

The radiation hypothesis was recently contradicted by the chronogram published by Ballesteros et al. (2021), which showed that Pantopoda diversification started as early as Ordovician and followed from that point a monotonic process of slowing diversification. Some calibration points used in this study are, however, problematic: i) most of them are poorly relevant to date nodes within Pantopoda as they belong to outgroup taxa: eight are within Euchelicerata, three others are within Mandibulata, and another one concerns the first emergence of arthropods in the fossil record; ii) *Palaeopycnogonides gracilis* was used to set the minimum age of Ammotheoidea, while its taxonomic assignation to Ammotheidae was regarded as doubtful by some authors because of the absence of characteristic cephalic appendages (Bamber, 2007; Charbonnier et al., 2007; Sabroux et al., 2019a); iii) the Silurian fossil *Haliestes dasos* (Siveter et al., 2004) was used to set a minimum age for Pantopoda; but this hypothesis is unsupported: *Haliestes dasos* abdomen was possibly segmented (Siveter et al., 2004),

while reduction of abdomen as an unsegmented terminal tagma is regarded as the main synapomorphy of Pantopoda (Bergström et al., 1980; Sabroux et al., 2019). Ballesteros et al. (2021) specified that the tree dating estimation only slightly changes when all sea spider fossils or *Haliestes dasos* alone were removed from calibration points. Their results suggest therefore that the 12 outgroup calibration points have an overwhelming effect on tree-dating. Ideally, the finding of new Pantopoda fossils would increase the number of calibration points in tree dating and provide more robust results. Unfortunately, sea spider fossils are peculiarly rare (Sabroux et al., 2019a). In the meantime of a new finding, testing the radiation hypothesis and unravelling Pycnogonida phylogenetic tree will benefit from the inclusion of new taxa with low levels of missing data (*e.g.*, Austrodecidae) and new molecular markers with higher resolving power for deep nodes than the mitogenome or 18S gene.

4.5 Evolution of Pycnogonida morpho-anatomy and behaviour

Sea spiders have an unusual body-plan (fig. 6A) that makes interpretation of homologies with other arthropods difficult (*e.g.*, Dunlop and Lamsdell, 2017). They are generally regarded as chelicerates (Chelicerata) (see a review in Dunlop and Arango, 2005), but they lack the subdivision into a prosoma and an opisthosoma, and present instead division into cephalon (containing the cephalic appendages and first pair of walking legs), trunk (three remaining pairs of walking legs, rarely four or five) and abdomen (a one-segmented, reduced tagma in extant species). Extant sea spiders present cylindrical legs (generally four pairs) composed of nine podomeres (including the terminal claw). Compared to other extant chelicerates, sea spiders have at least one additional pair of appendages, that may be homologous with chilariae of horseshoe crabs (Manuel et al., 2006; Dunlop and Lamsdell, 2017).

4.5.1 Cephalic appendages lability

Cephalic appendages are chelifores, palps and ovigers (fig. 6). While they are always present at larval stage (fig. 6B; see also Brenneis et al., 2017), they are not recovered in adults of different families. For example, adults among Nymphonidae present all the three pairs of appendages, while Phoxichilidiidae

lack the palps and have ovigers in males only; Rhynchothoracidae instead miss chelifores, and Pycnogonidae miss both chelifores and palps and have ovigers only in males, or no ovigers at all depending on species (fig. 6C). Until the late 1990's, it was commonly admitted that pycnogonid underwent a body-plan "simplification" in loosing progressively the cephalic appendages (Munilla, 1999; Stock, 1994). However, even under this hypothesis, several phylogenetic trees based on the morphological characters of cephalic appendages are equiparsimonious, and all of them include convergences.

Although resolution of sea spider phylogeny is yet limited to the four superfamilies Ammotheoidea, Colossendeoidea, Nymphonoidea, and Phoxichilidioidea (fig. 5), the lability of cephalic appendages within these clades shed light on their homoplasticity (fig. 6C). In our study, Ammotheidae and Pallenopsidae are grouped into superfamily Ammotheoidea, while Ammotheidae are traditionally considered as close to Ascorhynchidae (e.g., Bamber, 2007; Hedgpeth, 1947; Stock, 1994) based on the reduction of chelae to bulges in most of adults and the presence of developed palps, while Pallenopsidae are often regarded as close to Callipallenidae (*e.g.*, Bamber, 2007; Hedgpeth, 1947) or Phoxichilidiidae (*e.g.*, Stock, 1994) based on the presence of developed chelifores and the reduction (or complete absence) of palps. Therefore, reduction of palps and of chelae occurred at least twice in Pantopoda evolution. The superfamily Phoxichilidioidea, comprising Phoxichilidiidae and Endeidae, is supported by the absence of palps in both sexes and of ovigers in females (as already noticed by Ballesteros et al., 2021 and Bamber, 2007). Conversely, only Endeidae have no chelifores in adults, a character shared with Austrodecidae, Colossendeidae, Pycnogonidae, and Rhynchothoracidae. This is most probably a convergence rather than a symplesiomorphy, since the presence of chelifores in Phoxichildiidae is shared with Ammotheoidea. Nymphonoidea show high variability on the presence/absence of palps (present, absent or absent in females only) as well as the number of podomeres (from one to five podomeres). Finally, within Colossendeoidea, Rhynchothoracidae and Colossendeidae share the presence of palps and ovigers in both sexes (as also found in *Pantopipetta*) while Pycnogonidae have no palps, and no ovigers in females as in males of some species. All the three families lack chelifores, except in the colossendeid genera *Decolopoda* and *Dodecolopoda*.

As for now it is not possible to know whether these appendages were present or absent in the common ancestor of Pantopoda. Palaeozoic fossils are of little help since their relationship with Pantopoda are not well understood. The fact that cephalic appendages are, as far as we know, always retained at larval stage (see Brenneis et al., 2017) and during juvenile stages makes reacquisition in adult through retention of larval characters as likely as loss.

4.5.2 Polymerous sea spiders

Among sea spiders, there are four decapodous genera (*Decolopoda*, *Pentacolossendeis*, *Pentanymphon*, *Pentapycnon*) and two dodecopodous genera (*Dodecolopoda*, *Sexanymphon*), for a total of nine species. *Decolopoda*, *Dodecolopoda* and *Pentacolossendeis* belong to Colossendeidae, *Pentanymphon* and *Sexanymphon* to Nymphonidae, *Pentapycnon* to Pycnogonidae (Fig. 6C). These sea spiders are generally called "polymerous" (Arnaud and Bamber, 1987). Relying on the hypothesis of progressive reduction of appendages sets of sea spiders through time, Bouvier (1910) suggested that polymery was plesiomorphic among sea spiders, as polymerous species diverged early within their respective lineages.

Our study includes three polymerous genera, *Decolopoda*, *Pentanymphon* and *Pentapycnon*. In disagreement with Bouvier's hypothesis, our analyses supported a nested placement of the three polymerous taxa: *Decolopoda* appeared as nested within *Colossendeis* in mt+18S-157 tree, *Pentanymphon* within *Nymphon* in M3-83 tree, and *Pentapycnon* within *Pycnogonum* in the mt+18S-157 tree. *Pentapycnon* is even found polyphyletic. These results are all robust and recovered with both mt and nu datasets. They indicate that the polymerous state is derived and that these species should not be included in a different genus from octopodous species. This was already foreshadowed by Hedgpeth (1947) who pointed out that *Pentapycnon* and *Pentanymphon* polymerous species were morphologically very close to octopodous species beyond their additional pair of legs. To this regard, *Decolopoda* and *Dodecolopoda* are the exception, since they markedly differ from all other colossendeids by presenting functional chelifores at adult stage. Polymery therefore occurred several times (at least four times) during Pantopoda evolution.

4.5.3 Paternal care and cement gland

In most sea spiders families, males perform parental care for eggs, and sometimes for larvae and juveniles (Arnaud and Bamber, 1987). In most species, larvae are free-living and leave males after hatching in order to find a host to parasitize, although in some exceptions (Callipallenidae and some Nymphonidae, Pallenopsidae or Ammotheidae) larvae are lecithotrophic and stay until a later stage on the father (Brenneis et al., 2017). Paternal care was never observed among Colossendeidae and Austrodecidae (Stock, 1958; Arango and Wheeler, 2007), despite important collections and/or *in situ* observations of Colossendeidae – we actually do not know any larva from these families.

In sea spiders, paternal care can be linked with one specific organ: the cement gland. The cement gland is a male-only organ generally found in single or multiple instars on femorae. It produces a cement that enables to past eggs on the father (Arnaud and Bamber, 1987). Most of the time, eggs are pasted on ovigers, but in *Pycnogonum* species without ovigers (subgenus *Nulloviger*) the eggs are pasted directly on the ventral surface (*e.g.*, Staples, 2002).

Cement glands have been identified in the families Ammotheidae, Endeidae, Pallenopsidae, Phoxichilidiidae and Rhynchothoracidae, and are absent in Colossendeidae. The cement glands are not known from every species of Ascorhynchidae and Nymphonoidea (*e.g.*, Arango & Wheeler 2007), but we suppose this is more due to the fact that they are inconspicuous and not described. Similarly, cement glands have not been observed for Pycnogonidae, although they should exist since species of this family do present egg-pasting behaviour. The coxal glands (Staples, 2002; Lee and Kim, 2020) could have the role of cement glands. However, they have not yet been studied, and were not observed in all species. More surprisingly, Austrodecidae have structures on femorae identified as cement gland spurs (*e.g.*, Child, 1994) despite the absence of observed paternal care. Two evolutionary scenarios can be proposed: either i) the role of cement glands in paternal care was ancestral within Pantopoda, and this behaviour has been lost twice, in Austrodecidae and in Colossendeidae; or conversely ii) the ancestral function of cement glands in Pantopoda was not linked to paternal care and the glands have been subsequently recruited for this role, either twice independently, or only once in the clade excluding

Austrodecidae and then lost in Colossendeidae (depending on the position of Austrodecidae in Pantopoda phylogeny). However, it is not to be excluded that Austrodecidae paternal care exists without being observed or, on the contrary, that the identification of cement glands in Austrodecidae is erroneous, so that the paternal care behaviour was always linked with cement glands, and these were lost all together and only once in Colossendeidae. It was suggested that absence of parental behaviour in Colossendeidae could be linked with a complete change in their developmental biology, including direct development (Arnaud & Bamber, 1987), but this has yet to be demonstrated.

4.5.4 Morphological synapomorphies supporting interfamilial relationships

Within the superfamily Phoxichilidioidea, comprising Phoxichilidiidae and Endeidae, females have no ovigers, except in a few species, like *Anoplodactylus* cf. *californicus*, which possess residual appendages. The nested position of *A.* cf. *californicus* within *Anoplodactylus*, supported by both the mitogenome and 18S, suggests that these residual appendages are a secondary reacquisition. In males of Phoxichilidioidea, the strigilis (a hook-like assemblage formed by ovigeral podomeres 7 to 10; fig. 6A) is completely missing but the first strigilis podomere (*i.e.*, seventh oviger podomere). This podomere is further lost (or fused with the sixth oviger podomere) in some species of *Anoplodactylus* (*e.g.*, *Anoplodactylus monotrema*; see Stock, 1979). Although the loss of ovigers in females and the loss of strigilis can be regarded as synapomorphies of Phoxichilidioidea, it must be noted that they may have occurred by convergence in Pycnogonidae.

Within the superfamily Ammotheoidea, Ammotheidae and Pallenopsidae share the same general shape of ovigers strigilis, with reduced number (or total absence) of compound spines, different and uneven shape of strigilis podomeres, and loss of the terminal claw. In these two families, ovigers show a marked sexual dimorphism, with females having unfunctional strigilis, and shorter fourth and fifth podomeres (though this latter character is also found in Nymphonoidea). Due to the absence of the strigilis in Phoxichilidioidea, we cannot exclude that these characters were present in the common ancestor of Ammotheoidea and Phoxichilidioidea.

In the superfamily Nymphonoidea, which unites Callipallenidae and Nymphonidae, ovigers have uniform structures with ten podomeres, though the terminal spine is missing in several callipallenid genera. Chelifores are well formed, often denticled, and always three-articled (scape – palm – dactylus). However, these characters are generally shared with other taxa (three-articled chelifores have instars in most of chelifore bearing families) and are possibly plesiomorphic. More specific to Nymphonoidea is the narrowed, and more or less elongated preocular neck, followed by a broaden basis to chelifores articulation; but a rather similar preocular neck can also be observed in some *Ascorhynchus* species, *e.g.*, *Ascorhynchus glaberrimus* (Kim and Hong, 1986). Nymphonidae and Callipallenidae also share the presence of a single row of compound spines on the strigilis (it is found in loose distribution in Ammotheidae, and in several rows or in field in Ascorhynchidae and Colossendeidae; Bamber, 2007) though no plesiomorphic state of Pantopoda can be readily identified. Another common feature of Nymphonoidea is the reduction of palps. However, this character shows various patterns: in Nymphonidae, the palps are 5-articled in both males and females while in Callipallenidae, the palps show between 0 and 4 articles in males, and are absent in females.

The superfamily Colossendeoidea, which groups Colossendeidae, Pycnogonidae, and Rhynchothoracidae, cannot be characterized by any unambiguous synapomorphy. In Colossendeoidea, the chelifores are absent in most species, but are present in the genera *Decolopoda* and *Dodecolopoda*. The ovigers show very different patterns among the three families: Colossendeidae have 11-articled ovigers (including the terminal claw) with strigilis bearing fields of compound spines (Colossendeinae: *e.g.*, Dietz et al., 2015, 2013) or several rows of spines (Hedgpethinae); Rhynchothoracidae have 11-articled ovigers with a specific strigilis shape bearing few spines and a ventral lamella on the 10th podomere; and Pycnogonidae have ovigers with variable number of podomeres (5 to 10, 0 in females as well as in males of some species; Bamber, 2007) and few spines. The palps are present in Colossendeidae and Rhynchothoracidae, but not in Pycnogonidae. As pointed out by Munilla (1999) and Arabi et al. (2010), a unique pair of gonopore on fourth legs is found in females of Pycnogonidae and Rhynchothoracidae, but not in females of Colossendeidae.

CRediT author statement

Romain Sabroux: Conceptualization, Data curation, Formal analyses, Investigation, Project administration, Validation, Visualization, Writing – original draft, Writing – review and editing

Laure Corbari: Data curation, Funding acquisition, Specimen curation, Resources, Supervision, Writing – original draft, Writing – review and editing

Alexandre Hassanin: Conceptualization, Funding acquisition, Methodology, Formal analyses, Supervision, Writing – original draft, Writing – review and editing

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The specimens originates from various deep sea cruises, conducted respectively by MNHN and *Institut de Recherche pour le Développement* (IRD) as part of the Tropical Deep-Sea Benthos program and organized by P. Bouchet, S. Samadi and L.C.: KANADEEP 1 (10.17600/17003800), KANACONO (10.17600/16003900), SALOMONBOA 3 (10.17600/7100070), MADANG 2012 (10.17600/12100140), MUSORSTOM 8 (10.17600/94100040), BIOCAL (10.17600/85002911), EXBODI (10.17600/11100080), BIOMAGLO (10.17600/17004000), MD208-Walters Shoal (10.17600/17002700); and *Planète Revisitée* program (http://laplaneterevisitee.org): ATIMO VATAE (10.17600/10110040), KARUBENTHOS 2012, PAPUA NIUGINI (10.17600/18000841), KAVIENG 2014 (10.17600/14004400), GUYANE 2014, MADIBENTHOS. Funders and sponsors of the expeditions include the French Ministry of Foreign Affairs, the Total Foundation, Prince Albert II of Monaco Foundation, Stavros Niarchos Foundation, and Richard Lounsbery Foundation. Specimens from hydrothermal vents were provided by the IFREMER: BICOSE cruise (financial support was

provided by the Eramet and Technip companies). The CEAMARC cruise on RSV Aurora Australis was funded by the Australian Antarctic Division, the French polar institute IPEV, CNRS and the ANR project ANTFLOCKs (G. Lecointre). Insightful comments from the two anonymous reviewers contributed to improve the previous version of the manuscript.

Funding

The project was funded by the *Actions Thématiques du Muséum* (ATM) and the *Insitut de Systématique, Evolution, Biodiversité* (ISYEB). The funding sources were not involved in the study design, in the collection, analysis, or interpretation of data, in the writing of the report or in the decision to submit the article for publication.

Data accessibility

All new sequences were made available on GenBank (see accession number in table 1). Supplementary data A-D including alignments and additional trees were made available on OSF (https://osf.io/hztcm/). [data will be released upon reception of the proofs]

Supplementary material

APPENDIX A: Neighbour Joining tree based on the alignment of mt-197 dataset with data from Zehnpfennig *et al.* (2022).

APPENDIX B: Bayesian tree based on 18S-157 dataset.

APPENDIX C: Bayesian tree based on M3-110 dataset.

APPENDIX D: ML bootstrap consensus tree based on mt+18S-211 dataset.

APPENDIX E: Shared synapomorphies between *Tanystylum californicum* and *Ammothea* (including *Acheliana*) in mt+18S-211 alignment provided in Supplementary data C.

SUPPLEMENTARY DATA A: mt-197 alignment in nexus format.

SUPPLEMENTARY DATA B: 18S-157 alignment in nexus format.

SUPPLEMENTARY DATA C: mt+18S-211 alignment in nexus format.

SUPPLEMENTARY DATA D: Trees from ML bootstrap consensus and BI analyses not presented in figures or appendices, in Newick format.

BIBLIOGRAPHY

Allio, R., Donega, S., Galtier, N., Nabholz, B., 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Mol. Biol. Evol. 34, 2762–2772. https://doi.org/10.1093/molbev/msx197

Arabi, J., Cruaud, C., Couloux, A., Hassanin, A., 2010. Studying sources of incongruence in arthropod molecular phylogenies: Sea spiders (Pycnogonida) as a case study. Comptes Rendus - Biol. 333, 438–453. https://doi.org/10.1016/j.crvi.2010.01.018

Arabi, J., Judson, M.L.I., Deharveng, L., Lourenço, W.R., Cruaud, C., Hassanin, A., 2012. Nucleotide composition of CO1 sequences in Chelicerata (arthropoda): Detecting new mitogenomic rearrangements. J. Mol. Evol. 74, 81–95. https://doi.org/10.1007/s00239-012-9490-7

Arango, C.P., 2003. Molecular approach to the phylogenetics of sea spiders (Arthropoda: Pycnogonida) using partial sequences of nuclear ribosomal DNA. Mol. Phylogenet. Evol. 28, 588–600. https://doi.org/10.1016/S1055-7903(03)00069-1

Arango, C.P., 2002. Morphological phylogenetics of the sea spiders (Arthropoda: Pycnogonida). Org. Divers. Evol. 2, 107–125. https://doi.org/10.1078/1439-6092-00035

Arango, C.P., Wheeler, W.C., 2007. Phylogeny of the sea spiders (Arthropoda, Pycnogonida) based on direct optimization of six loci and morphology. Cladistics 23, 255–293.

https://doi.org/10.1111/j.1096-0031.2007.00143.x

Arnaud, F., 1978. A new species of *Ascorhynchus* (Pycnogonida) found parasitic on an opisthobranchiate mollusc. Zool. J. Linn. Soc. 63, 99–104.

Arnaud, F., Bamber, R.N., 1987. The Biology of Pycnogonida. Adv. Mar. Biol. 24, 1–96.

Audzijonyte, A., Väinölä, R., 2006. Phylogeographic analyses of a circumarctic coastal and a boreal lacustrine mysid crustacean, and evidence of fast postglacial mtDNA rates. Mol. Ecol. 15, 3287–3301. https://doi.org/10.1111/j.1365-294X.2006.02998.x

Ballesteros, J.A., Setton, E.V.W., Santibáñez-López, C.E., Arango, C.P., Brenneis, G., Brix, S., Corbett, K.F., Cano-Sánchez, E., Dandouch, M., Dilly, G.F., Eleaume, M.P., Gainett, G., Gallut, C., McAtee, S., McIntyre, L., Moran, A.L., Moran, R., López-González, P.J., Scholtz, G., Williamson, C., Woods, H.A., Zehms, J.T., Wheeler, W.C., Sharma, P.P., 2021. Phylogenomic resolution of sea spider diversification through integration of multiple data classes. Mol. Biol. Evol. 38, 686–701. https://doi.org/10.1093/molbev/msaa228

Bamber, R.N., 2013. Deep-water Pycnogonida from recent cruises to Papua New Guinea and Melanesia, with an appendix of new records from Polynesia and descriptions of five new species. Zoosystema 35, 195–214. https://doi.org/10.5252/z2013n2a5

Bamber, R.N., 2007. A holistic re-interpretation of the phylogeny of the Pycnogonida Latreille, 1810 (Arthropoda). Zootaxa 312, 295–312.

Bamber, R.N., El Nagar, A., Arango, C.P., 2022. PycnoBase: World Pycnogonida Base [WWW Document]. URL http://www.marinespecies.org/pycnobase/index.php (accessed 1.15.22).

Bergström, J., Stürmer, W., Winter, G., 1980. *Palaeoisopus*, *Palaeopantopus* and *Palaeothea*, pycnogonid arthropods from the Lower Devonian Hunsrück Slate, West Germany. Paläontologische Zeitschrift 54, 7–54. https://doi.org/10.1007/BF02985882

Bouvier, E.L., 1910. Les Pycnogonides à cinq paires de pattes recueillis par la mission antarctique Jean Charcot à bord du Pourquoi Pas? Comptes rendus Hebd. des séances l'Académie des Sci.

Paris 151, 26–32.

Brenneis, G., Bogomolova, E. V., Arango, C.P., Krapp, F., 2017. From egg to "no-body": An overview and revision of developmental pathways in the ancient arthropod lineage Pycnogonida. Front. Zool. 14, 1–22. https://doi.org/10.1186/s12983-017-0192-2

Carapelli, A., Torricelli, G., Nardi, F., Frati, F., 2013. The complete mitochondrial genome of the Antarctic sea spider *Ammothea carolinensis* (Chelicerata; Pycnogonida). Polar Biol. 36, 593–602. https://doi.org/10.1007/s00300-013-1288-6

Charbonnier, S., Vannier, J., Riou, B., 2007. New sea spiders from the Jurassic La Voulte-sur-Rhône Lagerstätte. Proc. R. Soc. B Biol. Sci. 274, 2555–2561. https://doi.org/10.1098/rspb.2007.0848

Child, C.A., 1998. The marine fauna of New Zealand. Pycnogonida (Sea Spiders). Niwa Biodivers. Mem. 109, 71 pp.

Child, C.A., 1994. Antarctic and Subantarctic Pycnogonida II. The family Austrodecidae. Antarct. Res. Ser. 49–99. https://doi.org/10.1029/ar063p0049

Chow, S., Niwa, K., Okamoto, S., Murauchi, Y., Hirai, R., Hibino, M., Wakui, K., Tomiyama, T., Yobayashi, Y., Toba, M., Kano, Y., 2012. Genetic diversity in three local populations of the parasitic sea spider *Nymphonella tapetis* and taxonomic position of the species. Nippon Suisan Gakkaishi 78, 895–902. https://doi.org/10.2331/suisan.78.895

Coddington, J.A., Giribet, G., Harvey, M.S., Prendini, L., Walter, D.E., 2004. Arachnida, in: Cracraft, J., Donoghue, M.J., Donoghue, M.M. (Eds.), Assembling the Tree of Life. Oxford University, pp. 296–318. https://doi.org/10.1111/j.1469-7998.1909.tb08558.x

Dietz, L., Dömel, J.S., Leese, F., Lehmann, T., Melzer, R.R., 2018. Feeding ecology in sea spiders (Arthropoda: Pycnogonida): What do we know? Front. Zool. 15, 1–16. https://doi.org/10.1186/s12983-018-0250-4

Dietz, L., Krapp, F., Hendrickx, M.E., Arango, C.P., Krabbe, K., Spaak, J.M., Leese, F., 2013. Evidence from morphological and genetic data confirms that *Colossendeis tenera* Hilton, 1943

(Arthropoda: Pycnogonida), does not belong to the *Colossendeis megalonyx* Hoek, 1881 complex. Org. Divers. Evol. 13, 151–162. https://doi.org/10.1007/s13127-012-0120-4

Dietz, L., Mayer, C., Arango, C.P., Leese, F., 2011. The mitochondrial genome of *Colossendeis megalonyx* supports a basal position of Colossendeidae within the Pycnogonida. Mol. Phylogenet. Evol. 58, 553–558. https://doi.org/10.1016/j.ympev.2010.12.016

Dietz, L., Pieper, S., Seefeldt, M.A., Leese, F., 2015. Morphological and genetic data clarify the taxonomic status of *Colossendeis robusta* and *C. glacialis* (Pycnogonida) and reveal overlooked diversity. Arthropod Syst. Phylogeny 73, 107–128.

Dunlop, J.A., Arango, C.P., 2004. Pycnogonid affinities: A review. J. Zool. Syst. Evol. Res. 43, 8–21. https://doi.org/10.1111/j.1439-0469.2004.00284.x

Dunlop, J.A., Lamsdell, J.C., 2017. Segmentation and tagmosis in Chelicerata. Arthropod Struct. Dev. 46, 395–418. https://doi.org/10.1016/j.asd.2016.05.002

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Frankowski, K., Miyazaki, K., Brenneis, G., 2022. A microCT-based atlas of the central nervous system and midgut in sea spiders (Pycnogonida) sheds first light on evolutionary trends at the family level. Front. Zool. 19, 1–34. https://doi.org/10.1186/s12983-022-00459-8

Fry, W.G., 1978. A classification within the pycnogonids. Zool. J. Linn. Soc. 63, 35–58. https://doi.org/10.1111/j.1096-3642.1978.tb02089.x

Golovatch, S., Geoffroy, J.-J., Mauriès, J.-P., VandenSpiegel, D., 2009. Review of the millipede family Haplodesmidae Cook, 1895, with descriptions of some new or poorly-known species (Diplopoda, Polydesmida). Zookeys 7, 1–53. https://doi.org/10.3897/zookeys.7.117

Guille, A., Soyer, J., 1967. Nouvelle signalisation du genre *Nymphonella* Ohshima à Banyuls-sur-Mer: *Nymphonella lecalvezi* n. sp. Vie Milieu 18, 345–353.

Hassanin, A., 2006. Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for

limiting the misleading effects of multiple changes in pattern and rates of substitution. Mol. Phylogenet. Evol. 38, 100–116. https://doi.org/10.1016/j.ympev.2005.09.012

Hassanin, A., Léger, N., Deutsch, J., 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoa, and consequences for phylogenetic inferences. Syst. Biol. 54, 277–298. https://doi.org/10.1080/10635150590947843

Hassanin, A., Veron, G., Ropiquet, A., van Vuuren, B.J., Lécu, A., Goodman, S.M., Haider, J., Nguyen, T.T., 2021. Evolutionary history of Carnivora (Mammalia, Laurasiatheria) inferred from mitochondrial genomes. PLoS One 16, 1–28. https://doi.org/10.1371/journal.pone.0240770

Hedgpeth, J.W., 1978. A reappraisal of the Palaeopantopoda with description of a species from the Jurassic. Zool. J. Linn. Soc. 63, 23–34. https://doi.org/10.1111/j.1096-3642.1978.tb02088.x

Hedgpeth, J.W., 1954. On the phylogeny of the Pycnogonida. Acta Zool. 35, 193–213. https://doi.org/10.1111/j.1463-6395.1954.tb00034.x

Hedgpeth, J.W., 1947. On the evolutionnary significance of the Pycnogonida. Smithson. Misc. Collect. 106, 53 pp.

Hugall, A.F., O'hara, T.D., Hunjan, S., Nilsen, R., Moussalli, A., 2016. An exon-capture system for the entire class ophiuroidea. Mol. Biol. Evol. 33, 281–294. https://doi.org/10.1093/molbev/msv216

Jeong, J.H., Kim, H., Ryu, S., Kim, W., 2020. The First Pycnogonid Draft Genome of *Nymphon striatum*. Front. Ecol. Evol. 8, 1–6. https://doi.org/10.3389/fevo.2020.554164

Kearse, M., Moir, R., Wilson, A., Stones-havas, S., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic : An integrated and extendable desktop software platform for the organization and analysis of sequence data 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kim, I.-H., Hong, J.-S., 1986. Korean Shallow-water pycnogonids based on the collections of the Korean Ocean Research and Development Institute. Korean J. Syst. Zool. 2, 35–52.

Kühl, G., Poschmann, M., Rust, J., 2013. A ten-legged sea spider (Arthropoda: Pycnogonida) from the Lower Devonian Hunsrück Slate (Germany). Geol. Mag. 150, 556–564. https://doi.org/10.1017/s0016756812001033

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054

Lee, D., Kim, W., 2020. New species of *Pycnogonum* (Pycnogonida: Pycnogonidae) from Green Island, Taiwan, with an additional note on the holotype of *P. spatium*. Zootaxa 4750, 122–130. https://doi.org/10.11646/zootaxa.4750.1.6

Mallatt, J., Craig, C.W., Yoder, M.J., 2012. Nearly complete rRNA genes from 371 Animalia: updated structure-based alignment and detailed phylogenetic analysis. Mol. Phylogenet. Evol. 64, 603–617.

Manuel, M., Jager, M., Murienne, J., Clabaut, C., Guyader, H. Le, 2006. Hox genes in sea spiders (Pycnogonida) and the homology of arthropod head segments. Dev. Genes Evol. 216, 481–491. https://doi.org/10.1007/s00427-006-0095-2

Masta, S.E., McCall, A., Longhorn, S.J., 2010. Rare genomic changes and mitochondrial sequences provide independent support for congruent relationships among the sea spiders (Arthropoda, Pycnogonida). Mol. Phylogenet. Evol. 57, 59–70. https://doi.org/10.1016/j.ympev.2010.06.020

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A Phylogeny of Birds Based on Over 1,500 Loci Collected by Target Enrichment and High-Throughput Sequencing. PLoS One 8. https://doi.org/10.1371/journal.pone.0054848

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees, in: Proceedings of the Gateway Computing Environments Workshop (GCE) 14 Nov. 2010. New Orleans, LA, pp. 1–8.

Munilla, T.L., 1999. Evolución y filogenia de los Picnógonidos. Boletín la Soc. entomológica Aragon. 26, 273–279.

Munilla, T.L., de Haro, A., 1981. An electrophoretical and immunological study of Pycnogonida, with phylogenetic considerations. Bijdr. tot Dierkd. 51, 191–198.

Nakamura, K., Kano, Y., Suzuki, N., Namatame, T., Kosaku, A., 2007. 18S rRNA phylogeny of sea spiders with emphasis on the position of Rhynchothoracidae. Mar. Biol. 153, 213–223. https://doi.org/10.1007/s00227-007-0803-0

Nosenko, T., Schreiber, F., Adamska, M., Adamski, M., Eitel, M., Hammel, J., Maldonado, M., Müller, W.E.G., Nickel, M., Schierwater, B., Vacelet, J., Wiens, M., Wörheide, G., 2013. Deep metazoan phylogeny: When different genes tell different stories. Mol. Phylogenet. Evol. 67, 223–233. https://doi.org/10.1016/j.ympev.2013.01.010

Park, S.J., Lee, Y.S., Hwang, U.W., 2007. The complete mitochondrial genome of the sea spider *Achelia bituberculata* (Pycnogonida, Ammotheidae): Arthropod ground pattern of gene arrangement. BMC Genomics 8, 1–17. https://doi.org/10.1186/1471-2164-8-343

Petzold, A., Hassanin, A., 2020. A comparative approach for species delimitation based on multiple methods of multi-locus DNA sequence analysis: A case study of the genus *Giraffa* (Mammalia, Cetartiodactyla). PLoS One 15, 28 pp.

Philippe, H., De Vienne, D.M., Ranwez, V., Roure, B., Baurain, D., Delsuc, F., 2017. Pitfalls in supermatrix phylogenomics. Eur. J. Taxon. 2017, 1–25. https://doi.org/10.5852/ejt.2017.283

Podsiadlowski, L., Braband, A., 2006. The complete mitochondrial genome of the sea spider *Nymphon gracile* (Arthropoda : Pycnogonida). BMC Genomics 7. https://doi.org/10.1186/1471-2164-7-284

Posada, D., 2008. jModelTest: Phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256. https://doi.org/10.1093/molbev/msn083

Poschmann, M., Dunlop, J.A., 2006. A new sea spider (Arthropoda: Pycnogonida) with a flagelliform telson from the Lower Devonian Hunsrück Slate, Germany. Palaeontology 49, 983–989. https://doi.org/10.1111/j.1475-4983.2006.00583.x

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Ropiquet, A., Li, B., Hassanin, A., 2009. SuperTRI: A new approach based on branch support analyses of multiple independent data sets for assessing reliability of phylogenetic inferences. Comptes Rendus - Biol. 332, 832–847. https://doi.org/10.1016/j.crvi.2009.05.001

Rota-Stabelli, O., Telford, M.J., 2008. A multi criterion approach for the selection of optimal outgroups in phylogeny: Recovering some support for Mandibulata over Myriochelata using mitogenomics. Mol. Phylogenet. Evol. 48, 103–111. https://doi.org/10.1016/j.ympev.2008.03.033

Roure, B., Baurain, D., Philippe, H., 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. Mol. Biol. Evol. 30, 197–214. https://doi.org/10.1093/molbev/mss208

Sabroux, R., Audo, D., Charbonnier, S., Corbari, L., Hassanin, A., 2019a. 150-million-year-old sea spiders (Pycnogonida: Pantopoda) of Solnhofen. J. Syst. Palaeontol. 0, 1–12. https://doi.org/10.1080/14772019.2019.1571534

Sabroux, R., Corbari, L., Krapp, F., Bonillo, C., Le Prieur, S., Hassanin, A., 2017. Biodiversity and phylogeny of Ammotheidae (Arthropoda: Pycnogonida). Eur. J. Taxon. 1–33. https://doi.org/10.5852/ejt.2017.286

Sabroux, R., Hassanin, A., Corbari, L., 2022. Sea spiders (Arthropoda: Pycnogonida) collected during Madibenthos expedition from Martinique shallow waters. Eur. J. Taxon. 851, 1–141. https://doi.org/https://doi.org/10.5852/ejt.2022.851.1999

Sabroux, R., Hassanin, A., Corbari, L., 2019b. Four times more species of sea spiders (Arthropoda: Pycnogonida) in Martinique Island (Lesser Antilles). Mar. Biodivers. 49, 1519–1535. https://doi.org/10.1007/s12526-019-00957-9

Simmons, M.P., 2017. Relative bene fi ts of amino-acid, codon, degeneracy, DNA, and purine-pyrimidine character coding for phylogenetic analyses of exons. J. Syst. Evol. 55, 85–109.

https://doi.org/10.1111/jse.12233

Siveter, Derek J., Sutton, M.D., Briggs, D.E.G., Siveter, David J., 2004. A Silurian sea spider. Lett. to Nat. 431, 978–980.

Smith, B.T., Mauck, W.M., Benz, B.W., Andersen, M.J., 2020. Uneven missing data skew phylogenomic relationships within the lories and lorikeets. Genome Biol. Evol. 12, 1131–1147. https://doi.org/10.1093/GBE/EVAA113

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Staples, D.A., 2002. *Pycnogonum* (Pycnogonida: Pycnogonidae) from Australia with descriptions of two new species. Mem. Museum Victoria 59, 541–553. https://doi.org/10.24199/j.mmv.2002.59.12

Stock, J.H., 1994. Indo-West Pacific Pycnogonida collected by some major Oceanographic Expeditions. Beaufortia 44, 17–77.

Stock, J.H., 1991. Deep-water Pycnogonida from the surroundings of New Caledonia, in: Crosnier, A. (Ed.), Résultats Des Campagnes MUSORSTOM, 8. Mémoires du Muséum national d'Histoire naturelle. Série A, Zoologie, pp. 125–212.

Stock, J.H., 1979. Pycnogonida from the mediolittoral and infralittoral zones in the Tropical Western Atlantic. Stud. Fauna Curaçao Other Caribb. Islands 184, 1–32.

Stock, J.H., 1958. The Pycnogonid genus *Rhopalorhynchus* Wood-Mason 1873. Tijdschr. voor Entomol. 101, 113–137.

Stork, N.E., 2018. How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth ? Annu. Rev. Entomol. 63, 31–45.

Swofford, D.L., Bell, C.D., 2017. PAUP 4*.

Tharme, R.E., Webb, S.C., Brown, A.C., 1996. Organisms associated with the sandy-beach bivalve Donax serra Röding, with a description of Cercaria serrae sp. nov. (Trematoda). South African J.

Zool. 31, 86–90. https://doi.org/10.1080/02541858.1996.11448398

Tomiyama, T., Miyazaki, K., Tamaoki, M., Yamada, K., 2015. 18S analysis of the taxonomic position of an endoparasitic pycnogonid, *Nymphonella tapetis* (Arthropoda: Pycnogonida: Ascorhynchidae). J. Crustac. Biol. 35, 491–494. https://doi.org/10.1163/1937240X-00002348

Tomiyama, T., Yamada, K., Wakui, K., Tamaoki, M., Miyazaki, K., 2016. Impact of sea spider parasitism on host clams: relationships between burial patterns and parasite loads, somatic condition and survival of host. Hydrobiologia 770, 15–26. https://doi.org/10.1007/s10750-015-2564-y

Xue, X., Dong, Y., Deng, W., Hong, X., Shao, R., 2017. Molecular Phylogenetics and Evolution The phylogenetic position of eriophyoid mites (superfamily Eriophyoidea) in Acariformes inferred from the sequences of mitochondrial genomes and nuclear small subunit (18S) rRNA gene. Mol. Phylogenet. Evol. 109, 271–282. https://doi.org/10.1016/j.ympev.2017.01.009

Yoshinaga, T., Kobayashi, Y., Toba, M., Miyama, Y., 2011. Development of a Sensitive Method for the Detection of Young Larvae of the Parasitic Pycnogonid *Nymphonella tapetis* in Manila Clam *Ruditapes philippinarum*. Fish Pathlology 46, 38–41.

Zehnpfennig, J.R., Varney, R.M., Halanych, K.M., Mahon, A.R., 2022. Mitochondrial genomes provide insight into interfamilial relationships within Pycnogonida. Polar Biol. 45, 1513–1522. https://doi.org/10.1007/s00300-022-03085-6

CAPTIONS

Figure 1. A glimpse of Pycnogonida diversity. *Ammothea* sp., Antarctica (a); *Ammothella exornata*, Martinique (b); *Nymphopsis muscosa*, Papua-New Guinea (c); *Ascorhynchus* sp., Papua-New Guinea (d); *Eurycyde kaiouti*, Martinique (e); *Austrodecus stocki*, Madagascar (f); *Pallenoides amazonicus*, French Guiana (g); *Austropallene cornigera*, Antarctica (h); *Colossendeis* cf. *macerrima*, Mozambique Channel (i); *Endeis* sp., Papua-New Guinea (j); *Nymphon australe*, Antarctica (k); *Pallenopsis schmitti*, Martinique (l); *Anoplodactylus* sp., Papua-New Guinea (m); *Pycnogonum cesairei*, Martinique (n); *Pentapycnon geayi*, Martinique (o); *Rhynchothorax crenatus*, Martinique (p). All scale bars 1 mm. pictures T.Y. Chan, L. Corbari, Z. Ďuriš, R. Sabroux, S. Soubzmaigne; ©REVOLTA-IPEV 1124, ©MNHN – La Planète Revisitée, ©MNHN – Tropical Deep-Sea Benthos.

Figure 2 (two pages). Bayesian tree based on mt-197 dataset. Support values at nodes are indicated (left value: posterior probabilities [PP], right value: bootstrap percentages [BP]). Maximal supports (PP = 1, BP = 100) are indicated by an asterisk (*). PP < 0.5 and BP < 50 are marked as "-". An "X" indicates that an alternative hypothesis is supported by BP. Outgroups were removed for better readability.

Figure 3 (two pages). Bayesian tree based on the mt+18S-157 dataset. Support values at nodes are indicated (first posterior probabilities [PP], second bootstrap percentages [BP], third mean posterior probabilities [MPP] and fourth superTRI Bootstrap percentages [SBP]). Maximal support values (PP = 1, BP = 100, MPP = 1, SBP = 100) are indicated by an asterisk (*). PP and MPP < 0.5, and SBP and BP < 50 are marked as "-". An "X" indicates that an alternative hypothesis is supported by BP or SBP. Node repeatability (Rep) in separate analyses of the 18S-157 and mt-157 alignments are indicated by branch thickening and colouring: red (Rep = 1), orange (Rep = 0.5, recovered in analyses on mt-157), blue (Rep = 0.5, recovered in analyses on 18S-157), and dashed black (Rep = 0).

Figure 4. Bayesian tree based on the M3-83 dataset. See figure 3 for signification of the legends. Nodes repeatability (Rep) in separate analyses of the mt-83, UCE-83, OG-83 and rib-83 datasets are indicated by branch thickening (dashed black line code for Rep = 0). Repeatability from at least one mitochondrial (mt) and one nuclear (nu) markers are marked by branches coloured in red, repeatability by only nu markers in blue, and only by the mt marker in orange.

Figure 5. Synthetic tree summarizing intergeneric relationships within Pycnogonida.
The ML bootstrap tree reconstructed from the mt+18S-211 dataset (Appendix D) was used as a framework to draw the synthetic tree. Other analyses were used to highlight the most reliable relationships. Thick grey branches indicate nodes supported by our 18S-157 subdataset (Appendix B) and at least one of our mitochondrial alignments mt-157 and mt-197-RY (fig. 3 and supplementary material D) (first repeatability criterion). Thick white branches indicate nodes supported by the mitochondrial alignment mt-83 and at least one of the three nuclear subdatasets of the M3-83 dataset (OG-83, rib-83 and UCE-83; fig. 4 and supplementary material D) (second repeatability criterion). Thick black branches indicate nodes for which the two criteria of repeatability were validated. Dashed branches indicate the least reliable relationships with none of the repeatability criteria validated. The letters refer to the results of previous studies: "A" for Arabi *et al*. (2010), "B" for Ballesteros *et al*. (2021) and "S" for Sabroux *et al*. (2017; only within Ammotheidae). They are green for highly supported nodes (PP $\geq$ 0.95 and/or BP $\geq$ 50%), orange for nodes weakly supported (PP < 0.95 and/or BP < 50%), and red when an alternative hypothesis was supported by PP $\geq$ 0.95 and/or BP $\geq$ 50%. Although not included in the two datasets, the genera of Austrodecidae and Rhynchothoracidae were tentatively placed in the tree (grey branches) using the results of the M3-110 tree (Appendix C)."

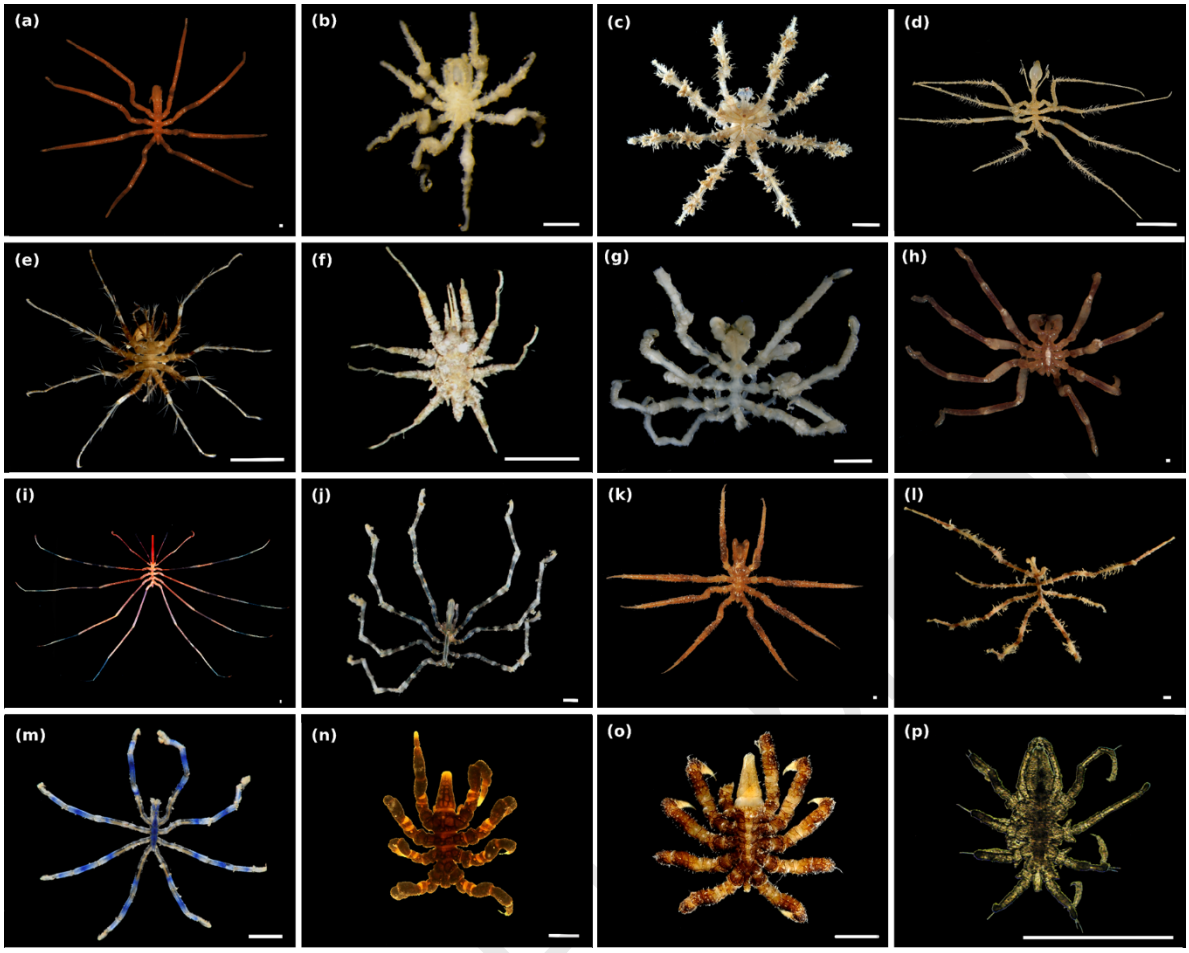
Figure 6. Body-plan of sea spiders, with organs and appendages highlighted. A: dorsal and ventral views of adult male of *Achelia echinata* (Ammotheidae), whith chelifores coloured in blue, palps coloured in red, ovigers coloured in green (including the strigilis, in dark green). B: protonymphon larva of *Nymphon brevirostre*, after Bogomolova (2007). C: distribution of the cephalic appendages and

cement gland characters in the synthetic tree of sea spider as recovered in fig. 5: blue circle representing the chelifores, red the palps, light green the ovigers, dark green the strigilis, orange the cement glands. The cephalic appendage or gland is present when the circle is coloured, absent when empty. When half coloured, it is present only in males; when quarter-coloured, only in males of some species/genera. When three-quarter coloured, it is absent only in some males. Circles marked with a drawbar indicate the reduction of chelifores terminal podomeres or their complete absence (for chelifores), reduction of palps to unfunctional bulges (for palps), a marked sexual dimorphism on the ovigers (for ovigers), and the absence of paternal care (for cement glands). In Pycnogonidae, paternal care is present, but the cement glands have not yet been identified, which is indicated with an orange question mark ("?"). An asterisk over a circle indicates that exceptions are known to occur within the group for this character. White asterisks on the tree's terminal branches indicate groups including polymerous species (*i.e.*, with ten or twelve walking legs).

Table 1. Table 1: Specimens sequenced for mt-genomes and 18S analyses (black) indicated by collection numbers (MNHN-IU-) or mined from GenBank (blue) with GenBank accession numbers.

Table 2. Summary of datasets used for phylogenetic analyses with quantification of missing data in the alignments.

- 152 mitogenomes and 120 18S ribosomal genes of sea spiders were sequenced.
- Phylogenetic signal extraction is impacted when high levels of missing data are included.
- Strong support for four superfamilies, six families, four subfamilies.
- Most cephalic appendage characters have evolved by convergence in different families.
- Poorly resolved deep relationships may be due to radiation before Jurassic.



TABLES

Table 1: Specimens sequenced for mt-genomes and 18S analyses (black) indicated by collection numbers (MNHN-IU-) or mined from GenBank (blue) with GenBank accession numbers.

Familly	**Genus**	**Species**	**locality**	**Collection number**	**mt-197**	**18S-157**
Ammothoeidae	*Ammothea*	*Ammothea adunca*	Antarctica	MNHN-IU-2007-231	OP985918	OQ065567
		Ammothea carolinensis			GU065293	
			Antarctica	MNHN-IU-2007-214	OP985917	
		Ammothea gigantea	Antarctica	MNHN-IU-2016-1421	OP998847	OQ065568
		Ammothea hilgendorfi			GU370075	
						DQ389936

		Ammothea cf. *tibialis*	Antarctica	MNHN-IU-2007-247	OP985915	OQ065569
		Ammothea sp.	Antarctica	MNHN-IU-2007-324	OP985919	OQ065570
	Acheliana	*Acheliana* sp.	Madagascar	MNHN-IU-2011-764	OP998844	OQ065572
			Madagascar	MNHN-IU-2011-659	OP998843	OQ065571
	Cilunculus	*Cilunculus scaurus*	New Caledonia	MNHN-IU-2016-6862	OP998846	OQ065573
		Cilunculus sewelli	Mozambique	MNHN-IU-2011-624	OP998842	OQ065574
	Sericosura	*Sericosura heteroscela*	Mid-Atlantic ridge	MNHN-IU-2013-15606	OP998845	KX536496
		Sericosura sp.	Mid-Atlantic ridge	MNHN-IU-2013-19239	OP998841	KX536422
	Achelia	*Achelia assimilis*	New Caledonia	MNHN-IU-2008-20589	OP998839	
		Achelia bituberculata			AY457170	
		Achelia gracilis	Martinique	MNHN-IU-2016-888	OP998840	
			Martinique	MNHN-IU-2016-851		OQ065575
			Martinique	MNHN-IU-2016-1276	OP988529	
		Achelia sawayai	Martinique	MNHN-IU-2016-859		
			Martinique	MNHN-IU-	OP988537	

				2016-1073		
			Martinique	MNHN-IU-2016-884		OQ065576
		Achelia sp.	Papua New Guinea	MNHN-IU-2013-18597	OP985913	KX536441
	Ammothella	*Ammothella di rbergi*	Martinique	MNHN-IU-2016-833 (holotype)	OP988530	OQ065579
			Martinique	MNHN-IU-2016-1091 (paratype)	OP988536	
		Ammothella exornata	Martinique	MNHN-IU-2016-829	OP988532	OQ065577
		Ammothella spinifera	Martinique	MNHN-IU-2016-826	OP988524	OQ065578
	Nymphopsis	*Nymphopsis curtiscapus*	Madagascar	MNHN-IU-2011-760	OP988527	OQ065581
		Nymphopsis duodorsospinosa	Martinique	MNHN-IU-2016-812	OP988528	OQ065580
		Nymphopsis muscosa	Papua New Guinea	MNHN-IU-2013-6600	OP988525	
			Papua New Guinea	MNHN-IU-2013-18640	OP988526	OQ065583
	Tanystylum	*Tanystylum acuminatum*	Martinique	MNHN-IU-2016-856	OP988531	OQ065583
		Tanystylum boucheti (holotype)	Martinique	MNHN-IU-2016-1074	OP988533	OQ065587
		Tanystylum hu mmelincki	Martinique	MNHN-IU-	OP988508	OQ065584

				2016-868		
		Tanystylum ingrallis (paratype)	Martinique	MNHN-IU-2016-867	OP988535	OQ065588
		Tanystylum orbiculare			GU370074	
						DQ389910
			Martinique	MNHN-IU-2016-872	OP988538	OQ065585
		Tanystylum tayronae	Martinique	MNHN-IU-2016-858	OP985914	OQ065586
		Ammotheidae gen. sp.	New Caledonia	MNHN-IU-2016-1393	OP988548	OQ065589
Ascorhynchidae	*Ascorhynchus*	*Ascorhynchus castelli*	French Guyana	MNHN-IU-2014-8275	OP988551	KX53646
		Ascorhynchus castellioides	Martinique	MNHN-IU-2016-883	OP985922	OQ065590
		Ascorhynchus iguanarum (holotype)	Martinique	MNHN-IU-2016-1047	OP988555	
		Ascorhynchus latipes	Martinique	MNHN-IU-2016-816	OP988557	OQ065591
		Ascorhynchus quartogibbus (holotype)	Solomon Is.	MNHN-IU-2008-20493	OP985921	OQ065592
		Ascorhynchus seticauda	New Caledonia	MNHN-IU-2016-6864	OP988553	OQ065593
		Ascorhynchus sp.	Madagascar	MNHN-IU-2011-748	OP988552	OQ065594
		Ascorhynchus sp.	Papua New Guinea	MNHN-IU-2013-6582	OP988550	OQ065595
	Bathyzetes	*Bathyzetes* sp.	New Caledonia	MNHN-IU-2016-6865	OP985920	OQ065596

	Eurycyde	*Eurycyde clitellaria*	Martinique	MNHN-IU-2016-824	OP988559	OQ065597
		Eurycyde kaiouti (paratype)	Martinique	MNHN-IU-2016-1187	OP988560	
		Eurycyde raphiaster	Martinique	MNHN-IU-2016-818	OP988554	OQ065598
		Eurycyde sp.	Martinique	MNHN-IU-2016-819	OP988558	OQ065599
		Eurycyde sp.	Papua New Guinea	MNHN-IU-2012-1248	OP988556	OQ065600
		Eurycyde sp.	Guadeloupe	MNHN-IU-2012-869	OP985968, OP985969	
		Ascorhynchidae gen. sp.	Papua New Guinea	MNHN-IU-2013-6550	OP988549	
Colossendeidae	*Colossendeis*	*Colossendeis australis*	Antarctica	MNHN-IU-2007-172	OP985930	OQ065601
		Colossendeis colossea	Vanuatu	MNHN-IU-2008-20591	OP988567	OQ065602
			New Caledonia	MNHN-IU-2008-20598	OP985934	OQ065603
		Colossendeis leptorhynchus	Vanuatu	MNHN-IU-2008-20651	OP985939	OQ065604
		Colossendeis macerrima	Solomon Is.	MNHN-IU-2008-20504	OP985928	OQ065605
			Vanuatu	MNHN-IU-2008-20580	OP985932	OQ065606
			Vanuatu	MNHN-IU-	OP985936	OQ065607

				2008-20494		
			Papua New Guinea	MNHN-IU-2011-1642	OP985927	OQ065608
			French Polynesia	MNHN-IU-2011-3664	OP988566	
		Colossendeis cf. *macerrima*	New Caledonia	MNHN-IU-2016-1473	OP988572	OQ065609
		Colossendeis megalonyx			HQ450773	
		Colossendeis minor	New Caledonia	MNHN-IU-2008-20602	OP985935	OQ065610
		Colossendeis pipetta	New Caledonia	MNHN-IU-2008-20509	OP985929	OQ065611
		Colossendeis cf. *pipetta*	New Caledonia	MNHN-IU-2016-1465	OP988571	OQ065612
		Colossendeis tenuipedis	Antarctica	MNHN-IU-2007-130	OP988564	OQ065613
		Colossendeis sp.	Antarctica	MNHN-IU-2007-212	OP985972, OP985973	OQ065614
		Colossendeis sp.	Glorioso Is.	MNHN-IU-2016-1412	OP988569	OQ065615
		Colossendeis sp.	Walters shoal	MNHN-IU-2016-1419	OP988568	OQ065616
	Decolopoda	*Decolopoda australis*	Antarctica	MNHN-IU-2007-219	OP985937	OQ065617
			Antarctica	MNHN-IU-2007-315	OP985931	OQ065618
	Hedgpethia	*Hedgpethia tibialis* (holotype)	New Caledonia	MNHN-IU-2007-4581	OP988570	

		Hedgpethia sp.	New Caledonia	MNHN-IU-2016-6868	OP985938	OQ065619
	Rhopalothynchus	*Rhopalorhynchus filipes*	New Caledonia	MNHN-IU-2016-1476	OP985933	OQ065620
		Colossendeidae gen. sp.	Papua New Guinea	MNHN-IU-2013-6603	OP988565	OQ065621
Endeidae	*Endeis*	*Endeis australis*	Antarctica	MNHN-IU-2007-207	OP988574	OQ065622
		Endeis flaccida	Martinique	MNHN-IU-2016-840	OP988578	OQ065623
		Endeis aff. *meridionalis*	Martinique	MNHN-IU-2016-1142	OP988577	OQ065624
		Endeis sp.	Antarctica	MNHN-IU-2007-135	OP988581	
		Endeis sp.	Guadeloupe	MNHN-IU-2012-846	OP988576	
		Endeis sp.	Papua New Guinea	MNHN-IU-2013-6605	OP988575	OQ065625
		Endeis sp.	Papua New Guinea	MNHN-IU-2013-18638	OP988573	KX536482
		Endeis sp.	Martinique	MNHN-IU-2016-863	OP988580	OQ065626
		Endeis sp.	Walters shoal	MNHN-IU-2016-1416	OP988579	OQ065627
Nymphonidae	*Nymphon*	*Nymphon aemulum*	Martinique	MNHN-IU-2016-1186	OP988582	
				MNHN-IU-2016-845		OQ065628

		Nymphon cf. *apicatum*	New Caledonia	MNHN-IU-2016-1460	OP985942	OQ065629
		Nymphon australe	Antarctica	MNHN-IU-2007-129	OP985940	OQ065630
		Nymphon charcoti	Antarctica	MNHN-IU-2007-327	OP985941	OQ065631
		Nymphon dorlis (holotype)	Martinique	MNHN-IU-2016-879	OP985978, OP985979	OQ065637
		Nymphon cf. *fortunatum*	New Caledonia	MNHN-IU-2021-7458	OP985946	OQ065632
		Nymphon giraffa	New Caledonia	MNHN-IU-2016-1429	OP985944	OQ065633
		Nymphon gracile	Roscoff (France)	MNHN-IU-2014-10214	OP985910	
					DQ666063	
						FJ862851
		Nymphon maculatum	Madagascar	MNHN-IU-2011-660	OP985976, OP985977	OQ065634
		Nymphon martinicum (holotype)	Martinique	MNHN-IU-2016-889	OP985945	OQ065638
		Nymphon striatum			SRR10993134	
		Nymphon surinamensis	French Guyana	MNHN-IU-2013-18615	OP985943	OQ065635
		Nymphon sp.	Guadeloupe	MNHN-IU-2012-976	OP988583	OQ065636
		Nymphon sp.	French Guyana	MNHN-IU-2014-8371	OP985966	
		Nymphon sp.	New Caledonia	MNHN-IU-2016-1472	OP985947	OQ065639

		Nymphon sp.			GU370076	
	Pentanymphon	*Pentanymphon antarcticum*	Antarctica	MNHN-IU-2007-336	OP985974, OP985975	OQ065640
Callipallenidae	*Austropallene*	*Austropallene cornigera*	Antarctica	MNHN-IU-2007-305	OP985924	OQ065641
	Pallenoides	*Pallenoides spinulosus*	Martinique	MNHN-IU-2016-846	OP988561	OQ065642
		Pallenoides cf. *amazonicus*	Martinique	MNHN-IU-2016-866	OP988562	
	Parapallene	*Parapallene bermudensis*	Martinique	MNHN-IU-2016-1221	OP985926	OQ065643
		Parapallene sp.	Mozambique	MNHN-IU-2011-605	OP985923	OQ065644
		Parapallene sp.	Madagascar	MNHN-IU-2011-675	OP985912	OQ065645
		Parapallene sp.	Madagascar	MNHN-IU-2011-707	OP985984, OP985985	
	Propallene	*Propallene* cf. *ardua*	Madagascar	MNHN-IU-2011-757	OP985925	OQ065646
			Madagascar	MNHN-IU-2011-768	OP985970, OP985971	OQ065647
			Madagascar	MNHN-IU-2011-785	OP988563	OQ065648
		Callipallenidae gen. sp.	Mozambique	MNHN-IU-2011-625	OP985967	OQ065649
Pallenopsidae	*Pallenopsis*	*Pallenopsis angusta*	New Caledonia	MNHN-IU-2016-6873	OP988591	OQ065650
		Pallenopsis candidoi	Martinique	MNHN-IU-	OP985948	OQ065651

				2016-814		
		Pallenopsis crosslandi	Madagascar	MNHN-IU-2011-693	OP988588	
		Pallenopsis patagonica	Antarctica	MNHN-IU-2007-134	OP988589	OQ065652
			New Caledonia	MNHN-IU-2008-20510	OP988585	OQ065653
		Pallenopsis pilosa	Antarctica	MNHN-IU-2007-310	OP988584	OQ065654
		Pallenopsis schmitti	Martinique	MNHN-IU-2016-813	OP988593	OQ065655
		Pallenopsis cf. *virgata*	New Caledonia	MNHN-IU-2016-6879	OP985916	OQ065656
		Pallenopsis sp.	New Caledonia	MNHN-IU-2016-1471	OP988595	OQ065657
		Pallenopsis sp.	New Caledonia	MNHN-IU-2016-1478	OP988597	
	Bathypallenopsis	*Bathypallenopsis mollissima*	Vanuatu	MNHN-IU-2008-20672	OP988587	OQ065658
		Bathypallenopsis sp.	Walters shoal	MNHN-IU-2016-1417	OP988592	OQ065659
		Bathypallenopsis sp.	New Caledonia	MNHN-IU-2016-1469	OP988596	OQ065660
		Pallenopsidae gen. sp.	Antarctica	MNHN-IU-2007-125	OP988594	
		Pallenopsidae gen. sp.	Mozambique	MNHN-IU-2009-2	OP988586	OQ065661
		Pallenopsidae gen. sp.	Papua New Guinea	MNHN-IU-	OP988590	OQ065662

				2013-6558		
Phoxichilidiidae	*Anoplodactylus*	*Anoplodactylus* cf. *californicus*	Martinique	MNHN-IU-2016-811	OP985954	OQ065663
			Martinique	MNHN-IU-2016-855	OP985951	OQ065664
		Anoplodactylus digitatus	Martinique	MNHN-IU-2016-876	OP985955	OQ065665
		Anoplodactylus ganchiformis	Martinique	MNHN-IU-2016-575	OP985952	
			Martinique	MNHN-IU-2016-854	OP985956	OQ065666
		Anoplodactylus glandulifer	Martinique	MNHN-IU-2016-877	OP985953	OQ065667
		Anoplodactylus insignis	Martinique	MNHN-IU-2016-808	OP988501	OQ065668
		Anoplodactylus massiliformis	Martinique	MNHN-IU-2016-1124	OP988502	OQ065669
		Anoplodactylus micros	Martinique	MNHN-IU-2016-804	OP988507	OQ065670
		Anoplodactylus pectinus	Martinique	MNHN-IU-2016-806	OP988503	OQ065671
		Anoplodactylus sp.	Madagascar	MNHN-IU-2011-699	OP988599	OQ065672
		Anoplodactylus madibenthos	Guadeloupe	MNHN-IU-2012-873	OP988500	
			Martinique	MNHN-IU-2016-1071 (holotype)	OP988505	

		Anoplodactylus sp.	Papua New Guinea	MNHN-IU-2012-1279 & MNHN-IU-2012-1266	OP985949	OQ065673 & OQ065674
		Anoplodactylus sp.	French Guyana	MNHN-IU-2013-18611	OP988504	KX535471
		Anoplodactylus sp.	Glorioso Is.	MNHN-IU-2016-1410	OP985950	OQ065675
		Phoxichilidiidae gen. sp.	Papua New Guinea	MNHN-IU-2013-6592	OP988598	OQ065676
		Phoxichilidiidae gen. sp.	Papua New Guinea	MNHN-IU-2013-6598	OP988509	OQ092425
		Phoxichilidiidae gen. sp.	French Guyana	MNHN-IU-2013-18544	OP988506	KX536457
Pycnogonidae	*Pycnogonum*	*Pycnogonum africanum*	Madagascar	MNHN-IU-2011-104	OP985961	OQ065677
		Pycnogonum cesairei (holotype)	Martinique	MNHN-IU-2016-10338	OP985982, OP985983	OQ065683
		Pycnogonum gaini	Antarctica	MNHN-IU-2007-126	OP985958	
		Pycnogonum madagascariensis	Madagascar	MNHN-IU-2011-717	OP985959	OQ065678
		Pycnogonum staplesi	New Caledonia	MNHN-IU-2016-6863	OP985962	OQ065679
		Pycnogonum sp.	Antarctica	MNHN-IU-2007-296	OP985957	OQ065680
		Pycnogonum sp.	Antarctica	MNHN-IU-2007-326	OP985965	OQ065681

		Pycnogonum sp.	Mozambique	MNHN-IU-2011-623	OP985960	OQ065682
	Pentapycnon	*Pentapycnon* cf. *bouvieri*	Antarctica	MNHN-IU-2007-330	OP985980, OP985981	OQ065684
		Pentapycnon geayi	Martinique	MNHN-IU-2016-4187	OP985963	OQ065685
		Pycnogonidae gen. sp.	Mozambique	MNHN-IU-2011-618	OP985964	OQ065686
Ricinulei		*Cryptocellus narino*			KC688690	
		Cryptocellus peckorum				JX951342
		Pseudocellus gertschi			KC688691	
		Pseudocellus pearsei			EU024483	
						U91489
		Ricinoides karschii			KC688692	
						JX951334
Amblypygi		*Damon diadema*			FJ204233	
						AY829907
		Phrynus sp.			EU520641	
		Phrynus goesii				JN018234
Mesothelae		*Liphistius erawan*			JQ407803	
		Liphistius bicoloripes				AF007104
		Heptathela hangzhouensis			AY309258	
		Heptathela kimurai				KY01651
Uropygi		*Mastigoproctus giganteus*			EU520643	
						AF005446
Ixodida		*Ixodes persulcatus*			KU935457	
						AY274888
		Ornithodoris turicata			MF818021	
						MG437266
Opiliones		*Opilio parietinus*			HM367070	
						AF124938
		Oligolophus tienmushanensis			KJ534551	
		Phalangium opilio			EU523757	
						AF124937
Solifugae		*Nothopuga* sp.			EU024482	
		Eremobates cf. *palpisetulosus*			EU520642	

	Eremobates sp.				AY859573
Limulida	*Carcinoscorpius rotundicauda*			JX437074	
					HQ588739
				JQ178358	
	Tachypleus tridentatus			JQ739210	
				FJ860267	
					HQ876480
	Limulus polyphemus			JX983598	
					HQ588741
Diplopoda	*Bachycybe lecontii*			NC 021934	
	Narceus annularus			AY055727	
	Narceus americanus				EU68519
	Abacion magnum			NC 021932	
Chilopoda	*Lithobius forficatus*			AF309492	
					EU024571
	Scolopocryptops sp.			KC200076	
	Scolopocryptops miersii				HQ402510
Pancrustacea	*Hydroporus obscurus*			KT876896	
	Hydroporus pubescens				AJ318734
	Mantis religiosa			KU201317	
					AY859586
	Nannophya pygmaea			KY402222	
	Leucorrhinia sp.				AY859584
	Pyrhila pisum			KU343210	
					Z25817
	Squilla mantis			AY639936	
					GQ328958
	Triops longicaudatus			GU475465	
					AF144219
Onychophora	*Epiperipatus biolleyi*			DQ666064	
				HM600781	
	Peripatoides sp.			HM600782	
	Peripatoides sympatrica			JF800075	
					MG973635
	Opisthopatus cinctipes			HM008997	
	Opisthopatus roseus				MG973642

1

2

Table 2: Summary of datasets used for phylogenetic analyses with quantification of missing data in the alignments.

dataset	Alignment size (nt)	Average missing data (%)	Taxon with the highest percentage of missing data
mt-197	**17989**	**0.6**	***Pycnogonum gaini* MNHN-IU-2007-126 (14.8%)**
mt+18S-157	**19705**	**0.2**	***Pentapycnon* cf. *bouvieri* MNHN-IU-2007-330 (4.1%)**
mt-157	*17989*	*0.6*	*Nymphon giraffa* MNHN-IU-2016-1429 (*12.2%*)
18S-157	*1711*	*1.3*	*Pentapycnon* cf. *bouvieri* MNHN-IU-2007-330 *(30.3%)*
M3-110	**24142**	**59.4**	***Austrodecus gordonae* PYC001 (96.9%)**
mt-110	*8286*	*52.1*	7 sequences *(100%)*
OG-110	*8696*	*62.0*	4 sequences *(100%)*
UCE-110	*3562*	*67.3*	7 sequences *(100%)*
rib-110	*3598*	*62.1*	16 sequences *(100%)*
M3-83	**24142**	**54.0**	***Rhynchothorax monnioti* PYC072 (83.3%)**
mt-83	*8286*	*45.3*	*Pentanymphon* PYC057 *(98.0%)*
OG-83	*8696*	*58.4*	*Pentanymphon* PYC116 *(77.6%*)
UCE-83	*3562*	*62.6*	*Pantopipetta armoricana* PYC002 *(93.3%)*
rib-83	*3598*	*53.8*	*Rhynchothorax monnioti* PYC072 (*95.6%)*
mt+18S-211	**19700**	**14.7**	***Pentanymphon* PYC112 (68.7%)**
mt-211	*17989*	*14.6*	*Nymphon molleri* PYC054 *(67.0%)*
18S-211	*1711*	*15.5*	*Colossendeis* cf. *glacialis* PYC098 *(68.7%)*

CRediT author statement

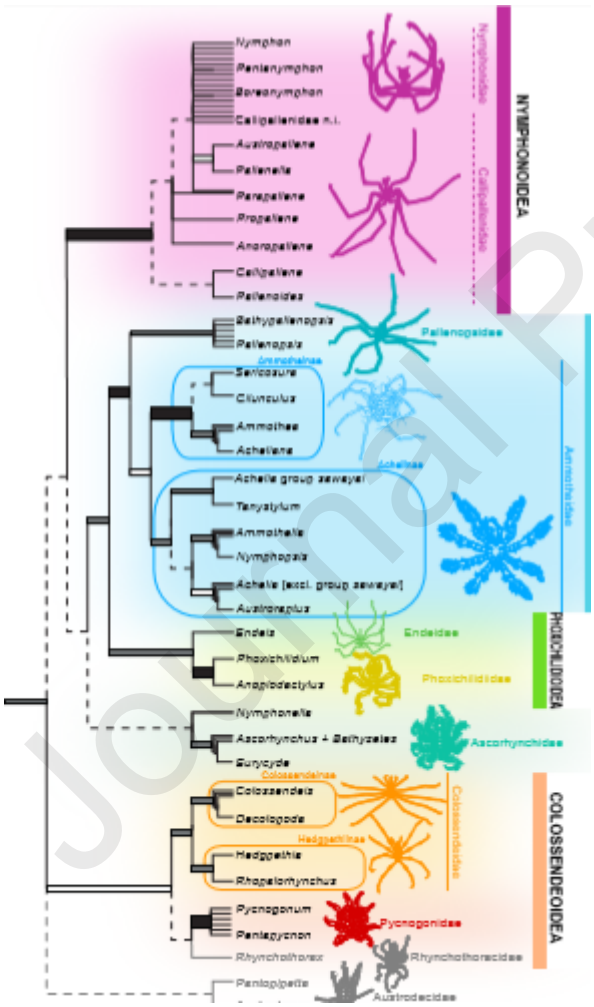
Romain Sabroux: Conceptualization, Data curation, Formal analyses, Investigation, Project administration, Validation, Visualization, Writing – original draft, Writing – review and editing

Laure Corbari: Data curation, Funding acquisition, Specimen curation, Resources, Supervision, Writing – original draft, Writing – review and editing

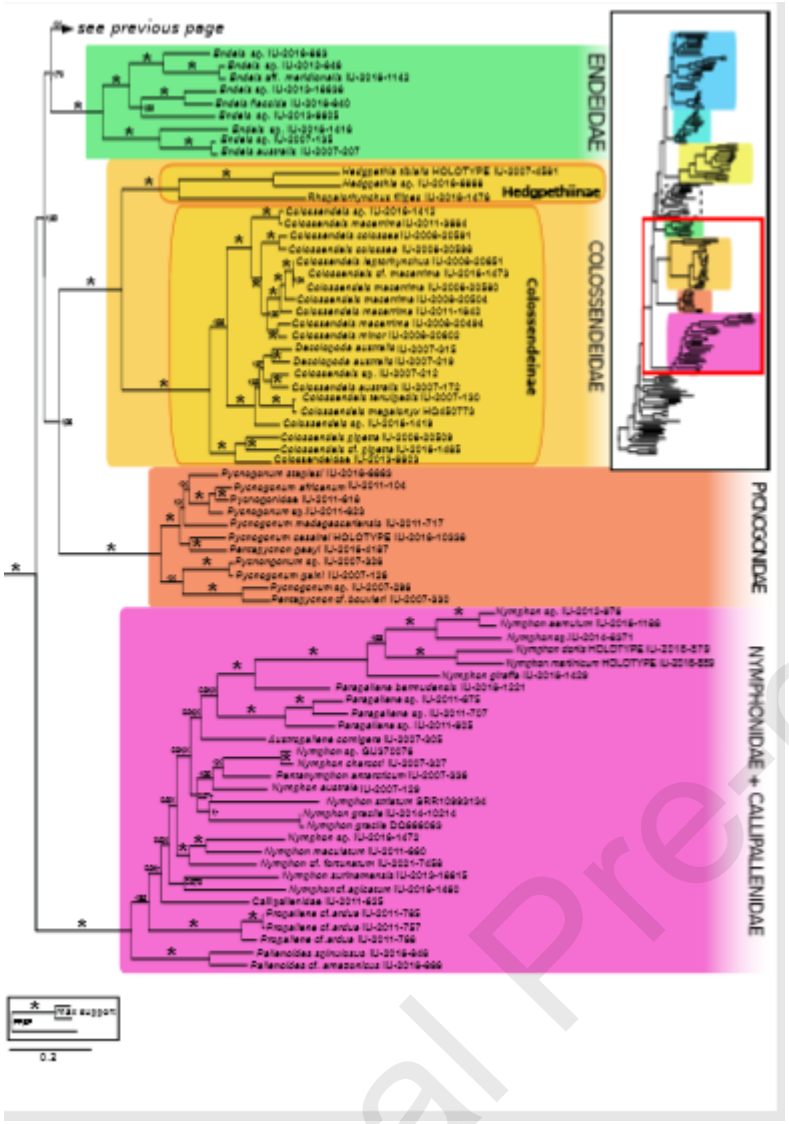
Alexandre Hassanin: Conceptualization, Funding acquisition, Methodology, Formal analyses, Supervision, Writing – original draft, Writing – review and editing

Declaration of competing interest

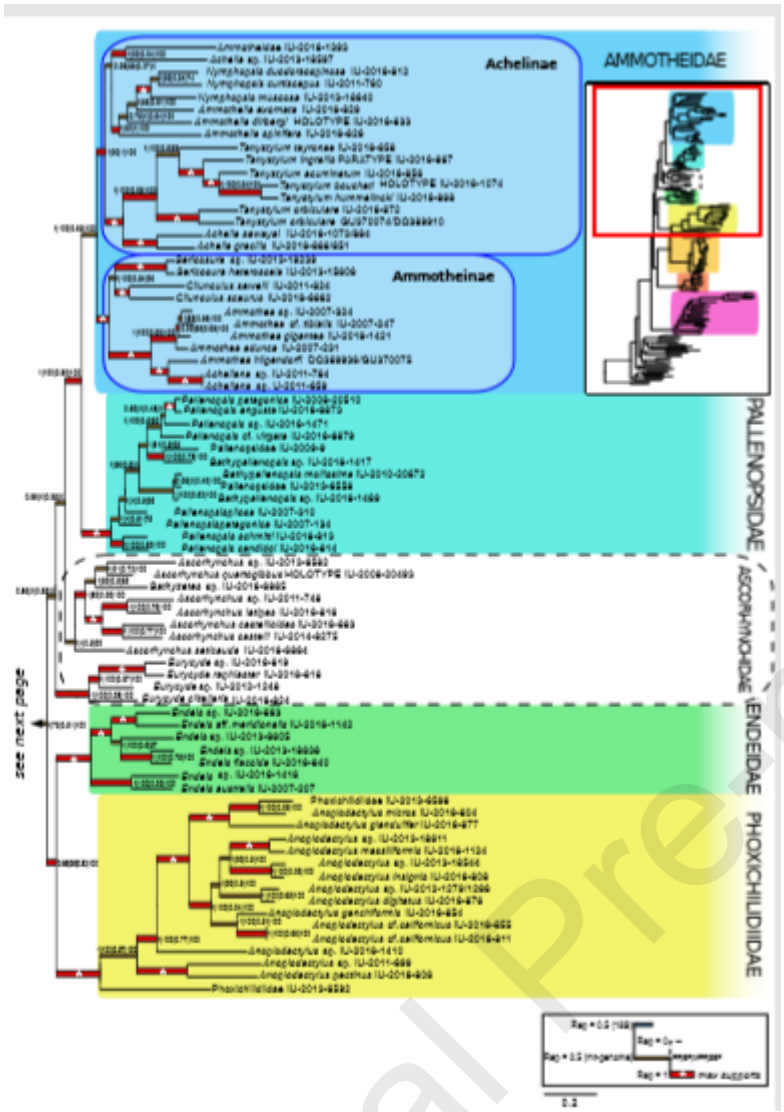
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.



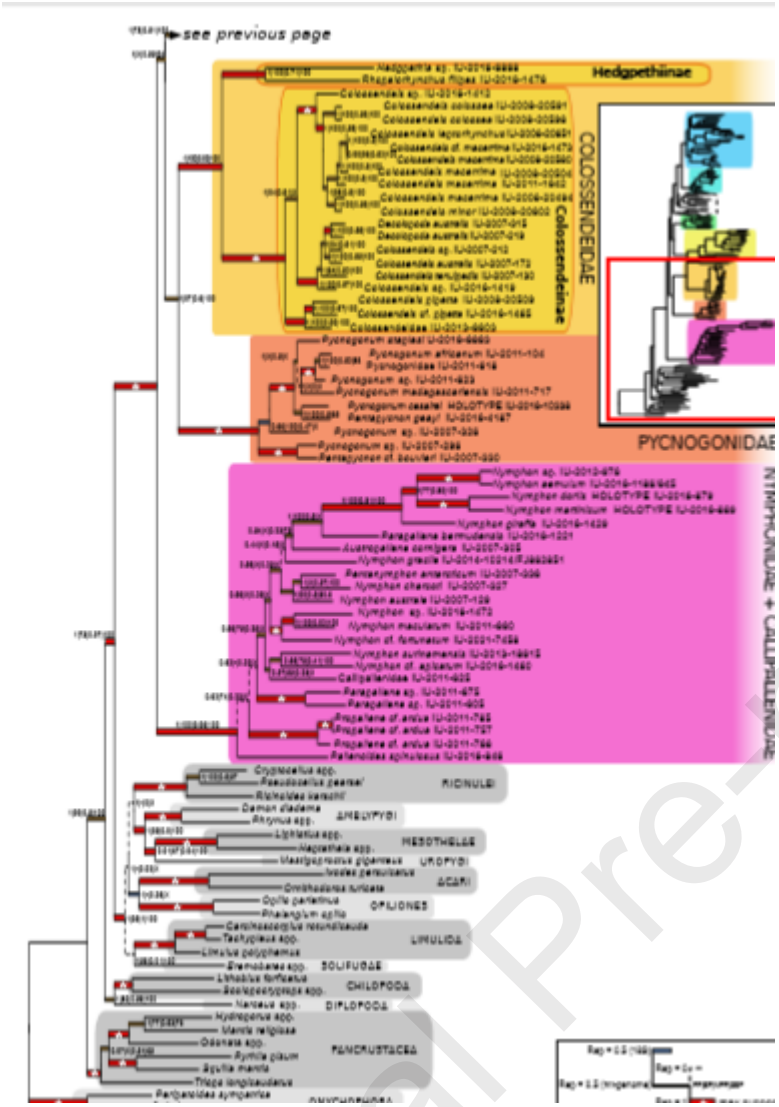
1



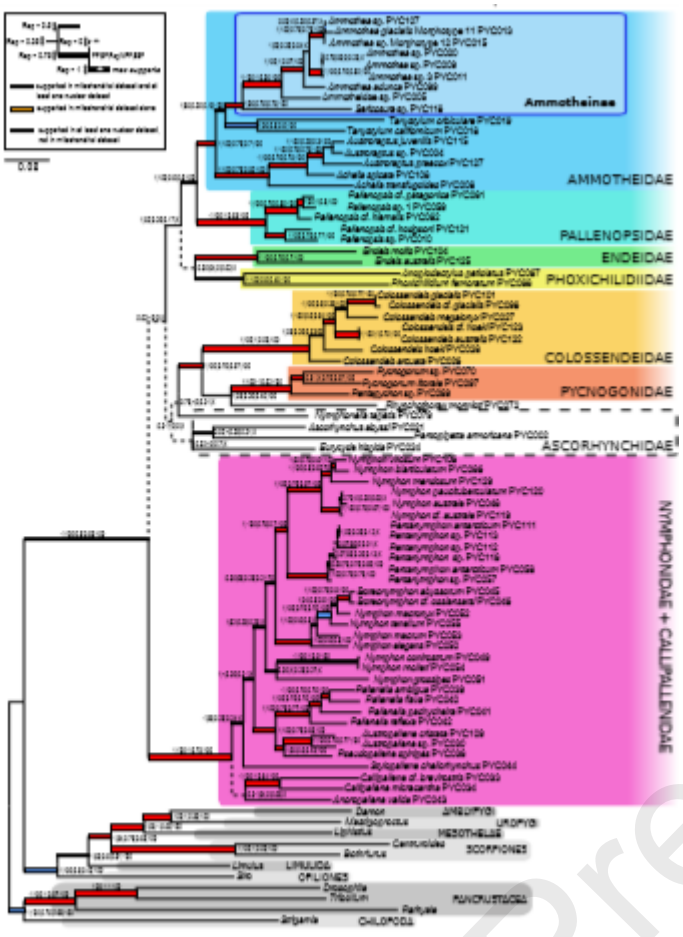
1



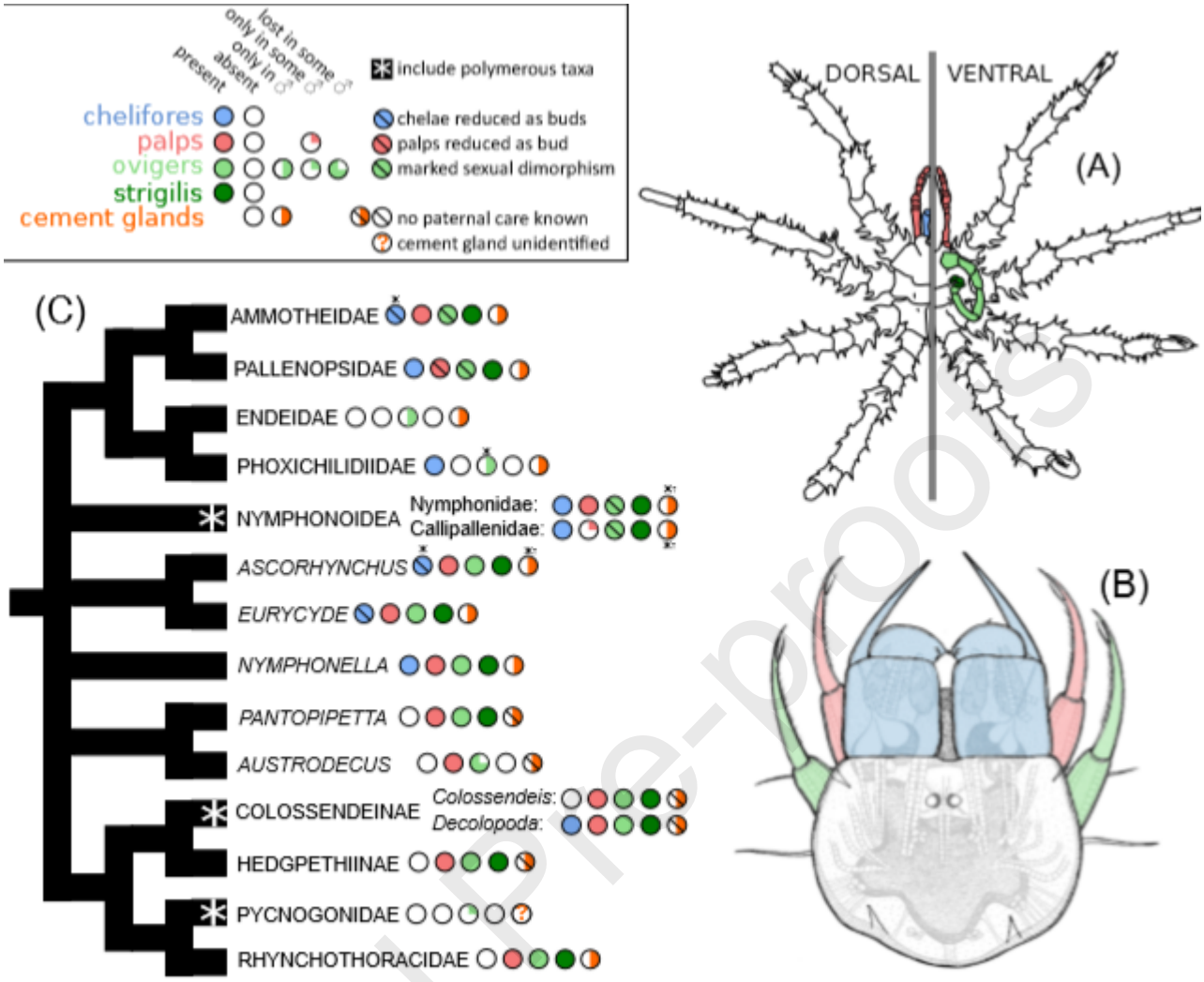
1


see previous page
Hedgpethiinae
COLOSSENDEIDAE
Colossendeinae
PYCNOGONIDAE
NYMPHONIDAE + CALLIPALLENIDAE
RICINULEI
AMBLYPYGI
MESOTHELAE
URODYDI
ACARI
OPILIONES
LIMULIDA
SOLIFUGAE
CHILOPODA
DIPLOPODA
PANCRUSTACEA

1



1



1

2