# Mitochondrial architecture rearrangements produce asymmetrical nonadaptive mutational pressures that subvert the phylogenetic reconstruction in Isopoda

Running title: Nonadaptive evolution and Isopoda mitogenomes

Dong Zhang[1,2#], Hong Zou[1#], Cong-Jie Hua[1], Wen-Xiang Li[1], Shahid Mahboob[3,4], Khalid Abdullah Al-Ghanim[3], Fahad Al-Misned[3], Ivan Jakovlić[5,*], Gui-Tang Wang[1,*]

[1] Key Laboratory of Aquaculture Disease Control, Ministry of Agriculture, and State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, P. R. China

[2] University of Chinese Academy of Sciences, Beijing 100049, P. R. China

[3] Department of Zoology, College of Science, King Saud University, P.O. Box 2455, Riyadh-11451, Saudi Arabia.

[4] Department of Zoology, GC University, Faisalabad, Pakistan

[5] Bio-Transduction Lab, Wuhan 430075, P. R. China

# These authors contributed equally to this study.

*Authors for Correspondence:

Gui-Tang Wang, Key Laboratory of Aquaculture Disease Control, Ministry of Agriculture and State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, 430072, China; Tel: +86-027 68780611, Email: gtwang@ihb.ac.cn

Ivan Jakovlić, Bio-Transduction Lab, Wuhan 430075, China; Tel: +86-15927373207, Email: ivanjakovlic@yahoo.com

## Abstract

The phylogeny of Isopoda, a speciose order of crustaceans, remains unresolved, with different datasets often producing starkly incongruent phylogenetic hypotheses. We hypothesised that extreme diversity in their life histories might be causing compositional heterogeneity/heterotachy in their mitochondrial genomes, and compromising the phylogenetic reconstruction. We tested the effects of different datasets (mitochondrial, nuclear, nucleotides, amino acids, concatenated genes, individual genes, gene orders), phylogenetic algorithms (assuming data homogeneity, heterogeneity, and heterotachy), and partitioning; and found that almost all of them produced unique topologies. As we also found that mitogenomes of Asellota and two Cymothoida families (Cymothoidae and Corallanidae) possess inversed base (GC) skew patterns in comparison to other isopods, we concluded that inverted skews cause long-branch attraction phylogenetic artefacts between these taxa. These asymmetrical skews are most likely driven by multiple independent inversions of origin of replication (i.e., nonadaptive mutational pressures). Although the PhyloBayes CAT-GTR algorithm managed to attenuate some of these artefacts (and outperform partitioning), mitochondrial data have limited applicability for reconstructing the phylogeny of Isopoda. Regardless of this, our analyses allowed us to propose solutions to some unresolved phylogenetic debates, and support Asellota are the most likely candidate for the basal isopod branch. As our findings show that architectural rearrangements can

43 produce major compositional biases even on short evolutionary timescales, the implications
44 are that proving the suitability of data via composition skew analyses should be a
45 prerequisite for every study that aims to use mitochondrial data for phylogenetic
46 reconstruction, even among closely related taxa.

47

48 **Key words:** base composition skew; GC skew; mitochondrial phylogenomics; Cymothoida;
49 replication origin inversion; compositional heterogeneity

50

## 51 Introduction

52 Significant taxonomic and phylogenetic uncertainty permeates the entire order of Isopoda.
53 The members of this highly speciose (>10,000) order of crustaceans (class Malacostraca)
54 exhibit a remarkable diversity in their life histories (comprising both free-living and parasitic
55 species) and occupy almost all habitats on the planet Earth (marine, freshwater and
56 terrestrial), from deep sea vents to the Antarctica. The traditional morphology-based
57 taxonomic classification and identification of isopods is further (aside from the speciosity)
58 hampered by great intraspecific morphological variation, sexual dimorphism, sequential
59 hermaphroditism, relatively flexible host preference, and global distribution of many species
60 (Joca, Leray, Zigler, & Brusca, 2015; L. S. F. Lins, Ho, Wilson, & Lo, 2012; Luana S.F. Lins, Ho, &
61 Lo, 2017; Rudy, Rendoš, Ľuptáčik, & Mock, 2018; Shen et al., 2017; Wetzer, 2002; George D.F.
62 Wilson, 2008). However, molecular data also appear to be an unreliable tool for the task, as

different datasets (mitochondrial genes, mitochondrial genomes, nuclear genes, combined mitonuclear data) often produce very different topologies (Richard C. Brusca, 1981; Hata et al., 2017; Kilpert, Held, & Podsiadlowski, 2012; Luana S.F. Lins et al., 2017; Martin, Bruce, & Nowak, 2016; Poore & Bruce, 2012; Wetzer, 2002; G D F Wilson, 2009; Yu, An, Li, & Boyko, 2018; Zou et al., 2018). As a result, even the identity of the basal isopod clade (defined as the sister-clade to all other isopod lineages (Krell & Cranston, 2004)) remains debated. Traditionally (morphology and single gene-based studies), Phreatoicidea was regarded as the basal clade (Kilpert et al., 2012; Wetzer, 2002; G D F Wilson, 2009), but some studies resolved Phreatoicidea+Aselotta at the base (Shen et al., 2017; G D F Wilson, 2009; George D.F. Wilson, 1999; Yu et al., 2018), one study found Limnoriidea (Luana S.F. Lins et al., 2017) at the base, whereas a few studies even resolved parasitic Cymothoidae and Corallanidae (suborder Cymothoida) at the base (Hua et al., 2018; L. S. F. Lins et al., 2012; Luana S.F. Lins et al., 2017; G D F Wilson, 2009; Zou et al., 2018). As the Cymothoida was traditionally regarded as the most derived isopod clade (Richard C. Brusca, 1981; Kilpert et al., 2012; Wetzer, 2002; G D F Wilson, 2009), these alternative hypotheses cannot be described as minor topological instability. The monophyly of suborder Cymothoida is generally supported by the morphological data, but rejected by the molecular data (Brandt & Poore, 2003; Hua et al., 2018; Kilpert et al., 2012; Luana S.F. Lins et al., 2017; Shen et al., 2017; G D F Wilson, 2009; Yu et al., 2018; Zou et al., 2018). Among a number of other unresolved phylogenetic issues permeating this order are the monophyly of the suborder Oniscidea (supported by morphology, sometimes rejected by molecular data) and the existence of several 'rogue' species/taxa, such as *Ligia oceanica* (Ligiidae), *Eurydice pulchra*, and *Limnoria*

*quadripunctata* (Limnoriidae), whose positions in the isopod clade often vary among studies (Kilpert et al., 2012; Luana S.F. Lins et al., 2017; Schmidt, 2008; Shen et al., 2017; Wetzer, Pérez-Losada, & Bruce, 2013; G D F Wilson, 2009; Yu et al., 2018).

Historically, variation found within gene sequences was typically considered to accumulate under a neutral equilibrium model, so commonly used phylogenetic reconstruction algorithms assume homogeneity in mutational rates. As this paradigm began to change during the last few decades (Wolff, Ladoukakis, Enríquez, & Dowling, 2014), this was accompanied by a growing amount of evidence that compositional heterogeneity can compromise phylogenetic reconstruction in some taxa and that evolutionary models operating under that prerequisite may not be suitable for all phylogenetic studies (Cameron, 2014; Hassanin, 2006; Kolaczkowski & Thornton, 2004; Lartillot, Brinkmann, & Philippe, 2007; Morgan et al., 2013; Phillips, McLenachan, Down, Gibb, & Penny, 2006; Sheffield, Song, Cameron, & Whiting, 2009; Zhong et al., 2011). However, the feud about the most suitable methodological approach to account for this heterogeneity remains unresolved, with most prominent contenders currently being the CAT models (Feuda et al., 2017), and partitioning schemes, i.e. different evolutionary models assigned to different character blocks, assuming homogeneity within each block (Whelan & Halanych, 2017). Although these two approaches account for rate heterogeneity across sites, they still assume that substitution rates for sites are constant across all included lineages. From the evolutionary perspective this is not a likely scenario, as substitution rates are likely to be both site- and lineage-specific (Crotty et al., 2017). Indeed, heterotachy, variations in lineage-specific evolutionary rates over time

106 (Lopez, Casane, & Philippe, 2002), is widespread in eukaryotes (Baele, Raes, Van De Peer, &
107 Vansteelandt, 2006).

108 Mitochondrial genomes (mitogenomes) generally provide much higher phylogenetic
109 resolution than traditionally used morphological and single gene-based molecular markers
110 (Nie et al., 2018), so mitochondrial phylogenomics is increasingly used to tackle phylogenetic
111 controversies (Bourguignon et al., 2018; Cameron, 2014; Der Sarkissian et al., 2015; Lan et al.,
112 2017; Li et al., 2017). Although the resolution of this approach is still limited by a very small
113 number of available mitogenomes in isopods, overview of published studies shows that they
114 also failed to produce results congruent with other approaches, failed to resolve the rogue
115 taxa issues, and taken as a whole generated more questions than answers (Hua et al., 2018;
116 Kilpert et al., 2012; Luana S.F. Lins et al., 2017; Shen et al., 2017; Yu et al., 2018; Zou et al.,
117 2018).

118 The evolutionary history of Isopoda abounds in independent (major) life history
119 innovations, such as free-living to parasitic lifestyle (Hata et al., 2017; Jones, Miller, Grutter,
120 & Cribb, 2008; Ketmaier, Joyce, Horton, & Mariani, 2008; Poore & Bruce, 2012), and radical
121 habitat expansions (L. S. F. Lins et al., 2012), such as sea to freshwater, or even water to land
122 (Broly, Deville, & Maillet, 2013; Hata et al., 2017; George D.F. Wilson, 2008). It has been
123 proposed that the Cymothoida may have originated in deep seas, subsequently expanded to
124 shallow seas, and then to brackish and freshwater (likely on several independent occasions)
125 (Hata et al., 2017). Signals of adaptation to high altitude (Hassanin, Ropiquet, Couloux, &
126 Cruaud, 2009; Mishmar et al., 2003; Scott et al., 2011), deep-sea environment (Almeida,
127 Maldonado, Vasconcelos, & Antunes, 2015), and shifts in physiological demands

(Botero-Castro et al., 2018; Hassanin, 2006) have been identified in mitogenomes of a range of animals. It is therefore highly likely that radical adaptations to life in different environments, from the anoxic environment of deep sea-inhabiting isopod species (L. S. F. Lins et al., 2012) to terrestrial species, would produce strikingly different evolutionary pressures on genomes of species, and result in disparate evolutionary rates of mitochondrial genes, which are central to energy production via the oxidative phosphorylation (Gawryluk et al., 2016). In agreement with this hypothesis are uneven evolutionary rates (dN/dS) observed among isopod mitogenomes (Shen et al., 2017) and different mutational rates of protein-coding genes (PCGs) encoded on the majority (or plus) strand among different lineages of isopods (Lloyd et al., 2015). Conflicting phylogenetic signals among different mitochondrial regions have been reported in a number of metazoan groups, which indicates that different mitochondrial regions can accumulate substitutions in ways that are difficult to model, which can result in biased estimates of phylogeny (Meiklejohn et al., 2014). There is evidence that this compositional heterogeneity may be comparatively highly pronounced in mitogenomes of some arthropod taxa (Cameron, 2014; Hassanin, 2006; Liu, Li, Jakovlić, & Yuan, 2017). We therefore hypothesised that the aforementioned extreme life history diversity of isopods might cause pronounced compositional heterogeneity/heterotachy in their mitogenomes, and interfere with the reconstruction of the Isopoda phylogeny. Although limitations of mitochondrial (and molecular in general) data for inferring phylogenies have been widely discussed (Ballard & Whitlock, 2004; Edwards, Potter, Schmitt, Bragg, & Moritz, 2016; Grechko, 2013; Hassanin, Léger, & Deutsch, 2005; Rubinoff, Holland, & Savolainen, 2005; Talavera et al., 2011; Willis, 2017), a review of the existing literature

reveals that most previous studies of evolutionary history of Isopoda ignored those limitations, or attempted to ameliorate them by using such strategies combined datasets (mtDNA, nuclear DNA, morphology) (Luana S.F. Lins et al., 2017; G D F Wilson, 2009), amino acid sequences (Kilpert et al., 2012; Luana S.F. Lins et al., 2017), or applying different models to each codon position (Hata et al., 2017). However, none of those studies attempted to use algorithms designed specifically to account for compositional heterogeneity/tachy, nor studied this problem directly. To test the hypothesis that compositional heterogeneity interferes with phylogenetic reconstruction in Isopoda, we used a number of different datasets: mitochondrial DNA (single genes, genomes, nucleotides, amino acids, gene orders) and nuclear DNA (*18S*); and methodological approaches: dataset partitioning, maximum likelihood (ML), Bayesian inference (BI), parsimony, PhyloBayes (PB) CAT-GTR model (heterogeneous), and GHOST (heterotachous).

# Materials and Methods

PhyloSuite (Zhang et al., 2018) was used to batch-download all selected mitogenomes from the GenBank, extract genomic features, translate genes into amino acid sequences, semi-automatically re-annotate ambiguously annotated tRNA genes with the help of the ARWEN (Laslett & Canbäck, 2008) output, automatically replace the GenBank taxonomy with the WoRMS database taxonomy, as the latter tends to be more up to date (Costello et al., 2013), generate comparative genome statistics tables, and conduct phylogenetic analyses (Flowchart mode) using a number of incorporated plug-in programs. Nucleotide and amino

acid sequences of protein-coding genes (PCGs) were aligned in batches (using codon and normal-alignment modes respectively) with '--auto' strategy, whereas rRNA genes (including the nuclear *18S*) were aligned using Q-INS-i algorithm, which takes secondary structure information into account, all implemented in MAFFT (Katoh & Standley, 2013; Katoh & Toh, 2008). Gblocks (Castresana, 2000; Talavera & Castresana, 2007) was used to remove ambiguously aligned regions from the concatenated alignments with default parameter settings. PhyloSuite was used to concatenate the alignments (all provided in the File S4). Data partitioning schemes were inferred using PartitionFinder2 (Lanfear, Calcott, Ho, & Guindon, 2012), and selection of the most appropriate evolutionary models for each partition was computed according to the Bayesian information criterion scores and weights using ModelFinder (Kalyaanamoorthy, Minh, Wong, Von Haeseler, & Jermiin, 2017). Chi-square ($\chi 2$) test for the homogeneity of character composition of aligned sequences was performed using IQ-TREE 1.6.8 (Trifinopoulos, Nguyen, von Haeseler, & Minh, 2016). Standard (homogeneous models) phylogenetic analyses were conducted using two programs integrated into PhyloSuite: MrBayes 3.2.6 (Bayesian inference, BI) (Ronquist et al., 2012) and IQ-TREE (Maximum Likelihood, ML). PAUP* 4.0 was used to conduct Parsimony analyses via heuristic searching (TBR branch swapping) and 500 random addition sequence replicates, and bootstrap branch support was calculated via heuristic searches on 1000 pseudo-replicate datasets (Swofford, 2002). The heterogeneous CAT-GTR model is implemented in PhyloBayes-MPI 1.7a (PB) (Lartillot et al., 2007), and the heterotachous GHOST model is implemented in IQ-TREE (Crotty et al., 2017). PhyloBayes was run on the beta version of the Cipres server (https://cushion3.sdsc.edu/portal2/tools.action) (Miller, Pfeiffer, & Schwartz,

2010), with default parameters (burnin = 500, invariable sites automatically removed from the alignment, two MCMC chains), and the analysis was stopped when the conditions considered to indicate a good run (PhyloBayes manual) were reached (maxdiff < 0.1 and minimum effective size > 300). The phylogenetic tree inferred from the gene-order (GO) dataset was reconstructed using MLGO (Hu, Lin, & Tang, 2014), with 1000 bootstrap replicates, and an input file generated by PhyloSuite. Phylograms and gene orders were visualized in iTOL (Letunic & Bork, 2007), and annotated using files generated by PhyloSuite. Skews were calculated and plotted using PhyloSuite and GraphDNA (Thomas, Horspool, Brown, Tcherepanov, & Upton, 2007).

# Results

## Mitochondrial datasets

As a majority of available isopod mitogenomes are incomplete, we were faced with the trade-off between the amount of data used in the analysis and the number of species used: after removing six (some incomplete and duplicates) of the 27 available isopod mitogenomes (Oct. 2018), the dataset comprised 8 complete and 13 partial sequences (Table 1). To attempt to resolve the debated issue of the basal Isopod clade with maximum resolution, we used a relatively large number of outgroups for phylogenetic analyses: a basal arthropod, *Limulus polyphemus* (Lavrov, Boore, & Brown, 2000), and a number of closely related non-isopod malacostracan taxa: three Decapoda, two Stomatopoda, two Amphipoda, one Mysida and one Euphausiacea species (Kilpert & Podsiadlowski, 2006; G D F Wilson, 2009). We conducted phylogenetic analyses using a number of different datasets: NUC - nucleotides

of concatenated 13 PCGs and two rRNA genes (*rrnL* and *rrnS*); AAs – concatenated amino acid sequences of 13 PCGs; 15 single-gene datasets (13 PCGs + 2 rRNAs); gene families (nad+atp and cox+cytb); and gene orders. We also tested the performance of data partitioning, by conducting the same analyses on both non-partitioned and partitioned datasets. The best partitioning scheme divided the NUC dataset by individual genes, with only *cox2*/*cox3*, *atp6*/*nad3*, and *rrnS*/*L* placed together in a same partition. Further details available in Supporting Information (File S1).

## Compositional heterogeneity tests

Best-fit model for the non-partitioned NUC dataset was GTR+F+R6. In the partitioned dataset, three *nad* family genes (*nad1*+*2*+*4*) and *atp8* were assigned the TVM+I+G model, and *nad6* the HKY+I+G model, whereas all nine remaining genes were assigned the GTR+I+G model (File S1). In the $\chi 2$ compositional homogeneity test of the non-partitioned NUC dataset, where a sequence was denoted 'failed' if its nucleotide composition significantly deviated from the average composition of the alignment, only an outgroup species *Penaeus vannamei* passed the test (30 sequences failed). However, only 11 species failed the test in the AAs dataset, indicating that the use of amino acids can attenuate compositional heterogeneity.

## Methodological approaches to infer phylogeny

We tested the performance of two phylogenetic methods relying on standard homogenous models, maximum likelihood (ML) and Bayesian inference (BI), on non-partitioned and partitioned data, and non-standard heterogeneous (CAT-GTR) and heterotachous (GHOST) models on our datasets. The latter two models require the input data to be non-partitioned.

The CAT-GTR site mixture model implemented in Phylobayes-MPI 1.7a (Lartillot et al., 2007) allows for site-specific rates of mutation, which is considered to be a more realistic model of amino acid evolution, especially for large multi-gene alignments (Maddock et al., 2016). GHOST model is an edge-unlinked mixture model consisting of several site classes with separate sets of model parameters and edge lengths on the same tree topology, thus naturally accounting for heterotachous evolution (Crotty et al., 2017).

## NUC dataset

GHOST, BI and ML analyses (both partitioned and non-partitioned) of the NUC dataset produced highly congruent topologies (referred to as the 'NUC-consensus' topology henceforth). Statistical support values were very high in BI (Fig. 1; all inferred topologies available in the Supporting information: File S2), relatively high in GHOST, and intermediate in ML analyses (low to high). Decapoda were rendered paraphyletic by the Euphausiacea nested within the clade (not in the ML-partitioned tree), and Mysida+Amphipoda were resolved as the sister-clade to Isopoda. In the isopod clade, Cymothoidae+Corallanidae families (Cymothoida) formed the basal clade, followed by Asellota and Phreatoicidea branches. The remaining isopods were divided into two sister-clades: Oniscidea and a 'catch-all' clade comprising Limnoriidea, Valvifera, Sphaeromatidea, *L. oceanica* (a 'rogue' Oniscidea species), and the remaining three Cymothoida species (*Gyge ovalis*, *Bathynomus* sp. and *Eurydice pulchra*), which did not cluster together (Fig. 1). Parsimony analysis produced a slightly rearranged topology, but crucial features were identical: Mysida+Amphipoda were the sister-clade to Isopoda, and Cymothoidae+Corallanidae were basal isopods. Notable differences were: Asellota forming a sister-clade with *G. ovalis* + *L.*

*quadripunctata*, and Phreatoicidea at the base of the catch-all clade. The PB analysis produced a notably different overall topology, with all non-isopod lineages forming a sister-clade to the isopods (monophyletic Decapoda), but the topology of the isopod clade was relatively similar to the NUC-consensus, apart from the rogue *Gyge ovalis* (Fig. 2).

## Single-gene and concatenated gene family datasets

As we hypothesised that the topological instability may be driven by conflicting signals produced by different genes (Meiklejohn et al., 2014), we conducted ML phylogenetic analyses on 15 single-gene datasets (13 PCGs + 2 rRNAs). Surprisingly, all these produced unique topologies (File S2). *Atp6* resolved Asellota+Cymothoidae+Corallanidae as the basal isopod clade; *atp8* (a very small gene) produced an almost nonsensical topology (defined as: in stark disagreement with any reasonable phylogenetic hypothesis); *nad1* resolved Asellota at the base; *nad2* produced a slightly rearranged NUC-consensus topology; *nad3* (also small, ≈350bp) a nonsensical topology; *nad4* produced a rearranged topology (compared to NUC-consensus), but Cymothoidae+Corralanidae were the basal isopod clade; *nad4L* (small gene) produced a highly rearranged isopod clade; despite its large size (≈1700bp); *nad5* produced a non-canonical isopod topology; *nad6* produced a nonsensical topology (paraphyletic Isopoda, rogue *G. ovalis*); *cox1* produced a highly rearranged topology, including both the non-isopod malacostraca and isopod clades, with Phreatoicidea+Cymothoidae+Corralanidae as the basal isopod clade, and non-canonical Valvifera position; *cox2* produced a slightly rearranged NUC-consensus topology; *cox3* produced a highly rearranged topology, with a unique Asellota+Corallanidae clade at the isopod base and *L. quadripunctata* nested within the Oniscidea; *cytb* produced basal

278 Phreatoicidea, and the remaining taxa divided into Oniscidea and 'catch all' sister-clades,
279 where the rogue Cymothoidae+Corallanidae (along with Asellota) were on a long branch in
280 the derived part of the clade; *rrnS* (or *12S*) topology was in some aspects similar to *cytb*, but
281 with Phreatoicidea+Limnoriidea as the basal isopod clade; *rrnL* (or *16S*) produced an almost
282 nonsensical topology, with paraphyletic Isopoda and rogue *G. ovalis*.

283 As some of the previous studies (see Introduction) used datasets combined of several
284 genes, and as some of our optimal partitioning strategy analyses indicated that different
285 gene families might evolve at somewhat congruent rates, we divided the PCG dataset into
286 two concatenated gene families: nad_atp (*nad1-6* and *atp6-8*) and cox_cytb (*cox1-3* and
287 *cytb*). ML analyses of both datasets also produced unique topologies. In comparison to
288 NUC-consensus tree, nad_atp dataset resolved Mysida as the sister-group to all other
289 Malacostraca, and rearranged Phreatoicidea, Limnoriidea and *G. ovalis*. Cox_cytb dataset
290 produced somewhat rearranged non-isopod Malacostraca, a minor discrepancy in the
291 position of *E. pulchra*, and slightly rearranged Oniscidea. Notably, none of the single-gene
292 topologies corresponded to the two family topologies.

293 **AAs dataset: amino acids of 13 PCGs**

294 Amino acids produced topologies that differed from the NUC-consensus one, with instable
295 topology of the non-isopod Malacostraca, including paraphyletic Eucarida
296 (Decapoda+Euphausiacea) in five (BI/ML, non-partitioned/partitioned, and GHOST; Fig. 3)
297 out of six (PB; Fig. 4) analyses. Sister-clade to Isopoda also varied: Mysida+Amphipoda in
298 Parsimony and both BI analyses, and Mysida in the remaining four analyses. In the isopod
299 clade, partitioning had a major effect on the BI analysis: non-partitioned dataset resolved

Cymothoidae+Corallanidae at the base, followed by Phreatoicidea+Asellota. Parsimony analysis produced a similar topology, but Phreatoicidea and *G. ovalis* + *L. quadripunctata* clades switched places. The partitioned BI dataset, as well the remaining five analyses, all resolved Phreatoicidea as the basal clade. In these, the sister-group to the remaining isopods (minus Phreatoicidea) was Asellota+Cymothoidae+Corallanidae in four analyses, and Asellota in PB. Five analyses (minus PB) produced a topology of the remainder of the isopod clade that was partially congruent with the NUC-consensus topology, but instead of it being divided into Oniscidea + all other taxa, here *G. ovalis*+*L. quadripunctata* were at the base (except in Parsimony: Phreatoicidea). Oniscidea (rendered paraphyletic by *L. oceanica*) clade topology was stable in all six, but the topology of the catch-all clade exhibited a number of permutations, none of which were identical to the NUC-consensus topology.

PB produced the only topology (Fig. 4) with monophyletic Oniscidea, with the rogue *L. oceanica* at the base of the clade. Also importantly, Cymothoidae+Corallanidae clade was placed on a long branch within the catch-all clade, together with *Eurydice pulchra* (all Cymothoida). The suborder was still rendered paraphyletic by the positions of *Bathynomus* sp. and rogue (polytomy) *G. ovalis*.

## Gene order

As gene orders are unlikely to exhibit homoplasy, it has been hypothesised that they may be able to resolve difficult (deep) phylogenies in some cases (Boore, 2006), so we tested this approach. As many tRNA genes were missing or we suspected that they may be misannotated, to test for the presence of false signals, we used two datasets: one excluding all tRNAs (PCGs+rRNAs) and one (PCGs+rRNAs+tRNAs) excluding only the missing and

ambiguously annotated tRNAs (such as tRNAL1 and 2). The two datasets produced incongruent topologies; different runs of the PCGs+rRNAs+tRNAs dataset produced identical topologies (Fig. 5), whereas different runs of PCGs+rRNAs did not (File S2: GO_PCGs+rRNAs_1 and 2). Both dataset produced a number of paraphyletic major clades and largely nonsensical topologies and largely low support values; for example, the more stable, PCGs+rRNAs+tRNAs, dataset produced paraphyletic Decapoda, Stomatopoda, Isopoda, Oniscidea and Cymothoida.

## Nuclear marker-based phylogeny (*18S*)

To view the evolutionary history of Isopoda from a non-mitochondrial perspective, we used the nuclear *18S* gene. This approach can also help us test whether the underlying reason for the conflicting phylogenetic signals between different studies might be mitochondrial introgression, which, recent evidence shows, is more widespread than previously thought (Edwards et al., 2016; Jakovlić, Wu, Treer, Šprem, & Gui, 2013; Mallet, Besansky, & Hahn, 2016). Some of the species used for mitogenomic analyses were not available, but we made sure to include representatives of all major taxa in the mitogenomic dataset (File S3). To obtain a more comparable dataset, we sequenced the *18S* gene of three parasitic Cymothoidae/Corallanidae species from the mtDNA dataset: *Cymothoa indica* (Cymothoidae; GenBank accession number MK079664), *Ichthyoxenos japonensis* (Cymothoidae; MK542857), and *Tachaea chinensis* (Corallanidae; MK542858); and an additional Cymothoidae species, *Asotana magnifica* (MK542856). We conducted a phylogenetic analysis using 55 malacostracan orthologues, with SYM+R4 selected as the best-fit model. Only one sequence (non-isopod: *Gammarus troglophilus*) failed the $\chi 2$ compositional homogeneity test. Results

344 produced by this dataset were unstable, i.e. different runs and datasets (addition and
345 removal of some taxa) would usually produce different topologies. However, some important
346 features were rather constant among most analyses: Asellota at the base (Limnoriidea in
347 Parsimony, but with very low support); highly derived Cymothoida, rendered paraphyletic by
348 the nested Limnoriidea (monophyletic in Parsimony); and Oniscidea rendered paraphyletic
349 by the rogue Ligiidae clade (*Ligia* sp. and *Ligidium* sp.). ML and BI analyses produced
350 relatively congruent topologies, with Isopoda rendered paraphyletic by the Amphipoda clade
351 nested within the large Cymothoida clade (File S2). As PB and Parsimony analyses produced
352 monophyletic Isopoda, with Amphipoda as the sister-clade (Fig. 6), we can confidently reject
353 this as a compositional heterogeneity artefact. Phreatoicidea largely clustered with the rogue
354 Ligiidae clade, but this was not supported by the PB analysis. Cymothoida were
355 monophyletic only in the Parsimony analysis, and divided into two clades in all topologies: a
356 stable monophyletic clade comprising Dajidae and Bopyridae; and a large instable (exhibiting
357 pervasive paraphyly) clade comprising Cirolanidae, Corallanidae and Cymothoidae families,
358 and aforementioned 'intruders' (Limnoriidea and Amphipoda). Corallanidae and Cirolanidae
359 were mostly paraphyletic (somewhat erratic behaviour of *Eurydice pulchra*), whereas
360 Cymothoidae (monophyletic) were highly derived and exhibited disproportionately long
361 branches.

## 362 Discussion

363 Most phylogenetic reconstruction algorithms assume homogeneity in nucleotide base
364 compositions, but mitochondrial genomes of some groups of Arthropoda exhibit

non-constant equilibrium nucleotide frequencies across different lineages, or compositional heterogeneity, which can produce artificial clustering in phylogenetic analysis (Hassanin, 2006; Lartillot et al., 2007; Morgan et al., 2013; Talavera et al., 2011). Having observed pervasive topological instability among previous attempts to resolve the phylogeny of Isopoda, we hypothesised that their remarkably different life histories may have produced asymmetrical adaptive evolutionary pressures on their (mito)genomes, which in turn resulted in compositional heterogeneity and heterotachy that interfere with the reconstruction of their evolutionary history.

We tested the performance of several methodologies commonly used to account for these compositional imbalances: the use of amino acid sequences (instead of nucleotide sequences), data partitioning, and phylogenetic algorithms designed for non-homogeneous data (heterogeneous CAT-GTR and heterotachous IQ-GHOST). We also tested commonly used algorithms (ML, BI and Parsimony), and a large number of different datasets. Apart from the heterotachous GHOST model, which tended to produce results identical to the common ML analysis, all other variables produced notable impacts on the topology, with unique topologies by far outnumbering identical topologies.

As regards the major unresolved issue of the isopod phylogeny, the sister-clade to all other isopods (basal clade) and the monophyly of Cymothoida, mitochondrial nucleotides (NUC) quite consistently produced Cymothoidae and Corallanidae as the basal clade, first followed by Asellota, then by Phreatoicidea, whereas other Cymothoida tended to be scattered throughout the central 'catch-all' clade. AAs, however, largely resolved Phreatoicidea as the basal isopod clade, with the remaining isopods split into two

sister-clades: 1) Asellota + Cymothoidae/Corallanidae and 2) all remaining taxa. The Parsimony method produced Cymothoidae+Corallanidae at the base, followed by Asellota + (*G. ovalis* + *L. quadripunctata*) using both datasets. PB analysis of AAs dataset produced a remarkably different topology, with Phreatoicidea at the base, followed by Asellota, but Cymothoidae+Corallanidae were relatively derived, and Cymothoida not so deeply paraphyletic. Nuclear (*18S* gene) topology was also strongly affected by the methodology, but consistently resolved Asellota as the basal isopod clade (Limnoriidea in Parsimony), and Cymothoida as paraphyletic (nested Limnoriidea, not in Parsimony), but highly derived. Even though it appears that we did not manage to reach a conclusion, as we failed to infer a stable topology, we did manage to identify another feature of isopod mitogenomes that may help us in this quest.

## Strand compositional bias: AT and GC skews

Organellar genomes often exhibit a phenomenon known as strand asymmetry, or strand compositional bias, where positive AT skew values indicate more A than T on the strand, positive GC skews indicate more G than C, and vice versa (Reyes, Gissi, Pesole, & Saccone, 1998; Wei et al., 2010). A likely cause for this is hydrolytic deamination of bases on the leading strand when it is single stranded, i.e. during replication and transcription (Bernt, Braband, Schierwater, & Stadler, 2013; Fonseca, Harris, & Posada, 2014; Reyes et al., 1998). Whereas other crustacean taxa usually exhibit positive overall AT skews for genes located on the plus strand (or majority strand) and negative GC skews for genes on the minus strand (minority strand) (Hassanin, 2006; Wei et al., 2010), isopod mitogenomes usually exhibit negative overall AT skews and positive GC skews of the majority strand (Kilpert et al., 2012;

Kilpert & Podsiadlowski, 2006; Yu et al., 2018). This is believed to be a consequence of an inversion of the replication origin (RO), where the changed replication order of two mitochondrial DNA strands consequently resulted in an inversed strand asymmetry (Bernt et al., 2013; Hassanin et al., 2005; Kilpert et al., 2012; Kilpert & Podsiadlowski, 2006; Wei et al., 2010).

It is known from before that *Asellus aquaticus* (Asellota) possesses an inversed skew in comparison to other isopod taxa (Kilpert & Podsiadlowski, 2006), but we found that the three available Cymothoidae and Corallanidae species (*C. indica*, *T. chinensis* and *I. japonensis*) also exhibit inversed skew patterns (Fig. 7, Table 1). As regards other studied non-isopod Malacostraca, they exhibit GC skews very similar to the isopod outliers, from -0.014 in *Metacrangonyx repens* to -0.332 in *Typhlatya miravetensis*, and mixed AT skews (negative in Mysida and Amphipoda, and mixed positive and negative in Stomatopoda and Decapoda). As intra-genomic and inter-specific variations in base composition have a strong power to bias phylogenetic analyses (Romiguier & Roux, 2017), and skew-driven LBA phylogenetic artefacts have been reported in arthropods (Hassanin, 2006; Hassanin et al., 2005) and other metazoans (Sun, Li, Kong, & Yu, 2018), we suspect that inversed skews in some isopods may produce such artefacts among the branches exhibiting similar skews.

## Collation of evolutionary hypotheses

The inversed skew in *Asellus aquaticus* led Kilpert and Podsiadlowski (Kilpert & Podsiadlowski, 2006) to speculate that this clearly suggests that this taxon branched off first in the isopod phylogeny, but a few years later Kilpert et al. (Kilpert et al., 2012) noticed that the basal position of Phreatoicidea causes a conflict in explaining the inversed isopod GC skew

(present in *Eophreatoicus* sp.; Fig. 4 – red stars). Our NUC dataset would mostly support a modified version of the first scenario, the replication origin inversion occurred in isopods after Cymothoidae+Corallanidae and Asellota branched off from the main isopod lineage (Figs. 1 and 2 – red star signs), whereas our AAs analyses would mostly support a scenario where the RO inversion occurred only in the common ancestor of Asellota+(Cymothoidae+Corallanidae) clade (Fig. 3). However, there is no support for either of these scenarios from nuclear (*18S*) or morphological (G D F Wilson, 2009) data, which relatively consistently indicate that Asellota (*18S*) or Asellota+Phreatoicidea (morphology) are the basal clade. Parsimony analyses also complicate these scenarios by Asellota forming a sister-clade with *G. ovalis* and *L. quadripunctata*. Importantly, our PB (heterogeneous algorithm) analysis of the AAs dataset produced an mtDNA topology that exhibited notable similarity to the *18S* and morphology-based topologies, where Cymothoidae+Corallanidae clustered with *E. pulchra* (Cirolanidae) in the relatively derived part of the isopod clade. From that, we can conclude that a combination of AAs dataset (which is expected to be less affected by skews than nucleotides) and a heterogeneous CAT-GTR model was most successful in attenuating the phylogenetic artefacts caused by compositional biases. We can therefore reject the above scenarios, and conclude that inversed skews of Asellota and Cymothoidae+Corallanidae are non-synapomorphic. The inversed skew of highly derived Cymothoidae+Corallanidae produces an LBA artefact of them clustering at the base of the isopod clade, phylogenetically close to other taxa with similar (homoplastic) skew patterns (Asellota and non-isopod Malacostraca).

452 Having established this, now we can use these results to infer the most parsimonious
453 hypothesis for the course of events in the evolutionary history of Isopoda. First, we can
454 reject with confidence the basal position of Cymothoidae+Corallanidae as an artefact. This
455 indicates that the basal isopod taxon is either Asellota (G D F Wilson, 2009), Phreatoicidea (R.
456 C. Brusca & Wilson, 1991), or Asellota+Phreatoicidea sister-clade (Dreyer & Wägele, 2001;
457 Kilpert et al., 2012; George D.F. Wilson, 1999). The latter two scenarios are less parsimonious,
458 as they would require at least three independent RO inversions in the evolutionary history of
459 Isopoda (in the ancestral isopod, in Asellota, and in Cymothoidae+Corallanidae), whereas the
460 first scenario is more parsimonious, as it requires only two (in the ancestral isopod after the
461 split of Asellota and in Cymothoidae+Corallanidae; Fig. 6 – red stars). Additionally, the *18S*
462 dataset relatively consistently resolved Asellota as the basal branch (disregarding
463 paraphyletic Isopoda and Parsimony analysis). Therefore, we can tentatively conclude that
464 multiple evidence supports the original hypothesis of Kilpert and Podsiadlowski (Kilpert &
465 Podsiadlowski, 2006): Asellota is the oldest isopod branch and RO inversion in isopods
466 occurred after the Asellota branched off. This scenario implies either a homoplastic nature of
467 the inverted skews in Asellota and Cymothoidae+Corallanidae, or an introgression event
468 from Asellota. Although the latter scenario would directly explain the phylogenetic affinity
469 between the two taxa, the fact that PB (and especially PB+AAs) analyses managed to
470 attenuate this artefact is a strong indication that we can reject this hypothesis, and conclude
471 that homoplastic skews in these taxa are driven by architectural rearrangements.

472 Although this resolves the issue of the deep paraphyly of Cymothoida, i.e., places the
473 rogue Cymothoidae+Corallanidae clade back within the remaining Cymothoida, *18S* data still

474 resolve the Cymothoida as divided into two clades (Dajidae+Bopyridae and
475 Cirolanidae+Corallanidae+Cymothoidae), and rendered paraphyletic by the nested
476 Limnoriidea. However, as Limnoriidea were resolved as the basal isopod taxon in the *18S*
477 Parsimony analysis, we suspect that this is an LBA between two 'rogue' taxa exhibiting
478 elevated evolutionary rates, and thus erratic phylogenetic behaviour:
479 Corallanidae+Cymothoidae and Limnoriidea. The monophyly of Corallanidae is unsupported
480 by our *18S* analyses, so it will be needed to sequence further mitogenomic (to identify skews)
481 and nuclear molecular data for these three families. This combination of skews and nuclear
482 data would enable us to identify the exact point in evolutionary history where the RO
483 inversion occurred in these taxa, and infer the most parsimonious topology and/or taxonomy,
484 i.e., the one that supports a single RO inversion (or introgression event), as opposed to those
485 that would require multiple events.

486 As regards other unresolved issues in the phylogeny of Isopoda, our analyses further
487 corroborate the existence of several rogue taxa that exhibit somewhat erratic topological
488 behaviour. The position of *Ligia oceanica* (nominally Oniscidea: Ligiidae), a recognized rogue
489 taxon (Luana S.F. Lins et al., 2017; Shen et al., 2017; G D F Wilson, 2009; Yu et al., 2018), was
490 mostly resolved at the base of the 'catch-all' clade in mtDNA analyses, but in the (putatively)
491 most reliable mitochondrial topology AAs+PB, it was resolved as the basal Oniscidea species.
492 Although this is in perfect agreement with morphological data, which resolve Ligiidae as the
493 most primitive Oniscidea clade (Schmidt, 2008), we cannot claim that this issue is fully
494 resolved, because the entire Ligiidae family exhibited rogue behaviour in the *18S* dataset as
495 well. Three Cymothoida taxa in the mtDNA dataset that exhibit standard isopod skews,

496 *Bathynomus* sp., *E. pulchra* (both Cirolanidae) and *G. ovalis* (Bopyridae), also exhibited rather
497 instable topological behaviour, and we did not find support for their monophyly using the
498 mtDNA data. Although the Cirolanidae are believed to be ancient (Wetzer, 2002) and highly
499 plesiomorphic within this suborder (Brandt & Poore, 2003), this is not supported by the *18S*
500 dataset. A putatively relevant observation is that all three species exhibit unique, highly
501 rearranged, gene orders (Fig. 8). As another rogue species, *L. quadripunctata* (Limnoriidea)
502 (Luana S.F. Lins et al., 2017; G D F Wilson, 2009; Zou et al., 2018), also exhibits a highly
503 rearranged gene order (Lloyd et al., 2015), and as there is evidence of a close positive
504 correlation between the mitogenomic architectural instability and the mutation rate
505 (Hassanin, 2006; Shao, Dowton, Murrell, & Barker, 2003; Xu, Jameson, Tang, & Higgs, 2006),
506 we hypothesise that frequent genome rearrangements may have resulted in an accelerated
507 mutational rate in these species. In agreement with this hypothesis, *E. pulchra* and *L.*
508 *quadripunctata* exhibited the highest evolutionary rates of all studied isopod (and decapod)
509 mitogenomes (Shen et al., 2017). However, as *L. quadripunctata* was resolved as the basal
510 isopod clade in one study (mitogenomic AAs dataset + BI analysis;
511 Cymothoidae+Corallanidae unavailable at the time) (Luana S.F. Lins et al., 2017), and
512 Limnoriidea exhibited rather erratic behaviour in our *18S* dataset, we can safely argue that
513 mitogenomic rearrangements are unsatisfactory explanation for some of the observed
514 phenomena. Accelerated rates of substitution in some Arthropoda were previously explained
515 by three main factors: genomic rearrangements, including duplication of the control region
516 and gene translocation, parasitic lifestyle, and small body size (Hassanin, 2006). We therefore
517 hypothesise that a proportion of these phylogenetic artefacts are driven by adaptive

518 evolution, which also produces compositional biases that are interfering with phylogenetic
519 reconstruction by producing disproportionately long branches. Furthermore, inverted
520 mitogenomic skews also do not explain the disproportionately long branch of Cymothoidae
521 in the *18S* dataset. We hypothesise that frequent life-history innovations in Cymothoida
522 (Hata et al., 2017; Luana S.F. Lins et al., 2017) may be producing accelerated substitution
523 rates in some of these taxa. Therefore, we can tentatively conclude that mitochondrial
524 evolution in isopods is a result of interplay between adaptive and nonadaptive evolutionary
525 pressures, where non-adaptive outweigh the adaptive in some taxa, such as Cymothoidae
526 and Corallanidae.

527 The topology of non-isopod Malacostraca was rather instable as well, with PB producing
528 notably different topologies from other analyses. As non-isopod Malacostraca also exhibit
529 notable variability in skews (Table 1), we conclude that compositional heterogeneity also
530 interfered with phylogenetic reconstruction. The most relevant question for this study is that
531 of the sister-group to Isopoda: Mysida+Amphipoda in most of our mtNUC analyses (except
532 PB, a unique topology), BI+AAs, and all Parsimony topologies; Mysida in most AAs analyses
533 (ML+GHOST+PB); and Amphipoda in the PB *18S* analysis (Amphipoda were nested within the
534 Isopoda in ML and BI *18S* topologies). As Amphipoda are considered to be the most
535 prominent contender for this position (G D F Wilson, 2009), this again indicates that a much
536 smaller marker in combination with the CAT-GTR model may be producing more reliable
537 results than complete mitogenomes.

## Methodological implications

Our findings indicate that phylogenetic reconstruction using mitochondrial data in Isopoda is severely hampered by compositional biases. This corroborates an earlier observation that in isopods mitochondrial sequence data are prone to producing artefactual LBA relationships, and thus a poor tool for phylogenetic reconstruction in this taxon (Wetzer, 2002). Intriguingly, these did not affect only the nucleotide dataset, but also the amino acid dataset. Amino acid datasets should be less affected by nonadaptive compositional biases, as nonsynonymous mutations are likely to be affected by the purifying selection (as opposed to synonymous mutations). However, it has been shown that mitochondrial strand asymmetry (skews) can have very large effects on the composition of the encoded proteins (Botero-Castro et al., 2018; Min & Hickey, 2007). Some groups of residues share similar physico-chemical properties and can be conservatively exchanged without significant functional impacts (Botero-Castro et al., 2018), and some parts of protein chains may evolve under lesser functional constraints, so many non-synonymous mutations may not affect the protein functionally. Finally, as argued above, adaptive evolutionary pressures are also likely to affect the composition of proteins in isopods.

All individual genes produced unique topologies, which has important implications for the interpretation of previous results inferred using single-gene datasets. This phenomenon has been observed in isopods before on a much smaller scale (Wetzer, 2002), and it is in agreement with the proposed mosaic nature of (mitochondrial) genomes, where different loci often produce conflicting phylogenetic signals (Degnan & Rosenberg, 2009; Pollard, Iyer, Moses, & Eisen, 2006; Romiguier & Roux, 2017). It should be noted that some genes

560 produced less pronounced phylogenetic artefacts, notably *cytb* and *12S* (File S2), which
561 indicates that these two genes might be evolving under a very strong purifying selection. The
562 latter gene (*12S*) produced a topology that exhibited remarkable congruence with the
563 nuclear *18S* topology, especially in placing Cymothoidae+Corallanidae in the derived part of
564 the clade. Although this corroborates the observation that some genes produce less biased
565 phylogenies than others (Romiguier, Ranwez, Delsuc, Galtier, & Douzery, 2013), the artefact
566 of Asellota clustering within the derived Cymothoida clade shows that compositional biases
567 affected this dataset as well, and that we can safely conclude that single-gene mitochondrial
568 markers are not a suitable tool for this task. Similarly, gene orders produced topological
569 instability, very low support, and almost nonsensical topologies, with paraphyletic Isopoda.
570 As highly rearranged gene orders were at the base of the isopod clade, we hypothesise that
571 the discontinuous evolution of mitogenomic architecture evolution (Zou et al., 2017)
572 produces phylogenetic artefacts, such as LBA. This corroborates the hypothesis that gene
573 orders are not a useful phylogenetic marker in lineages exhibiting destabilised mitogenomic
574 architecture (Zhang et al., 2017; Zou et al., 2017).

575 All of the tested standard models (BI, ML and Parsimony) were very sensitive to
576 compositional biases, and produced highly misleading artefacts. We also tested the
577 performance of a new (experimental) GHOST heterotachous model (Crotty et al., 2017), but
578 we found that it produces results almost identical to the common ML algorithm, so we
579 conclude that on this dataset the model appears to be largely useless. Importantly, as
580 regards the aforementioned unresolved feud about the most suitable methodological
581 approach to account for compositional heterogeneity (Feuda et al., 2017; Whelan &

582 Halanych, 2017), our results indicate that (in isopods) the CAT-GTR model by far outperforms
583 the partitioning (assigning different evolutionary models to different partitions). Although
584 we discourage the use of mitochondrial data as a tool for phylogenetic reconstruction in
585 isopods, there are other available methodological approaches designed to account for this
586 problem (Hassanin, 2006; Richards, Brown, Barley, Chong, & Thomson, 2018; Sheffield et al.,
587 2009; Yang, Li, Dang, & Bu, 2018), so future studies may attempt to test their performance.
588

589 # Conclusions

590 With respect to our working hypothesis, we can accept the first part of it: asymmetrical
591 mutational pressures generate compositional heterogeneity in isopod mitogenomes and
592 interfere with phylogenetic reconstruction. However, we were mistaken in assuming that
593 these mutational pressures are primarily adaptive, i.e., caused by their radically diverse life
594 histories. Our results imply that mitochondrial evolution in isopods is a result of interplay
595 between adaptive and non-adaptive evolutionary pressures, where non-adaptive outweigh
596 the adaptive in some taxa (Cymothoidae and Corallanidae). This is in agreement with a
597 recent observation that mitogenomes in isopods mutate at a rate independent of life history
598 traits (Saclier et al., 2018), and contributes to our understanding of the interplay of adaptive
599 and nonadaptive processes in shaping the mitochondrial genomes of Metazoa (Bernt et al.,
600 2013; Smith, 2016). We can conclude that mitogenomic architectural instability (comprising
601 RO inversions) generates strong compositional biases that render mitogenomic sequence
602 data a very poor tool for phylogenetic reconstruction in Isopoda. The deeply contrasting

603 phylogenetic signals that we identified, not only between nuclear and mitochondrial
604 datasets, but also among different mitogenomic datasets, have important implications both
605 for the interpretation of past studies and for scientists who plan to study the phylogeny of
606 these taxa in the future. Regardless of this, simply by rejecting the previous contradictory
607 hypotheses inferred using mitochondrial data, we managed to at least partially resolve
608 several contentious issues in the phylogeny of Isopoda. As regards the methodological
609 approaches used here to account for compositional heterogeneity, we can conclude that
610 none of the tools managed to fully revolve these biases, but CAT-GTR algorithm
611 outperformed partitioning, and best results were achieved by combining it with the amino
612 acids dataset. Although we discourage the use of mitochondrial data for this purpose, any
613 future study that would aim to rely (even if partially) on mitogenomic data would have to
614 first unambiguously prove that they used tools that successfully account for it. As mtDNA
615 data have played a major role in our current understanding of the evolutionary history of life
616 on Earth (Rubinoff et al., 2005), implications of this study are much broader than its original
617 scope. As our findings show that architectural rearrangements can produce major
618 compositional biases even on short evolutionary timescales, the implications of this study
619 are that proving the suitability of data via GC and AT skew analyses should be a prerequisite
620 for every study that aims to use mitochondrial data for phylogenetic reconstruction, even
621 among closely related taxa. These findings should not discourage scientists from sequencing
622 further isopod mitogenomes, as their architectural hypervariability still makes them a useful
623 tool for unravelling the conundrums of evolution of mitochondrial architecture, and as

624 mitochondrial skews can be used as an additional phylogenetic tool to infer the most
625 parsimonious phylogenetic hypotheses.
626
627
628 **Acknowledgements**

629 This work was supported by the Earmarked Fund for China Agriculture Research System
630 (CARS-45-15); the National Natural Science Foundation of China (31872604, 31572658); and
631 the Deanship of Scientific Research at King Saud University (RGP 1435-012). The funders had
632 no role in the design of the study, collection, analysis and interpretation of data, and in
633 writing the manuscript. The authors declare that they have no conflict of interest. The
634 authors would like to express their sincere appreciation to Chen Rong (Bio-Transduction Lab)
635 for helping us to conduct some of the wet lab experiments, and the Deanship of Scientific
636 Research at King Saud University for funding this research.

637 # References

638 Almeida, D., Maldonado, E., Vasconcelos, V., & Antunes, A. (2015). Adaptation of the
639 mitochondrial genome in cephalopods: Enhancing proton translocation channels and
640 the subunit interactions. *PLoS ONE*, *10*(8), 1–29. doi: 10.1371/journal.pone.0135405
641 Baele, G., Raes, J., Van De Peer, Y., & Vansteelandt, S. (2006). An improved statistical method
642 for detecting heterotachy in nucleotide sequences. *Molecular Biology and Evolution*,
643 *23*(7), 1397–1405. doi: 10.1093/molbev/msl006
644 Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria.
645 *Molecular Ecology*, *13*(4), 729–744. doi: 10.1046/j.1365-294X.2003.02063.x
646 Bernt, M., Braband, A., Schierwater, B., & Stadler, P. F. (2013). Genetic aspects of
647 mitochondrial genome evolution. *Molecular Phylogenetics and Evolution*, *69*(2),
648 328–338. doi: 10.1016/j.ympev.2012.10.020

649 Boore, J. L. (2006). The use of genome-level characters for phylogenetic reconstruction.
650 *Trends in Ecology and Evolution*, *21*(8), 439–446. doi: 10.1016/j.tree.2006.05.009
651 Botero-Castro, F., Tilak, M.-K., Justy, F., Catzeflis, F., Delsuc, F., & Douzery, E. J. P. (2018). In
652 cold blood: compositional bias and positive selection drive the high evolutionary rate
653 of vampire bats mitochondrial genomes. *Genome Biology and Evolution*. doi:
654 10.1093/gbe/evy120
655 Bourguignon, T., Tang, Q., Ho, S. Y. W., Juna, F., Wang, Z., Arab, D. A., … Lo, N. (2018).
656 Transoceanic Dispersal and Plate Tectonics Shaped Global Cockroach Distributions:
657 Evidence from Mitochondrial Phylogenomics. *Molecular Biology and Evolution*, *35*(4),
658 970–983.
659 Brandt, A., & Poore, G. C. B. (2003). Higher classification of the flabelliferan and related
660 Isopoda based on a reappraisal of relationships. *Invertebrate Systematics*, *17*(6),
661 893–923. doi: 10.1071/IS02032
662 Broly, P., Deville, P., & Maillet, S. (2013). The origin of terrestrial isopods (Crustacea: Isopoda:
663 Oniscidea). *Evolutionary Ecology*, *27*(3), 461–476.
664 Brusca, R. C., & Wilson, G. D. F. (1991). A phylogenetic analysis of the Isopoda with some
665 classificatory recommendations. *Memoirs of the Queensland Museum*, *31*, 143–204.
666 Retrieved from http://www.vliz.be/en/imis?module=ref&refid=140503
667 Brusca, Richard C. (1981). A monograph on the Isopoda Cymothoidae (Crustacea) of the
668 eastern Pacific. *Zoological Journal of the Linnean Society*, *73*(2), 117–199. doi:
669 10.1111/j.1096-3642.1981.tb01592.x
670 Cameron, S. L. (2014). Insect mitochondrial genomics: implications for evolution and
671 phylogeny. *Annu Rev Entomol*, *59*, 95–117. doi:
672 10.1146/annurev-ento-011613-162007
673 Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use
674 in Phylogenetic Analysis. *Molecular Biology and Evolution*, *17*(4), 540–552. doi:
675 10.1093/oxfordjournals.molbev.a026334
676 Costello, M. J., Bouchet, P., Boxshall, G., Fauchald, K., Gordon, D., Hoeksema, B. W., …
677 Appeltans, W. (2013). Global Coordination and Standardisation in Marine
678 Biodiversity through the World Register of Marine Species (WoRMS) and Related
679 Databases. *PLOS ONE*, *8*(1), e51629. doi: 10.1371/journal.pone.0051629
680 Crotty, S. M., Minh, B. Q., Bean, N. G., Holland, B. R., Tuke, J., Jermiin, L. S., & Haeseler, A. von.
681 (2017). GHOST: Recovering Historical Signal from Heterotachously-evolved Sequence
682 Alignments. *BioRxiv*. doi: 10.1101/174789
683 Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and
684 the multispecies coalescent. *Trends in Ecology & Evolution*, *24*(6), 332—340. doi:
685 10.1016/j.tree.2009.01.009
686 Der Sarkissian, C., Vilstrup, J. T., Schubert, M., Seguin-Orlando, A., Eme, D., Weinstock, J., …
687 Orlando, L. (2015). Mitochondrial genomes reveal the extinct Hippidion as an
688 outgroup to all living equids. *Biology Letters*, *11*(3), 20141058. doi:
689 10.1098/rsbl.2014.1058

690 Dreyer, H., & Wägele, J.-W. (2001). Parasites of crustaceans (Isopoda: Bopyridae) evolved
691 from fish parasites: Molecular and morphological evidence. *Zoology*, *103*(January
692 2001), 157–178. doi: 0944-2006/01/103/03-04-157
693 Edwards, S. V., Potter, S., Schmitt, C. J., Bragg, J. G., & Moritz, C. (2016). Reticulation,
694 divergence, and the phylogeography–phylogenetics continuum. *Proceedings of the*
695 *National Academy of Sciences*, *113*(29), 8025–8032. doi: 10.1073/pnas.1601066113
696 Feuda, R., Dohrmann, M., Pett, W., Philippe, H., Rota-Stabelli, O., Lartillot, N., … Pisani, D.
697 (2017). Improved Modeling of Compositional Heterogeneity Supports Sponges as
698 Sister to All Other Animals. *Current Biology*, *27*(24), p3864-3870.e4. doi:
699 10.1016/j.cub.2017.11.008
700 Fonseca, M. M., Harris, D. J., & Posada, D. (2014). The Inversion of the Control Region in
701 Three Mitogenomes Provides Further Evidence for an Asymmetric Model of
702 Vertebrate mtDNA Replication. *PLOS ONE*, *9*(9), e106654. doi:
703 10.1371/journal.pone.0106654
704 Gawryluk, R. M. R., Kamikawa, R., Stairs, C. W., Silberman, J. D., Brown, M. W., & Roger, A. J.
705 (2016). The Earliest Stages of Mitochondrial Adaptation to Low Oxygen Revealed in a
706 Novel Rhizarian. *Current Biology*, *26*(20), 2729–2738. doi:
707 10.1016/J.CUB.2016.08.025
708 Grechko, V. V. (2013). The problems of molecular phylogenetics with the example of
709 squamate reptiles: Mitochondrial DNA markers. *Molecular Biology*, *47*(1), 55–74. doi:
710 10.1134/S0026893313010056
711 Hassanin, A. (2006). Phylogeny of Arthropoda inferred from mitochondrial sequences:
712 Strategies for limiting the misleading effects of multiple changes in pattern and rates
713 of substitution. *Molecular Phylogenetics and Evolution*, *38*(1), 100–116. doi:
714 10.1016/j.ympev.2005.09.012
715 Hassanin, A., Léger, N., & Deutsch, J. (2005). Evidence for multiple reversals of asymmetric
716 mutational constraints during the evolution of the mitochondrial genome of metazoa,
717 and consequences for phylogenetic inferences. *Systematic Biology*, *54*(2), 277–298.
718 doi: 10.1080/10635150590947843
719 Hassanin, A., Ropiquet, A., Couloux, A., & Cruaud, C. (2009). Evolution of the mitochondrial
720 genome in mammals living at high altitude: New insights from a study of the tribe
721 Caprini (Bovidae, Antilopinae). *Journal of Molecular Evolution*, *68*(4), 293–310. doi:
722 10.1007/s00239-009-9208-7
723 Hata, H., Sogabe, A., Tada, S., Nishimoto, R., Nakano, R., Kohya, N., … Kawanishi, R. (2017).
724 Molecular phylogeny of obligate fish parasites of the family Cymothoidae (Isopoda,
725 Crustacea): evolution of the attachment mode to host fish and the habitat shift from
726 saline water to freshwater. *Marine Biology*, *164*(5). doi: 10.1007/s00227-017-3138-5
727 Hu, F., Lin, Y., & Tang, J. (2014). MLGO: phylogeny reconstruction and ancestral inference
728 from gene-order data. *BMC Bioinformatics*, *15*(1), 354. doi:
729 10.1186/s12859-014-0354-6
730 Hua, C. J., Li, W. X., Zhang, D., Zou, H., Li, M., Jakovlić, I., … Wang, G. T. (2018). Basal position
731 of two new complete mitochondrial genomes of parasitic Cymothoida (Crustacea:

Isopoda) challenges the monophyly of the suborder and phylogeny of the entire order. *Parasites & Vectors*, *11*(1), 628. doi: 10.1186/s13071-018-3162-4

Jakovlić, I., Wu, Q.-J., Treer, T., Šprem, N., & Gui, J.-F. (2013). Introgression evidence and phylogenetic relationships among three (Para)Misgurnus species as revealed by mitochondrial and nuclear DNA markers. *Archives of Biological Sciences*, *65*(4), 1463–1467. doi: 10.2298/ABS1304463J

Joca, L. K., Leray, V. L., Zigler, K. S., & Brusca, R. C. (2015). A new host and reproduction at a small size for the "snapper-choking isopod" Cymothoa excisa (Isopoda: Cymothoidae). *Journal of Crustacean Biology*, *35*(2), 292–294. doi: 10.1163/1937240X-00002312

Jones, C. M., Miller, T. L., Grutter, A. S., & Cribb, T. H. (2008). Natatory-stage cymothoid isopods: Description, molecular identification and evolution of attachment. *International Journal for Parasitology*, *38*(3–4), 477–491. doi: 10.1016/j.ijpara.2007.07.013

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. doi: 10.1038/nmeth.4285

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. doi: 10.1093/molbev/mst010

Katoh, K., & Toh, H. (2008). Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics*, *9*(1), 212. doi: 10.1186/1471-2105-9-212

Ketmaier, V., Joyce, D. A., Horton, T., & Mariani, S. (2008). A molecular phylogenetic framework for the evolution of parasitic strategies in cymothoid isopods (Crustacea). *Journal of Zoological Systematics and Evolutionary Research*, *46*(1), 19–23. doi: 10.1111/j.1439-0469.2007.00423.x

Kilpert, F., Held, C., & Podsiadlowski, L. (2012). Multiple rearrangements in mitochondrial genomes of Isopoda and phylogenetic implications. *Molecular Phylogenetics and Evolution*, *64*(1), 106–117. doi: 10.1016/j.ympev.2012.03.013

Kilpert, F., & Podsiadlowski, L. (2006). The complete mitochondrial genome of the common sea slater, Ligia oceanica (Crustacea, Isopoda) bears a novel gene order and unusual control region features. *BMC Genomics*, *7*, 241. doi: 10.1186/1471-2164-7-241

Kolaczkowski, B., & Thornton, J. W. (2004). Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature*, *431*, 980.

Krell, F.-T., & Cranston, P. S. (2004). Which side of the tree is more basal? *Systematic Entomology*, *29*, 279–281. doi: 10.1111/j.0307-6970.2004.00262.x

Lan, T., Gill, S., Bellemain, E., Bischof, R., Nawaz, M. A., & Lindqvist, C. (2017). Evolutionary history of enigmatic bears in the Tibetan Plateau-Himalaya region and the identity of the yeti. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1868), 20171804. doi: 10.1098/rspb.2017.1804

773 Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). PartitionFinder: Combined
774 Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses.
775 *Molecular Biology and Evolution*, *29*(6), 1695–1701. doi: 10.1093/molbev/mss020
776 Lartillot, N., Brinkmann, H., & Philippe, H. (2007). Suppression of long-branch attraction
777 artefacts in the animal phylogeny using a site-heterogeneous model. *BMC*
778 *Evolutionary Biology*, *7*(Suppl 1), S4. doi: 10.1186/1471-2148-7-S1-S4
779 Laslett, D., & Canbäck, B. (2008). ARWEN: A program to detect tRNA genes in metazoan
780 mitochondrial nucleotide sequences. *Bioinformatics*, *24*(2), 172–175. doi:
781 10.1093/bioinformatics/btm573
782 Lavrov, D. V, Boore, J. L., & Brown, W. M. (2000). The complete mitochondrial DNA sequence
783 of the horseshoe crab Limulus polyphemus. *Molecular Biology and Evolution*, *17*(5),
784 813–824.
785 Letunic, I., & Bork, P. (2007). Interactive Tree Of Life (iTOL): An online tool for phylogenetic
786 tree display and annotation. *Bioinformatics*, *23*(1), 127–128. doi:
787 10.1093/bioinformatics/btl529
788 Li, H., Leavengood, J. M., Chapman, E. G., Burkhardt, D., Song, F., Jiang, P., … Cai, W. (2017).
789 Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the
790 diversification of true bugs. *Proceedings of the Royal Society B: Biological Sciences*,
791 *284*(1862). doi: 10.1098/rspb.2017.1223
792 Lins, L. S. F., Ho, S. Y. W., Wilson, G. D. F., & Lo, N. (2012). Evidence for Permo-Triassic
793 colonization of the deep sea by isopods. *Biology Letters*, *8*(6), 979–982. doi:
794 10.1098/rsbl.2012.0774
795 Lins, Luana S.F., Ho, S. Y. W., & Lo, N. (2017). An evolutionary timescale for terrestrial isopods
796 and a lack of molecular support for the monophyly of Oniscidea (Crustacea: Isopoda).
797 *Organisms Diversity and Evolution*, *17*(4), 813–820. doi: 10.1007/s13127-017-0346-2
798 Liu, F.-F., Li, Y.-P., Jakovlić, I., & Yuan, X.-Q. (2017). Tandem duplication of two tRNA genes in
799 the mitochondrial genome of Tagiades vajuna (Lepidoptera: Hesperiidae). *European*
800 *Journal of Entomology*, *114*(1), 407–415. doi: 10.14411/eje.2017.052
801 Lloyd, R. E., Streeter, S., Foster, P. G., Littlewood, D. T. J., Huntley, J., Beckham, G. T., … Cragg,
802 S. M. (2015). The complete mitochondrial genome of Limnoria quadripunctata
803 Holthuis (Isopoda Limnoriidae). *Mitochondrial DNA*, *26*(6), 825–826.
804 Lopez, P., Casane, D., & Philippe, H. (2002). Heterotachy, an important process of protein
805 evolution. *Molecular Biology and Evolution*, *19*(1), 1–7. doi:
806 10.1093/oxfordjournals.molbev.a003973
807 Maddock, S. T., Briscoe, A. G., Wilkinson, M., Waeschenbach, A., San Mauro, D., Day, J. J., …
808 Gower, D. J. (2016). Next-generation mitogenomics: A comparison of approaches
809 applied to caecilian amphibian phylogeny. *PLoS ONE*, *11*(6), e0156757. doi:
810 10.1371/journal.pone.0156757
811 Mallet, J., Besansky, N., & Hahn, M. W. (2016). How reticulated are species? *BioEssays*, *38*(2),
812 140–149. doi: 10.1002/bies.201500149
813 Martin, M. B., Bruce, N. L., & Nowak, B. F. (2016). Review of the fish-parasitic genus
814 Cymothoa Fabricius, 1793 (Crustacea: Isopoda: Cymothoidae) from Australia.
815 *Zootaxa*, *4119*(1), 1–72. doi: 10.11646/zootaxa.4119.1.1

816 Meiklejohn, K. A., Danielson, M. J., Faircloth, B. C., Glenn, T. C., Braun, E. L., & Kimball, R. T.
817 (2014). Incongruence among different mitochondrial regions: a case study using
818 complete mitogenomes. *Molecular Phylogenetics and Evolution*, *78*, 314–323. doi:
819 10.1016/j.ympev.2014.06.003
820 Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for
821 inference of large phylogenetic trees. *2010 Gateway Computing Environments*
822 *Workshop, GCE 2010*. doi: 10.1109/GCE.2010.5676129
823 Min, X. J., & Hickey, D. A. (2007). DNA asymmetric strand bias affects the amino acid
824 composition of mitochondrial proteins. *DNA Research: An International Journal for*
825 *Rapid Publication of Reports on Genes and Genomes*, *14*(5), 201–206. doi:
826 10.1093/dnares/dsm019
827 Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A. G., Hosseini, S., … Wallace, D. C.
828 (2003). Natural selection shaped regional mtDNA variation in humans. *Proceedings*
829 *of the National Academy of Sciences*, *100*(1), 171–176. doi:
830 10.1073/pnas.0136972100
831 Morgan, C. C., Foster, P. G., Webb, A. E., Pisani, D., McInerney, J. O., & O'Connell, M. J. (2013).
832 Heterogeneous models place the root of the placental mammal phylogeny.
833 *Molecular Biology and Evolution*, *30*(9), 2145–2156. doi: 10.1093/molbev/mst117
834 Nie, R. E., Breeschoten, T., Timmermans, M. J. T. N., Nadein, K., Xue, H. J., Bai, M., … Vogler, A.
835 P. (2018). The phylogeny of Galerucinae (Coleoptera: Chrysomelidae) and the
836 performance of mitochondrial genomes in phylogenetic inference compared to
837 nuclear rRNA genes. *Cladistics*, *34*(2), 113–130. doi: 10.1111/cla.12196
838 Phillips, M., McLenachan, P., Down, C., Gibb, G., & Penny, D. (2006). Combined mitochondrial
839 and nuclear DNA sequences resolve the interrelations of the major Australasian
840 marsupial radiations. *Systematic Biology*, *55*(1), 122–137. doi:
841 10.1080/10635150500481614
842 Pollard, D. A., Iyer, V. N., Moses, A. M., & Eisen, M. B. (2006). Widespread discordance of
843 gene trees with species tree in Drosophila: evidence for incomplete lineage sorting.
844 *PLoS Genetics*, *2*(10), e173. doi: 10.1371/journal.pgen.0020173
845 Poore, G. C. B., & Bruce, N. L. (2012). Global Diversity of Marine Isopods (Except Asellota and
846 Crustacean Symbionts). *PLoS ONE*, *7*(8), e43529. doi: 10.1371/journal.pone.0043529
847 Reyes, A., Gissi, C., Pesole, G., & Saccone, C. (1998). Asymmetrical directional mutation
848 pressure in the mitochondrial genome of mammals. *Molecular Biology and Evolution*,
849 *15*(8), 957–966. doi: 10.1093/oxfordjournals.molbev.a026011
850 Richards, E. J., Brown, J. M., Barley, A. J., Chong, R. A., & Thomson, R. C. (2018). Variation
851 across mitochondrial gene trees provides evidence for systematic error: How much
852 gene tree variation is biological? *Systematic Biology*. doi: 10.1093/sysbio/syy013
853 Romiguier, J., Ranwez, V., Delsuc, F., Galtier, N., & Douzery, E. J. P. (2013). Less Is More in
854 Mammalian Phylogenomics: AT-Rich Genes Minimize Tree Conflicts and Unravel the
855 Root of Placental Mammals. *Molecular Biology and Evolution*, *30*(9), 2134–2144. doi:
856 10.1093/molbev/mst116
857 Romiguier, J., & Roux, C. (2017). Analytical Biases Associated with GC-Content in Molecular
858 Evolution. *Frontiers in Genetics*, *8*. doi: 10.3389/fgene.2017.00016

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., … Huelsenbeck, J. P. (2012). Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*(3), 539–542. doi: 10.1093/sysbio/sys029

Rubinoff, D., Holland, B. S., & Savolainen, V. (2005). Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. *Systematic Biology*, *54*(6), 952–961. doi: 10.1080/10635150500234674

Rudy, J., Rendoš, M., Ľuptáčik, P., & Mock, A. (2018). Terrestrial isopods associated with shallow underground of forested scree slopes in the Western Carpathians (Slovakia). *ZooKeys*, *801*, 323–335. doi: 10.3897/zookeys.801.24113

Saclier, N., François, C. M., Konecny-Dupré, L., Lartillot, N., Guéguen, L., Duret, L., … Lefébure, T. (2018). Life history traits impact the nuclear rate of substitution but not the mitochondrial rate in isopods. *Molecular Biology and Evolution*, msy184–msy184.

Schmidt, C. (2008). Phylogeny of the terrestrial Isopoda (Oniscidea): a review. *Arthropod Systematics & Phylogeny*, *66*(2), 191–226.

Scott, G. R., Schulte, P. M., Egginton, S., Scott, A. L. M., Richards, J. G., & Milsom, W. K. (2011). Molecular Evolution of Cytochrome c Oxidase Underlies High-Altitude Adaptation in the Bar-Headed Goose. *Molecular Biology and Evolution*, *28*(1), 351–363. doi: 10.1093/molbev/msq205

Shao, R., Dowton, M., Murrell, A., & Barker, S. C. (2003). Rates of Gene Rearrangement and Nucleotide Substitution Are Correlated in the Mitochondrial Genomes of Insects. *Molecular Biology and Evolution*, *20*(10), 1612–1619. doi: 10.1093/molbev/msg176

Sheffield, N. C., Song, H., Cameron, S. L., & Whiting, M. F. (2009). Nonstationary evolution and compositional heterogeneity in beetle mitochondrial phylogenomics. *Systematic Biology*, *58*(4), 381–394. doi: 10.1093/sysbio/syp037

Shen, Y., Kou, Q., Zhong, Z., Li, X., He, L., He, S., & Gan, X. (2017). The first complete mitogenome of the South China deep-sea giant isopod Bathynomus sp. (Crustacea: Isopoda: Cirolanidae) allows insights into the early mitogenomic evolution of isopods. *Ecology and Evolution*, *7*(6), 1869–1881. doi: 10.1002/ece3.2737

Smith, D. R. (2016). The mutational hazard hypothesis of organelle genome evolution: 10 years on. *Molecular Ecology*, *25*, 3769–3755. doi: 10.1111/mec.13742

Sun, S., Li, Q., Kong, L., & Yu, H. (2018). Multiple reversals of strand asymmetry in molluscs mitochondrial genomes, and consequences for phylogenetic inferences. *Molecular Phylogenetics and Evolution*, *118*, 222–231. doi: 10.1016/J.YMPEV.2017.10.009

Swofford, D. L. (2002). PAUP. Phylogenetic analysis using parsimony (and other methods) (Version 4.0). Sunderland, MA, USA: Sinauer Associates.

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–577. doi: 10.1080/10635150701472164

Talavera, G., Vila, R., Igea, J., Castresana, J., Nakano, H., Poustka, A., … Manuel, M. (2011). What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. *BMC Evolutionary Biology*, *11*(1), 315. doi: 10.1186/1471-2148-11-315

902 Thomas, J. M., Horspool, D., Brown, G., Tcherepanov, V., & Upton, C. (2007). GraphDNA: a
903 Java program for graphical display of DNA composition analyses. *BMC Bioinformatics*,
904 *8*, 21. doi: 10.1186/1471-2105-8-21
905 Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: a fast
906 online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*,
907 *44*(W1), W232–W235. doi: 10.1093/nar/gkw256
908 Wei, S. J., Shi, M., Chen, X. X., Sharkey, M. J., van Achterberg, C., Ye, G. Y., & He, J. H. (2010).
909 New views on strand asymmetry in insect mitochondrial genomes. *PLoS ONE*, *5*(9),
910 1–10. doi: 10.1371/journal.pone.0012708
911 Wetzer, R. (2002). Mitochondrial genes and isopod phylogeny (Peracarida: Isopoda). *Journal*
912 *of Crustacean Biology*, *22*(1), 1–14. doi: 10.2307/1549602
913 Wetzer, R., Pérez-Losada, M., & Bruce, N. (2013). Phylogenetic relationships of the family
914 Sphaeromatidae Latreille, 1825 (Crustacea: Peracarida: Isopoda) within
915 Sphaeromatidea based on 18S-rDNA molecular data. *Zootaxa*, *3599*(2), 161–177.
916 Whelan, N. V., & Halanych, K. M. (2017). Who let the CAT out of the bag? Accurately dealing
917 with substitutional heterogeneity in phylogenomic analyses. *Systematic Biology*,
918 *66*(2), 232–255. doi: 10.1093/sysbio/syw084
919 Willis, S. C. (2017). One species or four? Yes!...and, no. Or, arbitrary assignment of lineages to
920 species obscures the diversification processes of Neotropical fishes. *PLoS ONE*, *12*(2),
921 e0172349. doi: 10.1371/journal.pone.0172349
922 Wilson, G D F. (2009). The phylogenetic position of the Isopoda in the Peracarida (Crustacea:
923 Malacostraca). *Arthropod Systematics & Phylogeny*, *67*(2), 159–198.
924 Wilson, George D.F. (1999). Some of the deep-sea fauna is ancient. *Crustaceana*, *72*(8),
925 1019–1030. doi: 10.1163/156854099503915
926 Wilson, George D.F. (2008). Global diversity of Isopod crustaceans (Crustacea; Isopoda) in
927 freshwater. *Hydrobiologia*, Vol. 595, pp. 231–240. doi: 10.1007/s10750-007-9019-z
928 Wolff, J. N., Ladoukakis, E. D., Enríquez, J. A., & Dowling, D. K. (2014). Mitonuclear
929 interactions: evolutionary consequences over multiple biological scales.
930 *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1646).
931 Xu, W., Jameson, D., Tang, B., & Higgs, P. G. (2006). The relationship between the rate of
932 molecular evolution and the rate of genome rearrangement in animal mitochondrial
933 genomes. *Journal of Molecular Evolution*, *63*(3), 375–392. doi:
934 10.1007/s00239-005-0246-5
935 Yang, H., Li, T., Dang, K., & Bu, W. (2018). Compositional and mutational rate heterogeneity in
936 mitochondrial genomes and its effect on the phylogenetic inferences of
937 Cimicomorpha (Hemiptera: Heteroptera). *BMC Genomics*, *19*(1), 264. doi:
938 10.1186/s12864-018-4650-9
939 Yu, J., An, J., Li, Y., & Boyko, C. B. (2018). The first complete mitochondrial genome of a
940 parasitic isopod supports Epicaridea Latreille, 1825 as a suborder and reveals the less
941 conservative genome of isopods. *Systematic Parasitology*. doi:
942 10.1007/s11230-018-9792-2
943 Zhang, D., Gao, F., Li, W. X., Jakovlić, I., Zou, H., Zhang, J., & Wang, G. T. (2018). PhyloSuite: an
944 integrated and scalable desktop platform for streamlined molecular sequence data

945 management and evolutionary phylogenetics studies. *BioRxiv*, 489088. doi:
946 10.1101/489088
947 Zhang, D., Zou, H., Wu, S. G., Li, M., Jakovlić, I., Zhang, J., … Li, W. X. (2017). Sequencing of the
948 complete mitochondrial genome of a fish-parasitic flatworm Paratetraonchoides
949 inermis (Platyhelminthes: Monogenea): tRNA gene arrangement reshuffling and
950 implications for phylogeny. *Parasites & Vectors*, *10*(1), 462. doi:
951 10.1186/s13071-017-2404-1
952 Zhong, B., Deusch, O., Goremykin, V. V, Penny, D., Biggs, P. J., Atherton, R. A., … Lockhart, P. J.
953 (2011). Systematic Error in Seed Plant Phylogenomics. *Genome Biology and Evolution*,
954 *3*, 1340–1348. doi: 10.1093/gbe/evr105
955 Zou, H., Jakovlić, I., Chen, R., Zhang, D., Zhang, J., Li, W.-X., & Wang, G.-T. (2017). The
956 complete mitochondrial genome of parasitic nematode Camallanus cotti: extreme
957 discontinuity in the rate of mitogenomic architecture evolution within the
958 Chromadorea class. *BMC Genomics*, *18*(1), 840. doi: 10.1186/s12864-017-4237-x
959 Zou, H., Jakovlić, I., Zhang, D., Chen, R., Mahboob, S., Al-Ghanim, K. A., … Wang, G. (2018).
960 The complete mitochondrial genome of Cymothoa indica has a highly rearranged
961 gene order and clusters at the very base of the Isopoda clade. *PLOS ONE*, *13*(9),
962 e0203089. doi: 10.1371/journal.pone.0203089
963

## Data Accessibility

DNA sequences: Genbank accessions MK079664, MK542856, MK542857 and MK542858. The remaining data are included within the article and its supplementary files

## Author Contributions

IJ, DZ, HZ, and GTW designed research. DZ, HZ, CJH, WXL, KAAG, FAM, and SM performed research. DZ contributed analytical tools; IJ, DZ, HZ, CJH, WXL, KAAG, FAM, SM and GTW analyzed data. IJ and DZ wrote the paper, and all authors revised it critically for important intellectual content.

## Figures

Figure 1. Mitochondrial phylogenomics of Isopoda (suborder information shown) reconstructed using partitioned nucleotide sequences of PCGs and rRNAs (NUC dataset) and BI algorithm. A set of nine non-isopod Malacostraca species and *Limulus polyphemus* were used as outgroups (order information shown). The scale bar corresponds to the estimated number of substitutions per site. Bayesian posterior support values are shown next to corresponding nodes. Star sign indicates a putative origin of replication inversion scenario implied by the topology (see Discussion).

Figure 2. A phylogram reconstructed using nonpartitioned NUC dataset and an algorithm designed to address compositional heterogeneity: CAT-GTR (PB). Posterior Bayesian support values are shown. See Figure 1 for other details.

Figure 3. A phylogram reconstructed using amino acid dataset (AAs; 13 PCGs) in combination with data partitioning strategy and BI algorithm. See Figure 1 for other details.

Figure 4. A phylogram reconstructed using AAs dataset and CAT-GTR algorithm designed for heterogeneous datasets (PB). See Figure 1 for other details.

985 Figure 5. A phylogram reconstructed using mitochondrial gene orders (PCGs+rRNAs+tRNAs). Bootstrap support values are shown next to corresponding
986 nodes. See Figure 1 for other details.
987
988 Figure 6. A phylogram inferred using the nuclear *18S* gene and CAT-GTR algorithm (PB). See Figure 1 for other details.
989
990 Figure 7. Cumulative GC skews of the majority strands of a selected subset of mitogenomes used for phylogenetic analyses.
991
992 Figure 8. Gene orders in the mitogenomes of Isopoda (and selected Malacostraca).
993
994
995

## Supporting Information

File S1: best partitioning scheme and model selection

File S2: Phylograms produced by all analyses.

File S3: The 18S dataset with taxonomy details.

File S4: Alignments used in this study.

1006 # Tables

1007 Table 1. Taxonomy, length (bp), base composition (%) and skews of mitogenomes used in the analysis. P column indicates whether a species is parasitic. H indicates

1008 the habitat: F - freshwater, M - marine, T - terrestrial, and I - intertidal.

1009

| Species | Suborder | Family | Acc. no. | Length | A+T | A | C | AT skew | GC skew | P | H |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Isopoda** | | | | | | | | | | | |
| *Asellus aquaticus* | Asellota | Asellidae | GU130252 | 13639 | 61.9 | 31 | 21.3 | 0.002 | -0.122 | | F |
| *Cymothoa indica* | Cymothoida | Cymothoidae | MH396438 | 14475 | 63.8 | 36 | 26.1 | 0.129 | -0.442 | √ | M |
| *Ichthyoxenos japonensis* | Cymothoida | Cymothoidae | MF419233 | 15440 | 72.7 | 37 | 18.7 | 0.026 | -0.375 | √ | F |
| *Tachaea chinensis* | Cymothoida | Corallanidae | MF419232 | 14616 | 72.8 | 38 | 18.5 | 0.055 | -0.354 | √ | F |
| *Eurydice pulchra* | Cymothoida | Cirolanidae | GU130253 | 13055 | 55.9 | 27 | 17.7 | -0.052 | 0.198 | | M, I |
| *Bathynomus* sp. | Cymothoida | Cirolanidae | KU057374 | 14965 | 58.7 | 27 | 17 | -0.093 | 0.175 | | M |
| *Gyge ovalis* | Cymothoida | Bopyridae | NC_037467 | 14268 | 59.6 | 27 | 17.8 | -0.093 | 0.118 | √ | M |
| *Eophreatoicus* sp. | Phreatoicidea | Amphisopidae | NC_013976 | 14994 | 69.6 | 31 | 11.4 | -0.104 | 0.25 | | F |
| *Sphaeroma serratum* | Sphaeromatidea | Sphaeromatidae | GU130256 | 13467 | 54.4 | 25 | 17.8 | -0.069 | 0.219 | | M |
| *Limnoria quadripunctata* | Limnoriidea | Limnoriidae | NC_024054 | 16515 | 66.3 | 30 | 13.6 | -0.104 | 0.188 | | M |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Idotea balthica* | Valvifera | Idoteidae | DQ442915 | 14247 | 61 | 28 | 16.3 | -0.076 | 0.163 | M |
| *Glyptonotus antarcticus* | Valvifera | Chaetiliidae | GU130254 | 13809 | 65.4 | 32 | 16.6 | -0.033 | 0.038 | M |
| *Trachelipus rathkii* | Oniscidea | Trachelipodidae | MF187612 | 14080 | 67.3 | 33 | 13.3 | -0.029 | 0.184 | T |
| *Porcellio dilatatus* | Oniscidea | Porcellionidae | KX289582 | 14103 | 65.6 | 31 | 13.4 | -0.067 | 0.224 | T |
| *Cylisticus convexus* | Oniscidea | Cylisticidae | KR013002 | 14154 | 67.8 | 33 | 12.9 | -0.035 | 0.194 | T |
| *Porcellionides pruinosus* | Oniscidea | Porcellionidae | KX289584 | 14078 | 60.5 | 28 | 14.9 | -0.09 | 0.248 | T |
| *Armadillidium album* | Oniscidea | Armadillidiidae | KX289585 | 13812 | 69.7 | 33 | 12.5 | -0.045 | 0.172 | T |
| *Armadillidium nasatum* | Oniscidea | Armadillidiidae | MF187611 | 13943 | 68.1 | 33 | 13.4 | -0.043 | 0.16 | T |
| *Armadillidium vulgare* | Oniscidea | Armadillidiidae | MF187614 | 13932 | 71.5 | 34 | 11.8 | -0.043 | 0.174 | T |
| *Armadillidium vulgare* | Oniscidea | Armadillidiidae | MF187613 | 13955 | 71.3 | 34 | 11.8 | -0.039 | 0.179 | T |
| *Ligia oceanica* | Oniscidea | Ligiidae | NC_008412 | 15289 | 60.9 | 29 | 17 | -0.041 | 0.134 | M, I, T |

| **Species** | **Order** | **Suborder** | **Acc. no.** | **Length** | **A+T** | **A** | **C** | **AT skew** | **GC skew** |
|---|---|---|---|---|---|---|---|---|---|
| **Non-isopod Malacostraca and *L. polyphemus*** | | | | | | | | | |
| *Atergatis floridus* | Decapoda | Pleocyemata | NC_037201 | 16180 | 69.3 | 33 | 20.3 | -0.036 | -0.319 |
| *Penaeus vannamei* | Decapoda | Dendrobranchiata | NC_009626 | 15990 | 67.7 | 33 | 19.2 | -0.026 | -0.192 |
| *Typhlatya miravetensis* | Decapoda | Pleocyemata | NC_036335 | 15865 | 66.2 | 36 | 22.5 | 0.076 | -0.332 |

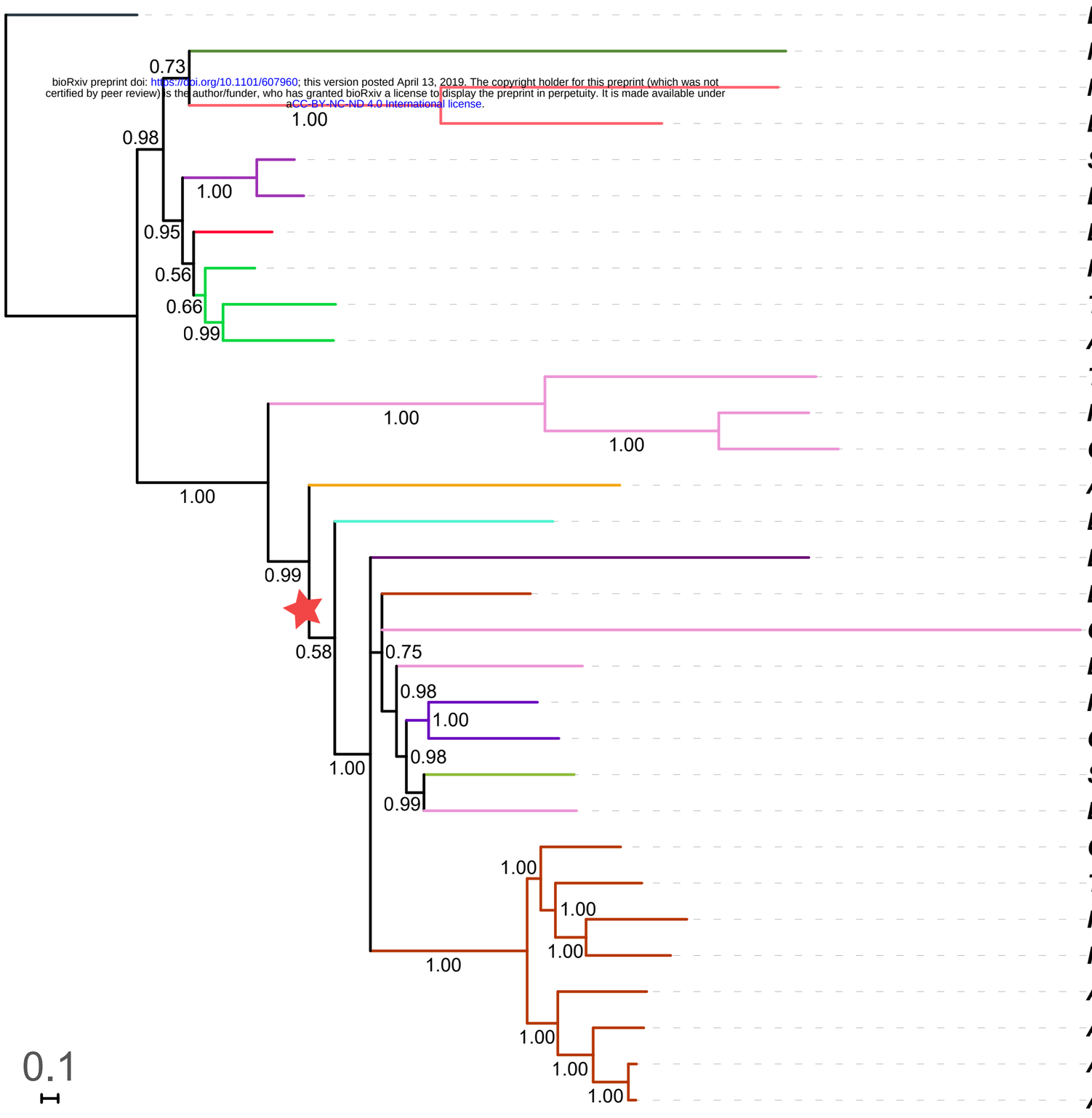
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Squilla mantis* | Stomatopoda | Unipeltata | NC_006081 | 15994 | 70.2 | 35 | 16.8 | -0.001 | -0.13 |
| *Lysiosquillina maculata* | Stomatopoda | Unipeltata | NC_007443 | 16325 | 63.9 | 33 | 21.4 | 0.026 | -0.185 |
| *Metacrangonyx repens* | Amphipoda | Senticaudata | NC_019653 | 14355 | 76.9 | 38 | 11.7 | -0.025 | -0.014 |
| *Neomysis japonica* | Mysida | NA | NC_027510 | 17652 | 74.5 | 37 | 13.8 | -0.021 | -0.085 |
| *Eulimnogammarus cyaneus* | Amphipoda | Senticaudata | NC_033360 | 14370 | 67.6 | 33 | 20.3 | -0.019 | -0.251 |
| *Euphausia pacifica* | Euphausiacea | NA | NC_016184 | 16898 | 72 | 36 | 16 | 0.004 | -0.145 |
| *Limulus polyphemus* | Xiphosurida | NA | NC_003057 | 14985 | 67.6 | 38 | 22.7 | 0.111 | -0.399 |

1010

*Limulus polyphemus*
*Lysiosquillina maculata*
*Squilla mantis*
*Atergatis floridus*
*Typhlatya miravetensis*
*Euphausia pacifica*
*Penaeus vannamei*
*Neomysis japonica*
*Eulimnogammarus cyaneus*
*Metacrangonyx repens*
*Tachaea chinensis*
*Cymothoa indica*
*Ichthyoxenos japonensis*
*Asellus aquaticus*
*Eophreatoicus sp. 14 FK-2009*
*Gyge ovalis*
*Limnoria quadripunctata*
*Ligia oceanica*
*Eurydice pulchra*
*Bathynomus sp.*
*Sphaeroma serratum*
*Glyptonotus cf. antarcticus FK-2009*
*Idotea baltica*
*Cylisticus convexus*
*Trachelipus rathkii*
*Porcellio dilatatus dilatatus*
*Porcellionides pruinosus*
*Armadillidium album*
*Armadillidium nasatum*
*Armadillidium vulgare*
*Armadillidium vulgare*

Xiphosurida
Stomatopoda
Decapoda
Euphausiacea
Decapoda
Mysida
Amphipoda
Cymothoida
Asellota
Phreatoicidea
Cymothoida
Limnoriidea
Oniscidea
Cymothoida
Sphaeromatidea
Valvifera
Oniscidea

1



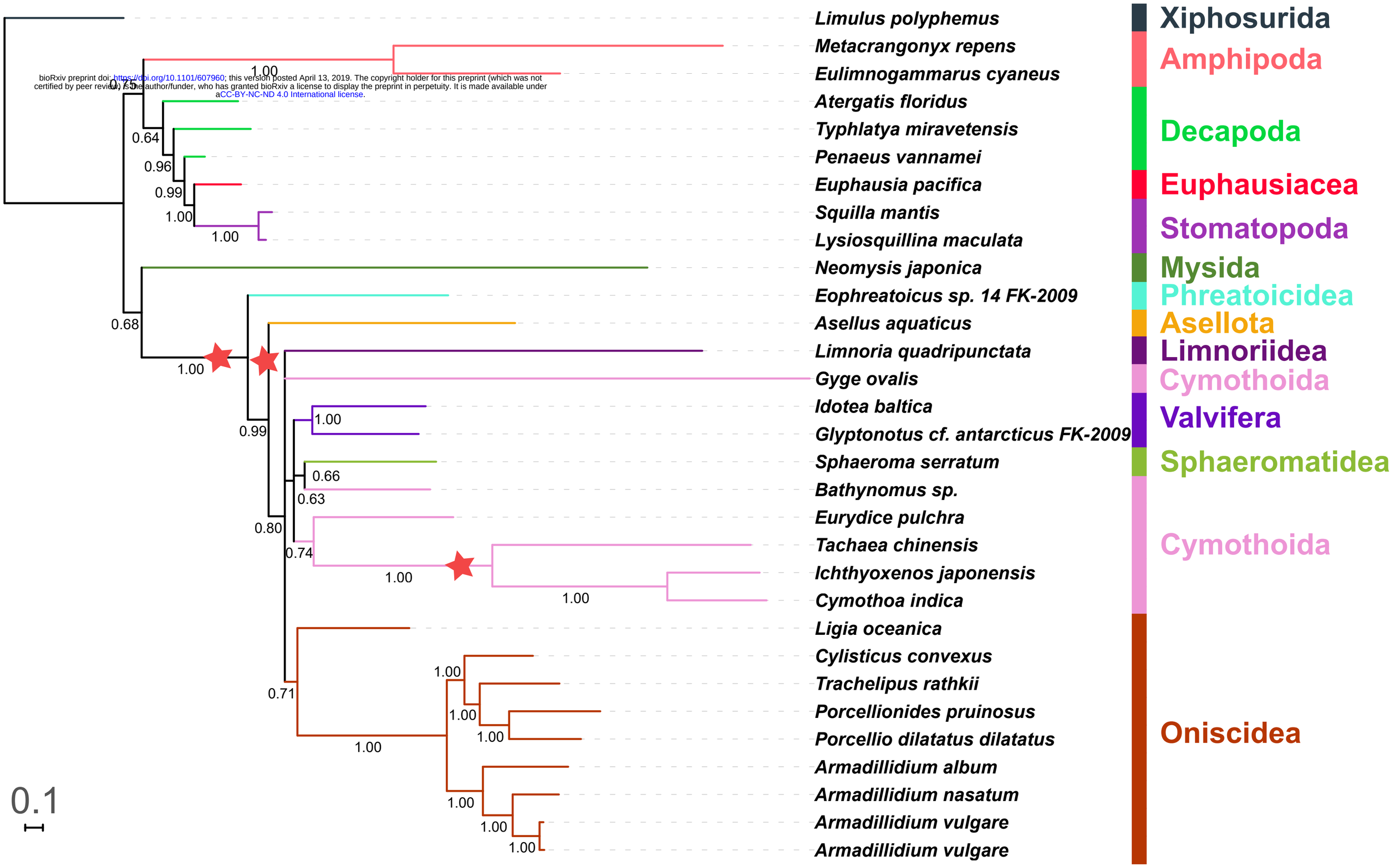

*Limulus polyphemus*
*Neomysis japonica*
*Metacrangonyx repens*
*Eulimnogammarus cyaneus*
*Squilla mantis*
*Lysiosquillina maculata*
*Euphausia pacifica*
*Penaeus vannamei*
*Typhlatya miravetensis*
*Atergatis floridus*
*Tachaea chinensis*
*Ichthyoxenos japonensis*
*Cymothoa indica*
*Asellus aquaticus*
*Eophreatoicus sp. 14 FK-2009*
*Limnoria quadripunctata*
*Ligia oceanica*
*Gyge ovalis*
*Eurydice pulchra*
*Idotea baltica*
*Glyptonotus cf. antarcticus FK-2009*
*Sphaeroma serratum*
*Bathynomus sp.*
*Cylisticus convexus*
*Trachelipus rathkii*
*Porcellionides pruinosus*
*Porcellio dilatatus dilatatus*
*Armadillidium album*
*Armadillidium nasatum*
*Armadillidium vulgare*
*Armadillidium vulgare*

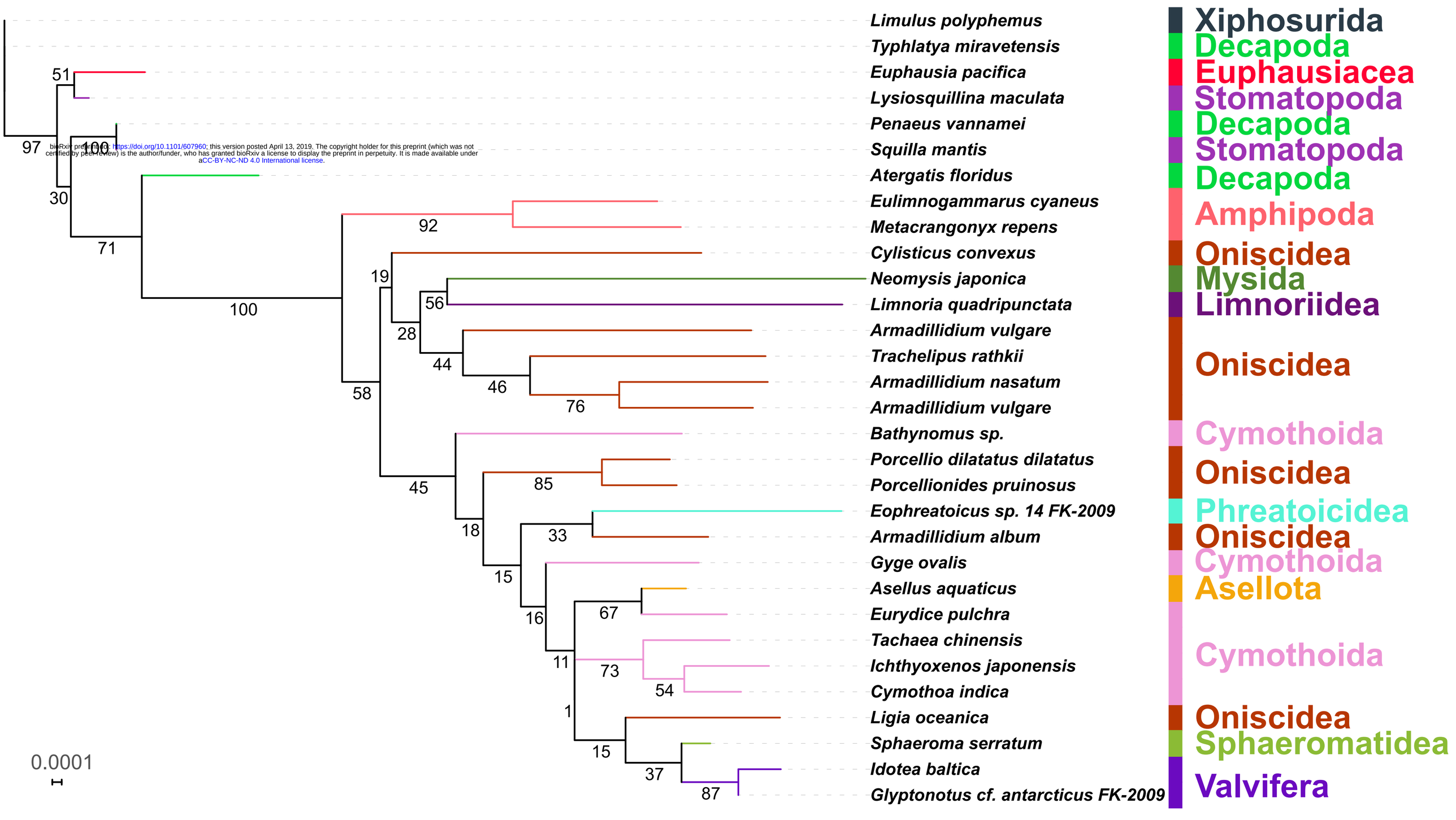
Xiphosurida
Mysida
Amphipoda
Stomatopoda
Euphausiacea
Decapoda
Cymothoida
Asellota
Phreatoicidea
Limnoriidea
Oniscidea
Cymothoida
Valvifera
Sphaeromatidea
Cymothoida
Oniscidea

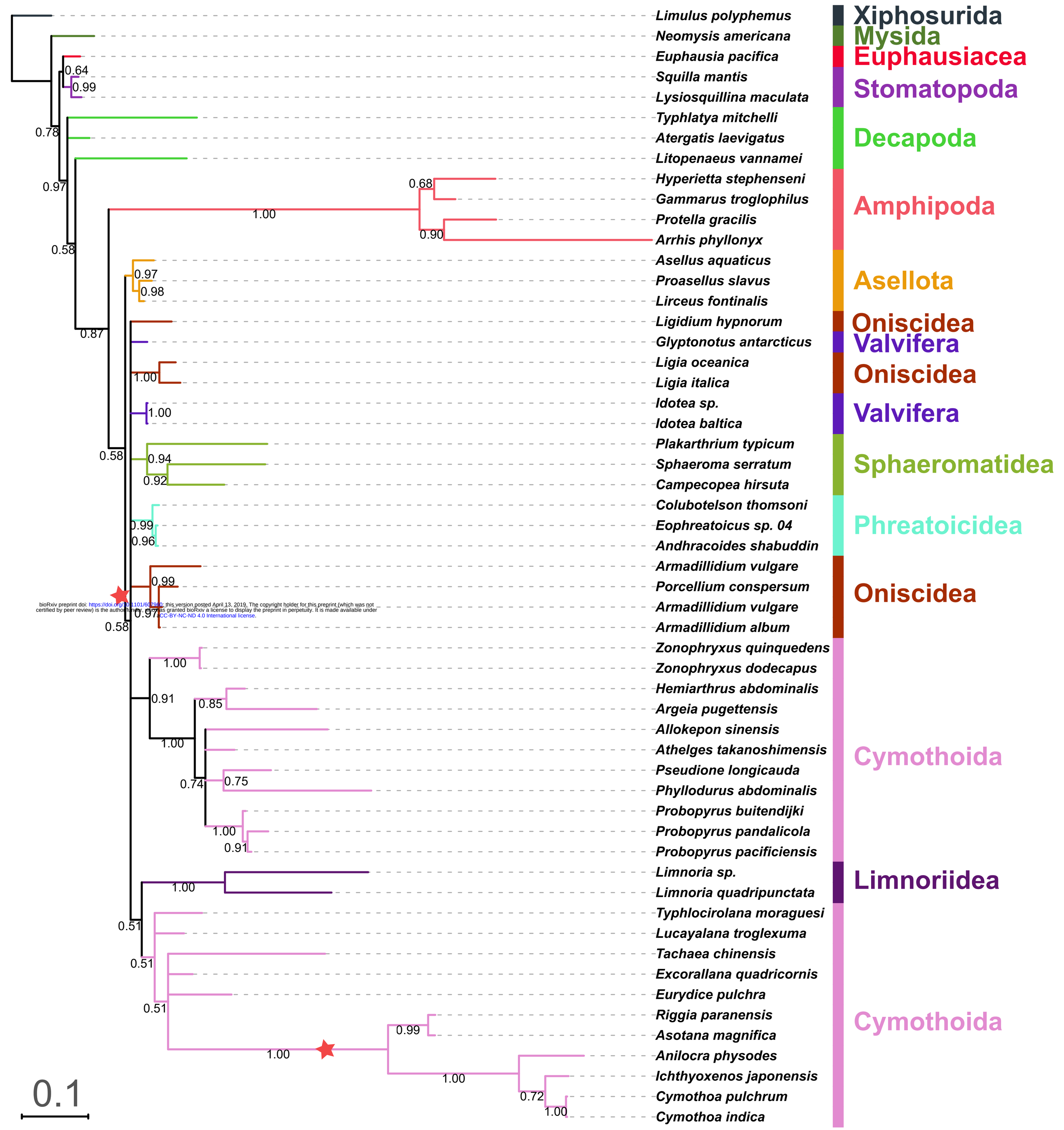
*Limulus polyphemus*
*Atergatis floridus*
*Typhlatya miravetensis*
*Penaeus vannamei*
*Euphausia pacifica*
*Lysiosquillina maculata*
*Squilla mantis*
*Neomysis japonica*
*Eulimnogammarus cyaneus*
*Metacrangonyx repens*
*Eophreatoicus sp. 14 FK-2009*
*Asellus aquaticus*
*Tachaea chinensis*
*Cymothoa indica*
*Ichthyoxenos japonensis*
*Gyge ovalis*
*Limnoria quadripunctata*
*Ligia oceanica*
*Glyptonotus cf. antarcticus FK-2009*
*Idotea baltica*
*Sphaeroma serratum*
*Bathynomus sp.*
*Eurydice pulchra*
*Cylisticus convexus*
*Trachelipus rathkii*
*Porcellio dilatatus dilatatus*
*Porcellionides pruinosus*
*Armadillidium album*
*Armadillidium nasatum*
*Armadillidium vulgare*
*Armadillidium vulgare*

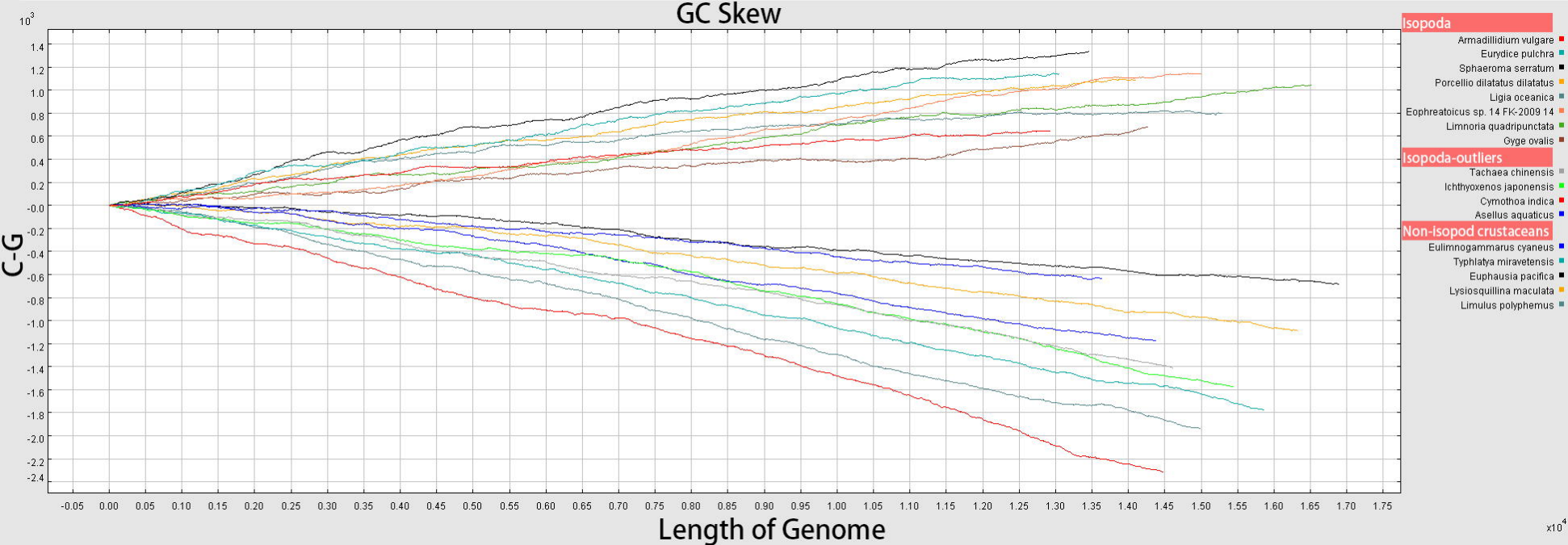
Xiphosurida
Decapoda
Euphausiacea
Stomatopoda
Mysida
Amphipoda
Phreatoicidea
Asellota
Cymothoida
Limnoriidea
Oniscidea
Valvifera
Sphaeromatidea
Cymothoida
Oniscidea

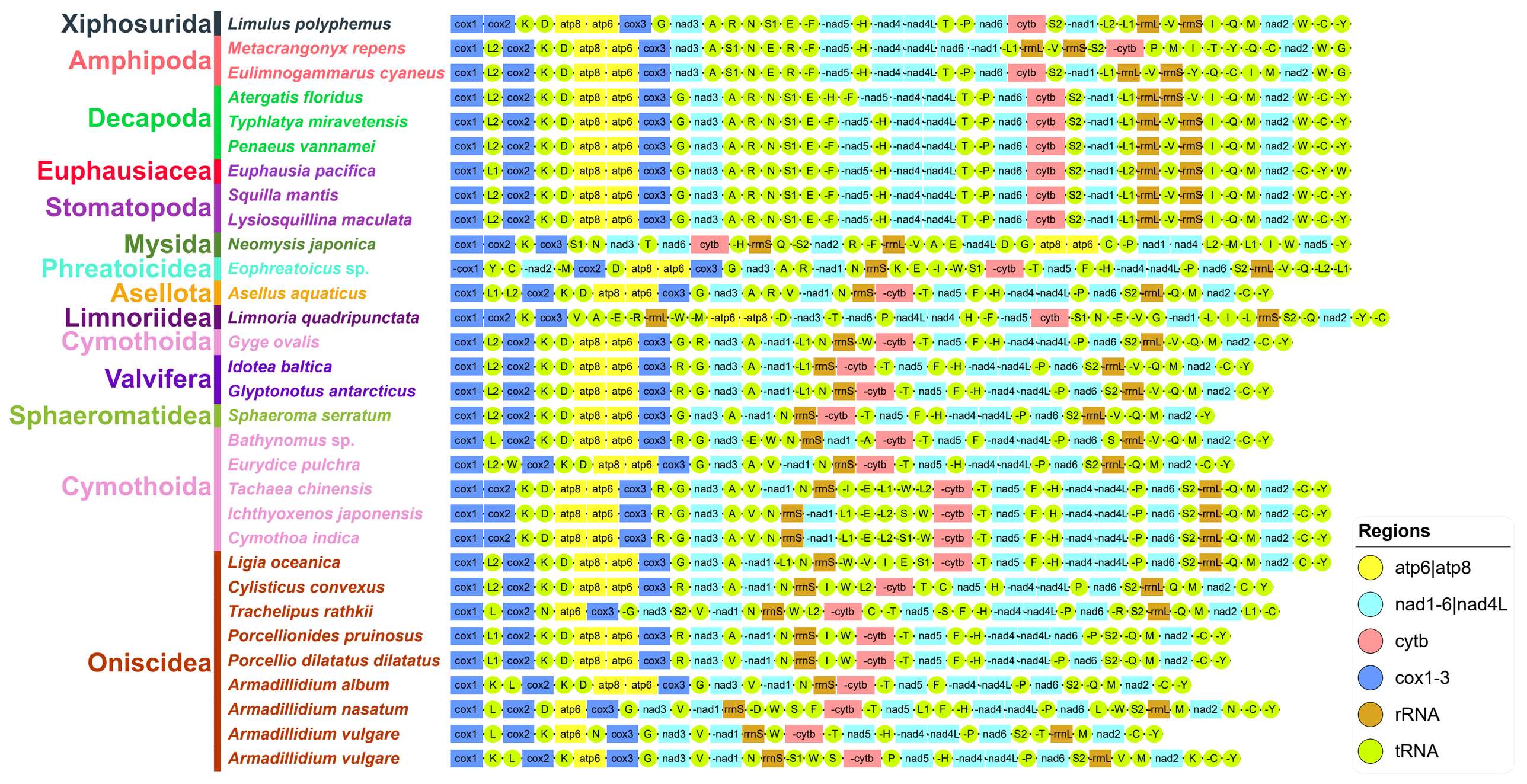
0.1








GC Skew
C-G
Length of Genome
Isopoda
Armadillidium vulgare
Eurydice pulchra
Sphaeroma serratum
Porcellio dilatatus dilatatus
Ligia oceanica
Eophreatoicus sp. 14 FK-2009 14
Limnoria quadripunctata
Gyge ovalis
Isopoda-outliers
Tachaea chinensis
Ichthyoxenos japonensis
Cymothoa indica
Asellus aquaticus
Non-isopod crustaceans
Eulimnogammarus cyaneus
Typhlatya miravetensis
Euphausia pacifica
Lysiosquillina maculata
Limulus polyphemus


Xiphosurida
Limulus polyphemus
Amphipoda
Metacrangonyx repens
Eulimnogammarus cyaneus
Decapoda
Atergatis floridus
Typhlatya miravetensis
Penaeus vannamei
Euphausiacea
Euphausia pacifica
Stomatopoda
Squilla mantis
Lysiosquillina maculata
Mysida
Neomysis japonica
Phreatoicidea
Eophreatoicus sp.
Asellota
Asellus aquaticus
Limnoriidea
Limnoria quadripunctata
Cymothoida
Gyge ovalis
Valvifera
Idotea baltica
Glyptonotus antarcticus
Sphaeromatidea
Sphaeroma serratum
Cymothoida
Bathynomus sp.
Eurydice pulchra
Tachaea chinensis
Ichthyoxenos japonensis
Cymothoa indica
Oniscidea
Ligia oceanica
Cylisticus convexus
Trachelipus rathkii
Porcellionides pruinosus
Porcellio dilatatus dilatatus
Armadillidium album
Armadillidium nasatum
Armadillidium vulgare
Armadillidium vulgare
Regions
atp6|atp8
nad1-6|nad4L
cytb
cox1-3
rRNA
tRNA