

Title: Phylogenomics illuminates the backbone of the Myriapoda Tree of Life and reconciles morphological and molecular phylogenies

Rosa Fernández^{1,2*}, Gregory D. Edgecombe³ and Gonzalo Giribet¹

¹ Museum of Comparative Zoology & Department of Organismic and Evolutionary Biology, Harvard University, 28 Oxford St., 02138 Cambridge MA, USA

² Current address: Bioinformatics & Genomics, Centre for Genomic Regulation, Carrer del Dr. Aiguader 88, 08003 Barcelona, Spain

³ Department of Earth Sciences, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

*Corresponding author: rfernandezgarcia@g.harvard.edu

The interrelationships of the four classes of Myriapoda have been an unresolved question in arthropod phylogenetics and an example of conflict between morphology and molecules. Morphology and development provide compelling support for Diplopoda (millipedes) and Pauropoda being closest relatives, and moderate support for Symphyla being more closely related to the diplopod-pauropod group than any of them are to Chilopoda (centipedes). In contrast, several molecular datasets have contradicted the Diplopoda–Pauropoda grouping (named Dignatha). Here we present the first transcriptomic data including a pauropod and both families of symphylans, allowing myriapod interrelationships to be inferred from phylogenomic data from representatives of all main lineages. Phylogenomic analyses consistently recovered Dignatha with strong support. Taxon removal experiments identified outgroup choice as a critical factor affecting myriapod interrelationships. Diversification of millipedes in the Ordovician and centipedes in the Silurian closely approximates fossil evidence whereas the deeper nodes of the myriapod tree date to various depths in the Cambrian-Early Ordovician, roughly coinciding with recent estimates of terrestrialisation in other arthropod lineages, including hexapods and arachnids. We provide a summary state-of-the-art Myriapoda Tree of Life at the ordinal level, the resolution of which has been improved by phylogenomics.

Keywords: Pauropoda; Diplopoda; Symphyla; Chilopoda; transcriptomics; Long Branch Attraction

Introduction

The evolutionary interrelationships between and within major arthropod groups were subject to much instability in the early years of molecular phylogenetics. Some hypotheses to emerge from that era – such as crustacean paraphyly with respect to Hexapoda (including insects) – have stood the test of time, whereas others have fallen by the wayside. Controversial results were exposed to be artefacts of insufficient amounts of data, flawed analytical methods, or systematic error. In recent years, phylogenomic approaches drawing on vastly expanded gene and taxon coverage, combined with improved analytical approaches, have seen stable, well supported molecular hypotheses being recovered [1-5], and these are increasingly consistent with classical morphological trees.

Transcriptome-based phylogenies drawing on hundreds or thousands of orthologues have assisted with phylogenetic analyses for the major groups of millipedes [6, 7] and centipedes [7, 8] but relationships between the four main myriapod groups have not been as rigorously tested (Fig. 1). A particular limitation is that only Sanger-sequenced data are available for pauropods—a group that lies at the crux of a molecular and morphological conflict within Myriapoda—probably due to their small size, difficulty in finding them, and cryptic behaviour. Analyses based on 62 nuclear protein coding genes [9, 10] underpin a formal taxonomic proposal that Pauropoda is most closely related to Symphyla, a putative clade named Edafopoda [11]. This grouping, also recovered using nuclear ribosomal genes [12] and mitochondrial genomes [13], is highly unexpected from the perspective of morphology because an alternative grouping of Pauropoda and Diplopoda has been widely accepted for more than a century [14-16]. This hypothesis is named Dignatha (=Collifera), referring to the mandibles and first maxillae being the only functional mouthparts, with the limbless postmaxillary segment forming a so-called collum that is not incorporated into the head. Other shared morphological characters include a first maxilla coalesced with a sternal intermaxillary plate, the *vas deferens* opening to the tips of conical penes between the second trunk leg pair, and the spiracles opening to a tracheal pouch that functions as an apodeme. Early post-embryonic development unites Dignatha based on a motionless pupoid stage immediately after hatching, followed by a hexapodous first free-living stage. The Dignatha hypothesis has also been supported by a few smaller molecular data sets [17], but it has been contradicted by Edafopoda in the analysis of larger data sets [e.g., 11].

In order to investigate the conflicting support for Dignatha versus Edafopoda – and therefore to shed light on the backbone of the Myriapoda Tree of Life – we present the first transcriptomic data including a pauropod and both families of symphylans. The new data are evaluated in a phylogenomic context specifically designed to test these hypotheses. Furthermore, we expand on

previous efforts to date the myriapod phylogenetic tree by coding a morphological character set for the same set of species as sampled transcriptomically as well as key fossil species for their preserved morphological characters in order to estimate the age of diversification within Myriapoda, particularly with reference to the likely timing of terrestrialisation.

Results & Discussion

Pauropoda is the sister group to Diplopoda

All the analyses (with the exception of two, in which Pauropoda was attracted to the outgroups, see below) recovered Pauropoda as the sister group to Diplopoda with high support (Fig. 2a,b). Given the unanimity of support for Dignatha/Collifera in morphological studies, this stable, well-supported result reconciles classical morphological studies with molecules. In analyses based on more intensively sampled or closely related outgroups (discussed below), the sister group of Dignatha is Symphyla, together forming the traditional clade Progoneata. The only two analyses not recovering Dignatha (both maximum likelihood analyses not accounting for among-site rate heterogeneity and including the most distant outgroups) positioned Pauropoda at the base of the ingroup due to a long branch attraction artefact (LBA) (Fig. 2c,d). In fact, one of them even recovered Myriapoda as non-monophyletic (Fig. 2d), with the pauropod spuriously clustering within Pancrustacea, highlighting the potential of LBA in this data set. Edafopoda was not recovered in either of these two analyses, as symphylans fall as the sister group of Diplopoda+Chilopoda in both cases (although without strong support in one of the analysis; Fig. 2c). The attraction of symphylans and pauropods as Edafopoda (the only hypothesis exclusively based on molecular information) is therefore probably due to artefacts during phylogenetic reconstruction, as discussed below.

Outgroup selection impacts on myriapod phylogeny

Despite Dignatha being recovered in most of our analyses, a result not often found in prior molecular analyses, the interrelationships between it and the other two main groups of myriapods varied across analyses, being sensitive to outgroup choice. In the PhyloBayes analyses, matrices 1 and 3 recovered Progoneata, whereas in matrix 2 (in which only the more distantly allied chelicerates were selected as outgroups) symphylans appeared as sister group of the other Myriapoda instead of centipedes, as in the previous cases (Fig. 2a,b). The latter clade formed by these three groups was also recovered in the ML analyses (with or without strong support, though; Fig. 2c,d). This is not the first study in which this result was obtained: in Miyazawa et al. [18]

symphylans were likewise recovered as sister group to all other myriapods, followed by pauropods as sister group to millipedes plus centipedes. However, that study was based on just three Sanger sequenced genes, and conflicts with other well resolved nodes in our phylogeny (i.e., Dignatha). The present study and Fernández et al. [7] suggest that outgroup selection is a major factor affecting phylogenetic reconstruction in myriapods. In addition, the latter study found that the most complete matrices were enriched in ribosomal proteins, and both factors strongly compromised the estimated relationships within the ingroup. In the present study, biases from ribosomal proteins were minimized by using a different orthology inference procedure, which ensures that only single copy genes are selected. In spite of this, it remains the case that the selection of only distant outgroups (chelicerates in this case) yields interrelationships of the myriapod classes that are less congruent with morphology than when closer and more comprehensively sampled outgroups are included. This study also highlights the importance of accounting for site-specific heterogeneity (through the CAT-GTR model of PhyloBayes) at least when taxon sampling is not dense for some of the groups, as even when only closer outgroups are included the long-branched pauropod is attracted to the equally long-branched Pancrustacea. The inclusion of more pauropods may alleviate this effect.

The timing of myriapod diversification

Diversification of Chilopoda (i.e., the basal split in the crown group) is dated to the Early Silurian (Fig. 3), not much earlier than the oldest fossil chilopods in the Late Silurian, these already being representatives of the chelopod crown group. Diversification of Diplopoda dates to the Middle Ordovician (autocorrelated rates)–earliest Silurian (uncorrelated rates). Though this is considerably older than the first millipede body fossils (from the Wenlock Series of the Silurian), it closely approximates the age of trace fossils that have been attributed to Diplopoda and especially compared to locomotion in Penicillata [19, 20]. In contrast, deeper nodes associated with the divergences between myriapod classes are substantially older than available fossil data. No plausible total-group myriapod body fossils are known from the Cambrian, but as in previous studies dating Myriapoda [3, 17], some deep splits are estimated to be of Cambrian age. Diversification of Dignatha is inferred to date to the latest Cambrian–Early Ordovician, Progoneata to the mid-late Cambrian, and Myriapoda to the early-middle Cambrian (auto- and uncorrelated rates, respectively). The shared terrestrial adaptations of all extant myriapods (e.g., tracheae, Malpighian tubules, uniramous trunk limbs) suggest that the common ancestor of each of these estimated Cambrian nodes was terrestrial, coinciding (although being slightly younger) with estimates of terrestrialisation for other arthropod lineages, including arachnids and hexapods [3, 5]. Although the trace fossil record is

consistent with amphibious arthropods by the mid Cambrian [21, 22], these molecular estimates for early or middle Cambrian crown-group myriapods continue to pose an unanswered question in arthropod terrestrialisation.

Towards a fully-resolved Myriapoda Tree of Life

The branching pattern of the four main groups of myriapods has been one of the last unresolved questions in arthropod phylogenetics, together with the interrelationships of the chelicerate orders and the exact sequence of crustaceans that led to the origins of hexapods. With the advent of phylogenomic methods, myriapod phylogeny has attracted attention during the last few years, with several studies devoted to shedding light on the interrelationships of the millipede [6] and centipede [8] orders, and more recently expanding taxon sampling to include most centipede families, most millipede orders and a couple of symphylans [7]. The different analyses of large data matrices combined in all these studies (as well as the current one) have allowed us to discern the main artefacts affecting phylogenetic reconstruction in this group of arthropods and to draw a stable myriapod phylogeny (Fig. 4). The tree, largely congruent with traditional hypotheses based on morphology and development, can now be seen as a well-resolved phylogeny with only a handful of uncertain or untested placements, including the unsampled pauropod order Hexamerocerata, and the unexplored position of the diplopod orders Siphoniulida, Siphonophorida and Siphonocryptida. Further efforts towards including these last missing taxa will set the road towards a complete Myriapoda Tree of Life.

Methods

Sample collection and molecular techniques

Fourteen species representing the four major groups of myriapods (Chilopoda, Diplopoda, Pauropoda and Symphyla) were included in this study. Building upon previous work [6, 7], our sampling was designed to maximize representation of all groups, including all orders of centipedes, both families of symphylans, the main clades of millipedes, and pauropods. New sequence data were generated from organisms targeted for their instability or lack of representation in prior analyses: a pauropod (*Pauropus huxleyi*) and a symphylan from the family Scolopendrellidae (Scutigereidae was already represented in earlier studies). Information on sampling localities and accession numbers in the Sequence Read Archive database for each transcriptome can be found in Table 1. The remaining 12 myriapods from Brewer and Bond [6] and our own published data [7]

were available from the Sequence Read Archive (SRA). The following taxa were included as outgroups: a crustacean (*Daphnia pulex*), two hexapods (*Drosophila melanogaster*, *Folsomia candida*), and three chelicerates (*Limulus polyphemus*, *Liphistius malayanus* and *Centruroides vittatus*). The new sequenced cDNA libraries are accessioned in SRA (Table 1). Tissue preservation and RNA sequencing are as described in Fernández et al. [8]. All data sets included in this study were sequenced with the Illumina platform.

Phylogenomic analyses

Single copy genes in arthropods were identified in our data sets with BUSCO v1 [23] based on hidden Markov model profiles. The homologous genes detected were screened to identify multiple hits (i.e., paralogues). Only one homologue per BUSCO single copy gene was selected in each case, assuming that they were single copy in our samples as well, and therefore orthologs. The genes were parsed from each sample and combined into individual files (i.e., one file per gene) with custom python scripts. Alignment, trimming and concatenation were done as in Fernández et al. [7]. As the selection of outgroups may be critical in resolving myriapod relationships we constructed three matrices with different outgroup composition: matrix 1 (300 genes, 49,576 amino acids), including all outgroups (i.e., chelicerates, hexapods and crustaceans); matrix 2 (229 genes, 49,075 amino acids), only with chelicerate outgroups; and matrix 3 (299 genes, 61,611 amino acids), only with pancrustacean outgroups. In all cases, we selected a high level of gene occupancy to ensure the selection of a relatively large amount of genomic information while minimizing missing data and computational burden (75% gene occupancy in matrix 1 and 88% in matrices 2 and 3). Bayesian analyses were conducted with PhyloBayes MPI 1.4e [24] selecting the site-heterogeneous CAT-GTR model of amino acid substitution [25]. Two independent Markov chain Monte Carlo (MCMC) chains were run for 5000–10,000 cycles. The initial 25% of trees sampled in each MCMC run prior to convergence (judged when maximum bipartition discrepancies across chains were < 0.1) were discarded as burn-in. Convergence of chains was assessed both at the level of the bipartition frequencies (with the command `bpcomp`) and the summary variables displayed in the trace files (with the command `tracecomp`). We considered that convergence was achieved when (i) the maximum difference of the frequency of all the bipartitions observed in the chains was < 0.1 , and (ii) when the maximum discrepancy observed for each column of the trace file was < 0.1 and the minimum effective size of 100. A 50% majority-rule consensus tree was then computed from the remaining trees. In order to further test for the effect of heterotachy and heterogeneous substitution rates, the matrices were also analysed in PhyML v.3.0.3 implementing the integrated length (IL) approach [26, 27]. In this

analysis, the starting tree was set to the optimal parsimony tree and the FreeRate model [28] was selected. Congruence between the different topologies was visualized with DensiTree v2.2.5 [29].

Molecular Dating

Divergence times for myriapods were estimated through molecular dating, constrained by the position of critical fossils using a morphological data set. Five Palaeozoic and Mesozoic myriapod fossils (two centipedes and three millipedes; Table 2) were included in our morphological data set of 187 characters. We also included three fossil outgroups: the crustacean *Rehbachella kinnekullensis*, the scorpion *Proscorpius osborni*, and the hexapod *Rhyniella praecursor*. The morphological data set was analysed under parsimony, resulting in 488 trees of 257 steps (Fig. 2e). Absolute dates follow the International Chronostratigraphic Chart v 2015/01. Justifications for age assignments of the fossils (Table 2) follow Wolfe et al. [30]. Divergence dates were estimated using the Bayesian relaxed molecular clock approach as implemented in PhyloBayes v.3.3f [24]. Both an auto-correlated and uncorrelated relaxed clock model were applied to our dataset. The calibration constraints specified above were used with soft bounds [31] under a birth-death prior in PhyloBayes. Two independent MCMC chains were run for 5000–7,000 cycles, sampling posterior rates and dates every 10 cycles. The initial 25% were discarded as burn-in. Posterior estimates of divergence dates were then computed from the remaining samples of each chain.

Author contributions

All authors designed the study. RF and GG collected the samples. RF executed molecular work and data analysis. GDE coded the morphological data. All authors wrote the manuscript.

Acknowledgements

Miquel Arnedo, Ligia Benavides, Brian Colby and Julia Cosgrove assisted with pauropod collection. This study was mainly supported by internal funds from the Museum of Comparative Zoology, Harvard University and by NSF grant DEB-1457539 to GG, which funded RF.

Competing financial interests statement

The authors declare that they have no competing interests.

Data availability statement

All data sets have been deposited to the NCBI SRA data base. Accession numbers are listed in Table 1.

References

1. Rota-Stabelli O, Campbell, L, Brinkmann, H, Edgecombe, GD, Longhorn, SJ, Peterson, KJ, Pisani, D, Philippe, H, Telford, MJ. 2011 A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support monophyletic Mandibulata. *Proc. R. Soc. B Biol. Sci.* **278**, 298-306. (doi:10.1098/rspb.2010.0590)
2. Rota-Stabelli O, Daley, AC, Pisani, D. 2013 Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Curr. Biol.* **23**, 392-398. (doi:10.1016/J.Cub.2013.01.026)
3. Lozano-Fernandez J, Carton, R, Tanner, AR, Puttick, MN, Blaxter, M, Vinther, J, Olesen, J, Giribet, G, Edgecombe, GD, Pisani, D. 2016 A molecular palaeobiological exploration of arthropod terrestrialisation. *Philos. Trans. Roy. Soc. B Biol. Sci.* **371**, 20150133. (doi:10.1098/rstb.2015.0133)
4. Misof B, Liu, S, Meusemann, K, Peters, RS, Donath, A, Mayer, C, Frandsen, PB, Ware, J, Flouri, T, Beutel, RG *et al.* 2014 Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763-767. (doi:10.1126/science.1257570)
5. Schwentner M, Combosch, DJ, Nelson, JP, Giribet, G. 2017 A phylogenomic solution to the origin of insects by resolving crustacean-hexapod relationships. *Curr. Biol.* (doi:10.1016/j.cub.2017.05.040)
6. Brewer MS, Bond, JE. 2013 Ordinal-level phylogenomics of the arthropod class Diplopoda (millipedes) based on an analysis of 221 nuclear protein-coding loci generated using next-generation sequence analyses. *PLoS One* **8**, e79935. (doi:10.1371/journal.pone.0079935)
7. Fernández R, Edgecombe, GD, Giribet, G. 2016 Exploring phylogenetic relationships within Myriapoda and the effects of matrix composition and occupancy on phylogenomic reconstruction. *Syst. Biol.* **65**, 871-889. (doi:10.1093/sysbio/syw041)
8. Fernández R, Laumer, CE, Vahtera, V, Libro, S, Kaluziak, S, Sharma, PP, Pérez-Porro, AR, Edgecombe, GD, Giribet, G. 2014 Evaluating topological conflict in centipede phylogeny using transcriptomic data sets. *Mol. Biol. Evol.* **31**, 1500-1513. (doi:10.1093/molbev/msu108)
9. Regier JC, Shultz, JW, Zwick, A, Hussey, A, Ball, B, Wetzer, R, Martin, JW, Cunningham, CW. 2010 Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**, 1079-1083. (doi:10.1038/nature08742)
10. Regier JC, Zwick, A. 2011 Sources of signal in 62 protein-coding nuclear genes for higher-level phylogenetics of arthropods. *PLoS One* **6**, e23408. (doi:10.1371/journal.pone.0023408)
11. Zwick A, Regier, JC, Zwickl, DJ. 2012 Resolving discrepancy between nucleotides and amino acids in deep-level arthropod phylogenomics: Differentiating serine codons in 21-amino-acid models. *PLoS One* **7**, e47450. (doi:10.1371/journal.pone.0047450)
12. Gai Y-H, Song, D-X, Sun, H-Y, Zhou, K-Y. 2006 Myriapod monophyly and relationships among myriapod classes based on nearly complete 28S and 18S rDNA sequences. *Zool. Sci.* **23**, 1101-1108.
13. Dong Y, Sun, H, Guo, H, Pan, D, Qian, C, Hao, S, Zhou, K. 2012 The complete mitochondrial genome of *Paupopus longiramus* (Myriapoda: Pauropoda): Implications on early diversification of the myriapods revealed from comparative analysis. *Gene* **505**, 57-65. (doi:10.1016/j.gene.2012.05.049)
14. Boudreaux HB. 1987 Phylogeny of the arthropod classes. In *Arthropod phylogeny, with special reference to insects*, pp. 82-122. Malabar, Florida: Robert E. Krieger Publishing Company.
15. Pocock RI. 1893 On the classification of the tracheate Arthropoda. *Zool. Anz.* **16**, 271-275.
16. Tiegs OW. 1947 The development and affinities of the Pauropoda, based on a study of *Paupopus silvaticus*. *Quarterly Journal of Microscopical Sciences* **88**, 275-336.
17. Rehm P, Meusemann, K, Borner, J, Misof, B, Burmester, T. 2014 Phylogenetic position of Myriapoda revealed by 454 transcriptome sequencing. *Mol. Phylogenet. Evol.* **77**, 25-33. (doi:10.1016/j.ympev.2014.04.007)

18. Miyazawa H, Ueda, C, Yahata, K, Su, ZH. 2014 Molecular phylogeny of Myriapoda provides insights into evolutionary patterns of the mode in post-embryonic development. *Sci. Rep.* **4**, 4127. (doi:10.1038/srep04127)
19. Johnson EW, Briggs, DEG, Suthren, RJ, Wright, JL, Tunnicliff, SP. 1994 Non-marine arthropod traces from the subaerial Ordovician Borrowdale Volcanic Group, English Lake District. *Geological Magazine* **131**, 395-406.
20. Wilson HM. 2006 Juliformian millipedes from the Lower Devonian of Euramerica: implications for the timing of millipede cladogenesis in the Paleozoic. *J. Paleontol.* **80**, 638-649.
21. MacNaughton RB, Cole, JM, Dalrymple, RW, Braddy, SJ, Briggs, DEG, Lukie, TD. 2002 First steps on land: Arthropod trackways in Cambrian-Ordovician Eolian sandstone, southeastern Ontario, Canada. *Geology* **30**, 391-394.
22. Collette JH, Gass, KC, Hagadorn, JW. 2012 *Protichnites eremita* unshelled? Experimental model-based neoichnology and new evidence for a euthycarcinoid affinity for this ichnospecies. *J. Paleontol.* **86**, 442-454. (doi:10.1666/11-056.1)
23. Simão FA, Waterhouse, RM, Ioannidis, P, Kriventseva, EV, Zdobnov, EM. 2015 BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212. (doi:10.1093/bioinformatics/btv351)
24. Lartillot N, Rodrigue, N, Stubbs, D, Richer, J. 2013 PhyloBayes MPI: Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611-615. (doi:10.1093/Sysbio/Syt022)
25. Lartillot N, Philippe, H. 2004 A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095-1109. (doi:10.1093/molbev/msh112)
26. Guindon S, Gascuel, O. 2003 A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696-704.
27. Guindon S. 2013 From trajectories to averages: an improved description of the heterogeneity of substitution rates along lineages. *Syst. Biol.* **62**, 22-34. (doi:10.1093/sysbio/sys063)
28. Soubrier J, Steel, M, Lee, MSY, Sarkissian, CD, Guindon, S, Ho, SYW, Cooper, A. 2012 The influence of rate heterogeneity among sites on the time dependence of molecular rates. *Mol. Biol. Evol.* **29**, 3345-3358. (doi:10.1093/Molbev/Mss140)
29. Bouckaert RR, Heled, J. 2014 DensiTree 2: Seeing trees through the forest. *bioRxiv*. (doi:10.1101/012401)
30. Wolfe JM, Daley, AC, Legg, DA, Edgecombe, GD. 2016 Fossil calibrations for the arthropod Tree of Life. *Earth-Sci. Rev.* **160**, 43-110. (doi:10.1101/044859)
31. Yang ZH, Rannala, B. 2006 Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212-226. (doi:10.1093/molbev/msj024)

Figures and Tables

Figure 1. The four main groups of myriapods. A, *Otostigmus* (*Parotostigmus*) *pococki* (Northern Range, Trinidad, Trinidad and Tobago) (Chilopoda, Scolopendromorpha); B, *Hanseniella* sp. (South Island, New Zealand) (Symphyla, Scutigereillidae); C, *Pauropus huxleyi* (Massachusetts, USA) (Pauropoda, Tetramerocerata); and D, *Platydesmus* sp. (La Selva, Costa Rica) (Diplopoda, Platydesmida).

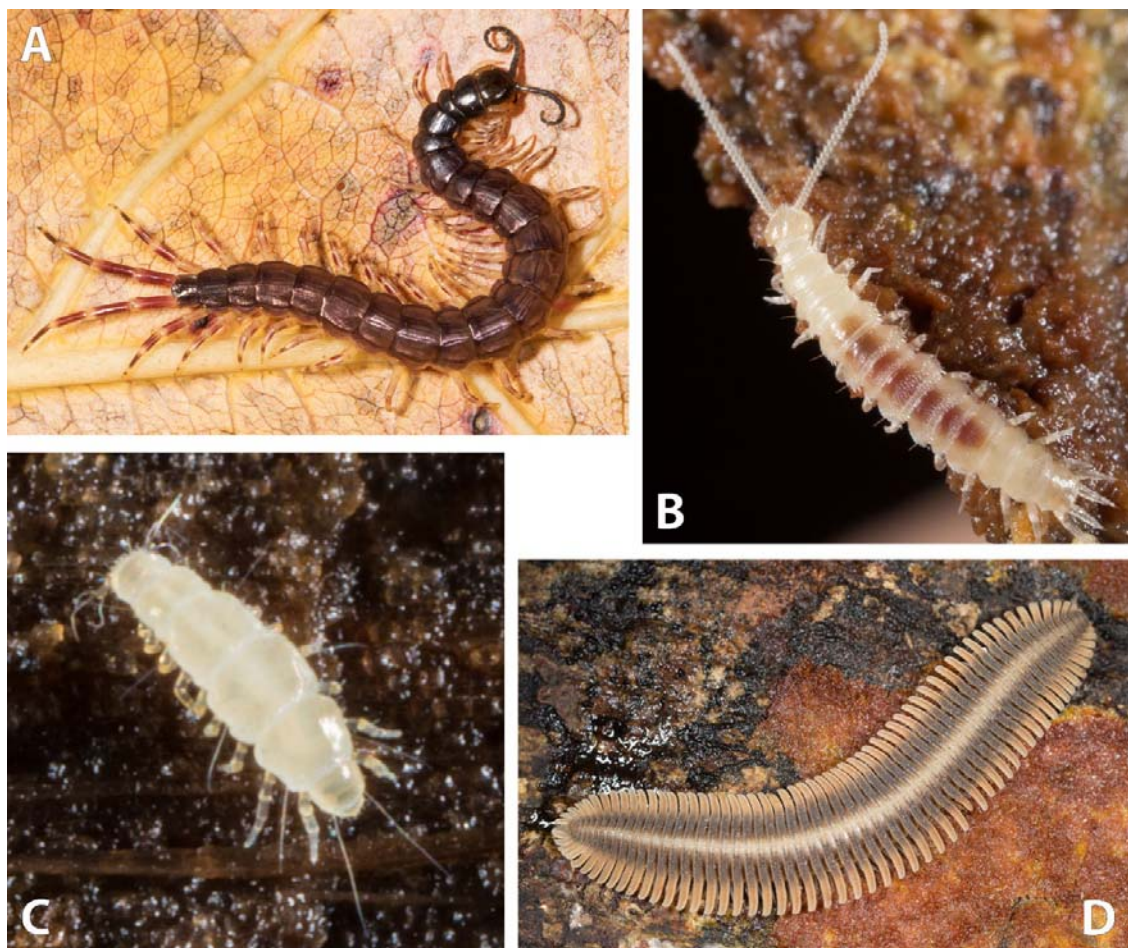


Figure 2A. Preferred phylogenetic hypothesis of myriapod interrelationships (PhyloBayes, matrix 1). **2B.** DensiTree visualization of the four most congruent analyses (PhyloBayes, matrices 2 and 3; PhyML, matrix 3). **2C, 2D.** Main conflicting alternative hypothesis (**2C**, PhyML, matrix 2; **2D**, PhyML, matrix 1). **2E.** Phylogenetic hypothesis of Myriapoda based on 232 morphological characters coded for both extant and extinct species (see Methods for further details); strict consensus of 488 trees of 257 steps; Fossil taxa are identified with a dagger symbol. Black circles in nodes represent high support (> 95% posterior probability, > 90% bootstrap support). CHE: Chelicerata. PAN: Pancrustacea. CHI: Chilopoda. SYM: Symphyla. PAU: Pauropoda. DIP: Diplopoda. Colour codes for each clade are maintained in all figures.

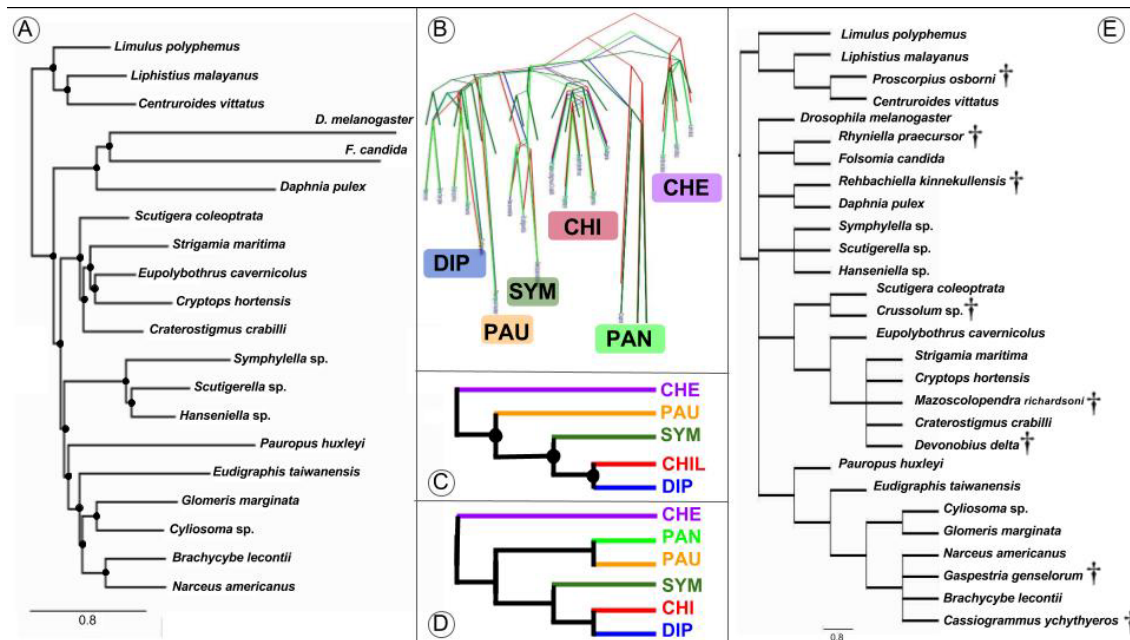


Figure 3. Chronogram of myriapod evolution based on matrix 1 (PhyloBayes analysis) with 95% highest posterior density (HPD) bars for the dating under the uncorrelated (blue) or autocorrelated (pink) model. Nodes that were calibrated with fossils are indicated with a diamond placed at the age of the fossil.

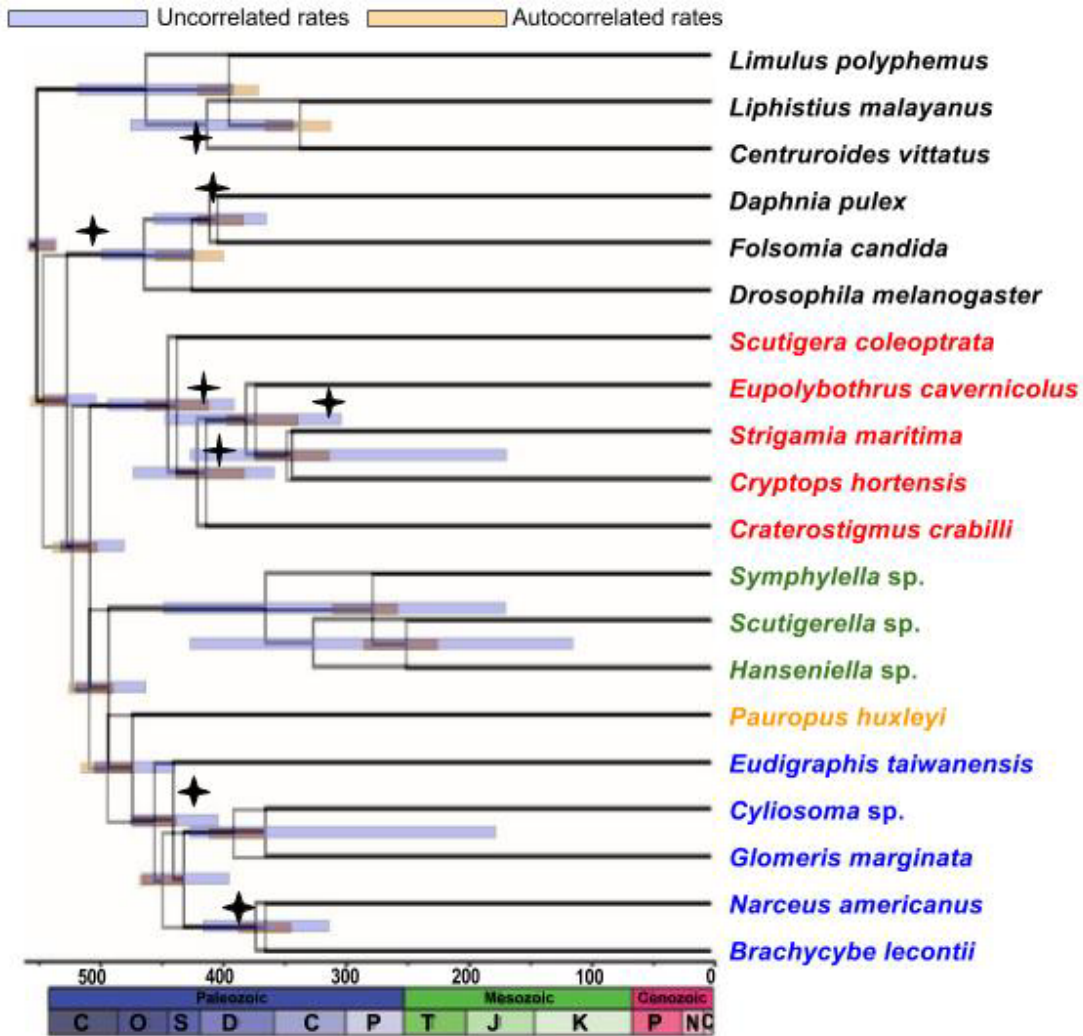


Figure 4. Summary of the Myriapod Tree of Life based as understood by us based on current data. Dashed lines represent unknown placements. Asterisks represent lineages yet to be sequenced with NGS technologies (Hexamerocerata – Pauropoda; Siphoniulida, Siphonophorida, Siphonocryptida – Diplopoda).

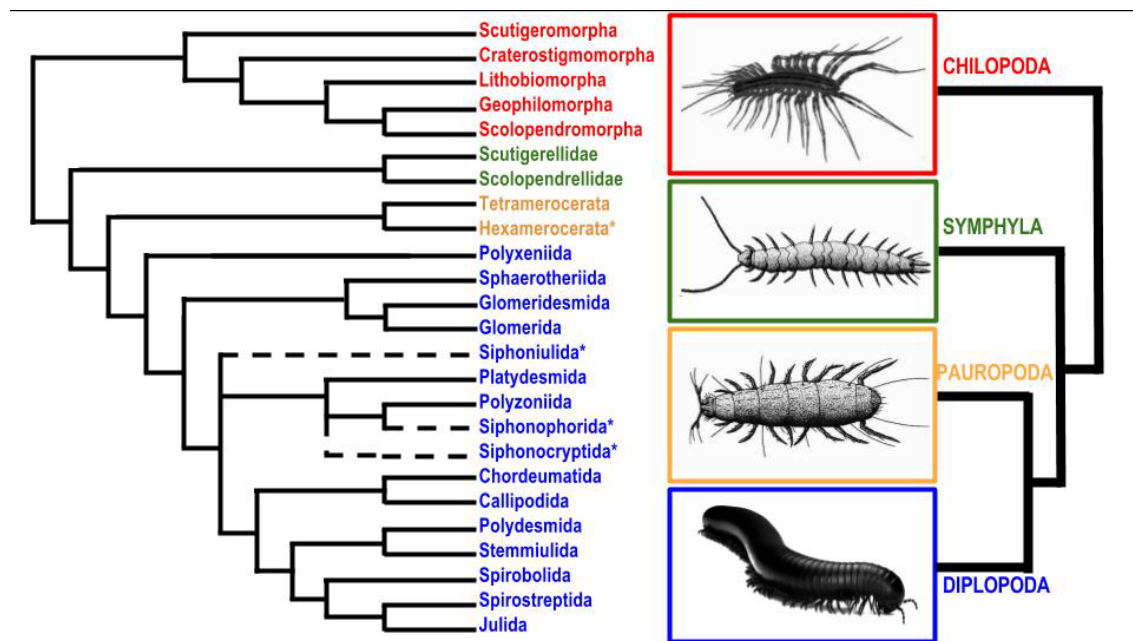


Table 1. List of taxa included in the present study. Location, Museum of Comparative Zoology (MCZ) and SRA accession numbers are indicated.

Species	Data	Source	MCZ Voucher	SRA #
Chilopoda				
<i>Scutigera coleoptrata</i>	Transcriptome	Fernández et al. (2014)	IZ-204015	SRR1158078
<i>Craterostigma crabilli</i>	Transcriptome	Fernández et al. (2016)	IZ-71256	SRR3232915
<i>Cryptops hortensis</i>	Transcriptome	Fernández et al. (2014)	IZ-130583	SRR1153457
<i>Strigamia maritima</i>	Genome	Chipman et al. 2014	-	
<i>Eupolybothrus cavernicolus</i>	Transcriptome	Stoev et al. (2013)	-	ERX311347
Symphyla				
<i>Hanseniella</i> sp.	Transcriptome	Fernández et al. (2016)	IZ-133580	SRR3458649
<i>Scutigarella</i> sp.	Transcriptome	Fernández et al. (2016)	IZ-46890	SRR3458650
<i>Symphylella</i> sp.	Transcriptome	Illumina HiSeq (this study)	IZ-141598	to be assigned
Pauropoda				
<i>Pauropus huxleyi</i>	Transcriptome	Illumina HiSeq (this study)	IZ-141222	to be assigned
Diplopoda				
<i>Eudigraphis taiwanensis</i>	Transcriptome	Fernández et al. (2016)	IZ-128912	SRR3458640
<i>Glomeris marginata</i>	Transcriptome	Fernández et al. (2016)	IZ-43690	SRR3233211
<i>Cyliosoma</i> sp.	Transcriptome	Fernández et al. (2016)	IZ-44064	SRR3458641
<i>Brachycybe</i> sp.	Transcriptome	Brewer & Bond (2013)	-	SRX326776
<i>Narceus americanus</i>	Transcriptome	Fernández et al. (2016)	IZ-44069	SRR3233222
Outgroups				
<i>Limulus polyphemus</i>	Transcriptome	Sharma et al. (2014)	IZ:29738	SRX450966
<i>Liphistius malayanus</i>	Transcriptome	Sharma et al. (2014)	IZ-29742	SRX450965
<i>Centruroides vittatus</i>	Transcriptome	Sharma et al. (2014)	IZ:49754	SRX451012
<i>Daphnia pulex</i>	Genome	Lentz et al. (2014)	-	
<i>Folsomia candida</i>	Genome		-	
<i>Drosophila melanogaster</i>	Genome		-	

Table 2. List of fossil taxa selected for the molecular dating analyses. Their age and placement are indicated.

Fossil taxon	Age (My)	Position
<i>Rehbachella kinnekullensis</i>	497	crown Branchiopoda + Hexapoda
<i>Rhyniella praecursor</i>	405	crown Hexapoda
<i>Casiogrammus ichthyeros</i>	426.9	crown Diplopoda
<i>Devonobius delta</i>	382.7	crown Pleurostigmophora
<i>Mazoscolopendra richardsoni</i>	309.9	crown Amalpighiata
<i>Gaspestria genselorum</i>	392.1	crown Helminthomorpha
<i>Proscorpius osborni</i>	416	crown Scorpiones
<i>Crussolum</i> sp.	408	crown Chilopoda