WILEY

Cladistics

Cladistics (2026) 1–10

doi: 10.1111/cla.70069

First genomic analysis of Platycopioida and phylogenomics across six copepod orders

Donggu Jeon[a], Chi-une Song[a], Junil Ko[a], Yusuke Kondo[b,†], Kenji Toyota[c], Man-Ki Jeong[d,e], Ho Young Soh[f], Susumu Ohtsuka[b,‡] and Seong-il Eyun[a,*]

[a]*Department of Life Science, Chung-Ang University, Seoul, Korea;* [b]*Setouchi Carbon Neutral Research Center, Hiroshima University, Hiroshima, Japan;* [c]*Department of Bioresource Science, Hiroshima University, Hiroshima, Japan;* [d]*Department of Smart Fisheries Resources Management, Chonnam National University, Yeosu, Korea;* [e]*ICT Fishery Resource Management Research Institute, Chonnam National University, Yeosu, Korea;* [f]*Department of Ocean Integrated Science, Chonnam National University, Yeosu, Korea*

Received 2 June 2025; Revised 31 August 2026; Accepted 4 September 2026

Abstract

Recent advances in bioinformatics have significantly expanded our understanding of the genetic properties of organisms and have provided valuable insights into their phylogenetic relationships and evolutionary histories. However, molecular phylogenetic studies of the class Copepoda remain limited, and the focus has been mainly on the major taxa, namely Calanoida, Cyclopoida, Harpacticoida and Siphonostomatoida. Consequently, a comprehensive phylogeny encompassing diverse copepod groups has yet to be established. In this study, we reconstructed the phylogeny of Copepoda using amino acid sequences of 40 orthologous genes, with particular emphasis on the previously understudied order Platycopioida. Our primary objectives were to confirm the generally recognized ancestral position of Platycopioida and to investigate subsequent phylogenetic changes in light of previous studies while incorporating additional copepod taxa. Our phylogenomic and statistical analyses revealed that Platycopioida represents the most basal lineage of Copepoda, which is consistent with previous morphological studies. In addition, we investigated the minor topological changes observed within the Podoplea clade, particularly in relation to the inclusion of additional copepod taxa and the use of more genetic information to improve the accuracy of phylogenetic inference.

Introduction

The small crustacean class Copepoda is among the most successful metazoan groups, with more than 14 500 known valid species worldwide (Walter and Boxshall, 2024). They inhabit virtually every type of aquatic environment, ranging from the deep sea (Wolff, 1959; Suárez-Morales et al., 2026) to high-elevation mountain waters (Jersabek et al., 2001), and from hypersaline (Bayly and Boxshall, 2009; Anufriieva, 2015) to freshwater ecosystems (Reid, 1986; Huys and Boxshall, 1991). This remarkable adaptability is evident in their immense morphological diversity and varied lifestyles, including planktonic, benthic and symbiotic forms. The class Copepoda is traditionally classified into 10 orders, although the validity of some orders has fluctuated over time and remains a subject of debate (Mikhailov and Ivanenko, 2021; Bernot et al., 2023). These orders were initially classified into two infraclasses, Progymnoplea and Neocopepoda, and the latter infraclass was further subdivided into two superorders, Gymnoplea and Podoplea. The Progymnoplea and Gymnoplea each contain a single order, namely Platycopioida and Calanoida, respectively, whereas the remaining eight orders belong to Podoplea.

*Corresponding author.
E-mail address: eyun@cau.ac.kr

†Present address: Hiroshima Department of Global Environmental Studies, Faculty of Environmental Studies, Hiroshima Institute of Technology, Hiroshima, Japan

‡Present address: Hiroshima Study Center, The Open University of Japan, Hiroshima, Japan.

Early Copepoda phylogenies were primarily based on morphological characteristics. However, the extreme morphological variability of copepods has become the main obstacle in phylogenetic inference, as it poses challenges in assessing morphological homology, determining character states and applying consistent weighting criteria to interpret morphological data (Kabata, 1979; Huys and Boxshall, 1991; Ho, 1994; Boxshall and Halsey, 2004; Eyun, 2017). In addition, the molecular phylogeny of copepod orders is poorly understood. Most early molecular studies relied on a single gene or a combination of several gene sequences, such as nuclear ribosomal RNAs (Huys et al., 2006, 2007; Ki et al., 2009; Takenaka et al., 2012). These limited genetic data have been used to infer relationships between species within a specific lineage or, at most, among closely related groups (Machida and Tsuda, 2010; Willett, 2012; Krajíček et al., 2016; Bradford-Grieve et al., 2017; Jeon et al., 2018), but is inadequate for reliably elucidating copepod phylogeny in large-scale analyses due to insufficient phylogenetic signal (Blanco-Bercial et al., 2011; Eyun, 2017). More recent molecular studies, in which a greater scope of genetic information has been utilized, are still limited in terms of addressing the so-called major copepod groups (i.e. Calanoida, Cyclopoida, Harpacticoida and Siphonostomatoida); each of these groups is represented by a few taxa, thus leaving the general Copepoda phylogeny somewhat incomplete (Bernot et al., 2021, 2025; Jeon et al., 2024). There is a lack of intensive molecular studies to reveal their phylogenetic positions and relationships with the rest copepod orders, in particular Platycopioida.

Platycopioida are composed of free-living copepod species, distinguished from the other groups by notable plesiomorphic features, including the gymnoplean body plan (i.e. prosome-urosome boundary lying between the fifth pedigerous and genital somites), and a maxillule with a lobate exite bearing two setae. Females also possess a single genital somite completely separate from the first abdominal somite (Huys and Boxshall, 1991; Boxshall and Halsey, 2004). Copepods in the order Platycopioida are encountered in hyperbenthic communities of relatively shallow seas or anchialine caves (Fosshagen, 1972; Fosshagen and Iliffe, 1985, 1988; Ohtsuka and Boxshall, 1994; Arbizu, 1997; Ohtsuka et al., 1998). Given their unique plesiomorphic features and ecology, Platycopioida is currently considered the most primitive order of copepods (Lang, 1948; Fosshagen and Iliffe, 1985; Huys and Boxshall, 1991). The origin of Platycopioida remains unknown; however, the distributional patterns of *Platycopia* species, which are restricted to the former Tethyan region, together with their presumed poor dispersal ability (Movies S1–S3), suggest that their ancestral lineage leading to this group might have already been present at least by the end of the Palaeozoic Era (Ohtsuka et al., 1998). Furthermore, several recent fossil-calibrated phylogenetic and phylogenomic studies have estimated ancient divergence times (408–521 mya) for Copepoda (Selden et al., 2010; Bernot et al., 2023; Jeon et al., 2024). Given that Platycopioida has traditionally been regarded as one of the most basal copepod lineages, its divergence time may also extend deep into the evolutionary history of Copepoda. However, Platycopioida has been poorly represented or simply overlooked in previous studies, leaving its evolutionary origin and divergence time unresolved.

In this study, we inferred the molecular phylogeny of the Copepoda, with a specific focus on the previously understudied order Platycopioida. The goal of this study was to reassess copepod phylogeny from a molecular perspective and explore evolutionary relationships and patterns by incorporating this ancient lineage of copepods. Although the relationships at higher taxonomic ranks aligned with previous findings, subtle topological changes were observed, particularly within Podoplea. The present results and further discussion on interpreting the phylogeny would provide valuable insights for understanding Copepoda phylogeny in general.

Materials and methods

Sample collection and next-generation sequencing analysis

Two samples containing *Platycopia orientalis* Ohtsuka & Boxshall, 1994 were collected on 21 May 2016 near Nagannu Island (Stn-7) and Ie Island (Stn-8), Japan. The Nagannu sample was obtained through two separate dredging operations, each lasting 5 min, at a depth of approximately 52 m. The operations covered the following coordinates: from 26°14′23″N, 127°32′05″ E to 26°14′26″N, 127°31′50″ E, and from 26°14′29″N, 127°31′29″ E to 26°14′33″N, 127°31′16″E. The Ie sample was collected using a sledge net towed for 20 min at depths ranging from 350 to 311 m, across coordinates from 26°40′26″N, 127°36′42″E to 26°40′57″N, 127°37′06″E. The collected samples were immediately preserved in 99% ethanol solution. In total, 17 individuals of *P. orientalis* were sorted out from these samples and next-generation sequencing (NGS) including DNA extraction and purification was conducted by DNA Link Inc. (Seoul, Korea) after sending the pooled specimen sample to their facility. Short-read whole-genome sequencing (WGS) data were generated using an Illumina NovaSeq6000 system (Illumina, San Diego, CA, USA). Paired-end libraries for WGS analysis were prepared using a TruSeq DNA Nano 350 bp kit (Illumina), and 151 bp sequencing was performed. The Illumina NGS data of *P. orientalis* have been deposited in the NCBI Sequence Read Archive (SRA) under the accession number SRR33263809.

Read filtering and de novo *genome assembly*

Raw read filtering was performed using fastp v0.23.4 (Chen et al., 2018) according to the following criteria: Trimmed reads are required to be at least 101 bp in length, contain no undetermined 'N' bases and have an average base quality Phred score of Q30 or higher. If both the reads in a pair met these conditions, they were

stored as cleaned forward and reverse data. If only one read in a pair qualified, it was stored as single-read data. Paired reads that shared sufficient overlap were merged into a single read and stored as merged-read data (Eyun et al., 2017; Jeon et al., 2023). All resulting datasets underwent read correction and contaminant removal steps through Tadpole (http://sourceforge.net/projects/bbmap) and KRAKEN2 v2.1.3 (Wood et al., 2019), and final read normalization was carried out using BBNorm (http://sourceforge.net/projects/bbmap). Genome assembly was performed using SPAdes v3.15.5 (Prjibelski et al., 2020) by applying multi *k*-mer parameters of 21-, 33-, 55- and 77-mers (Jung et al., 2020; Choi et al., 2025).

BUSCO gene mining

Single-copy orthologous genes of *P. orientalis* were identified using Benchmarking Universal Single-Copy Orthologs v5.6.1 (BUSCO) (Manni et al., 2021) analysis based on the Arthropoda lineage dataset (*arthropoda_odb10*, as of 8 January 2024; $n = 1013$). Similarly, BUSCO genes from 21 copepod taxa representing five orders (Calanoida, Cyclopoida, Siphonostomatoida, Harpacticoida and Monstrilloida) were obtained. Additionally, two malacostracan species, three thecostracan species and two insect species were analysed, with the non-crustacean insect group serving as an outgroup in the phylogenetic analysis. The BUSCO genes of taxa other than *P. orientalis* were retrieved from our legacy transcriptome and genome assemblies used in a previous study (Jeon et al., 2024). In the phylogenetic analysis, BUSCO genes were selected based on the following criteria: The gene must be present in the *P. orientalis* genome assembly and found in at least 20 of the 29 taxa analysed. Ultimately, 40 complete and single-copy BUSCO genes were assembled (Table S1).

Phylogenetic analysis

Protein sequences of the 40 BUSCO genes were individually aligned using MAFFT v7.520 (Katoh and Standley, 2013) under the L-INS-i algorithm. Poorly aligned and/or gapped regions were removed using TrimAl v1.4.rev15 (Capella-Gutiérrez et al., 2009). After trimming, all the alignments were concatenated into a single dataset (Full Matrix). The best-fit substitution model and optimal partitioning scheme were determined using IQ-TREE2 v2.1.2 (Minh et al., 2020). Phylogenetic relationships among the copepods were inferred using RAxML-NG v1.2.2 (Kozlov et al., 2019) and IQ-TREE2 for maximum-likelihood (ML) analysis, and using MrBayes v3.2.7a (Ronquist et al., 2012) and PhyloBayes-MPI v1.9 (Lartillot and Philippe, 2004) for Bayesian inference (BI). In the ML analysis using RAxML-NG, the best tree was selected from 30 random and 30 parsimony trees, and nodal support values (BP_R) were calculated using 3000 bootstrap replicates with the 'autoMRE' option, which allows the analysis to stop automatically when convergence drops below 0.03 within given the cycles. An additional ML analysis using the ultrafast bootstrapping algorithm in IQ-TREE2 was conducted by calculating the nodal support values (BP_I) with 3000 bootstrap replicates. The BI analysis of MrBayes was run for 1 042 000 generations, with sampling every 100 generations, and the first 25% of the generations were discarded for the final tree reconstruction. Another BI analysis was performed using PhyloBayes-MPI under the CAT-GTR, given that this model is considered particularly effective in addressing artefacts caused by long-branch attraction and site-specific compositional heterogeneity (Lartillot et al., 2007). Analysis was conducted using two independent chains. The final BI tree and convergence matrices were inferred using the bpcomp and tracecomp command in PhyloBayes-MPI by sampling every tree and discarding the first 350 trees as burn-in. Detailed information regarding the PhyloBayes-MPI runs is summarized in Figs. S1–S4 and Table S2. Bayesian nodal support values were provided as posterior probabilities (PP_M and PP_P, respectively). Lastly, the same ML (in IQ-TREE2) and BI (in PhyloBayes) analyses were repeated for three alternative datasets: one excluding four incomplete taxa (Matrix A), another excluding Monstrilloida (Matrix B) and one excluding Harpacticoida (Matrix C). These analyses aimed to assess whether taxa with low gene coverage or long branch leading to Monstrilloida could influence the overall tree inference. All phylogenies were visualized in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree) by using non-crustacean insect species as outgroups (see also Discussion).

Alternative tree topology test

A statistical test comparing the present copepod subtree topology (TP1) with that of our previous study (Jeon et al., 2024) (TP2), which was estimated using a similar molecular approach, was conducted using IQ-TREE2. Because the previous tree lacked Platycopioida, the present phylogenetic tree was edited to exclude this order, ensuring that the comparison focused solely on the differences within Neocopepoda (Gymnoplea + Podoplea). Two alignment datasets composed exclusively of copepod sequences were prepared from the Full Matrix generated in the present study (CopMatrix A) and the full alignment data from Jeon et al. (2024) (CopMatrix B). For statistical assessment, the Kishino-Hasegawa (KH) (Kishino and Hasegawa, 1989), Shimodaira-Hasegawa (SH) (Shimodaira and Hasegawa, 1999) and Approximately Unbiased (AU) (Shimodaira, 2002) tests were performed with 10 000 re-samplings using the RELL method, based on two protein sequence alignments, one from this study and the other from a previous study.

Results

Genome assembly using Illumina short-read WGS data

We obtained 63.5 Gbp of WGS short-read data (paired-end; 151 × 2) from a pooled sample containing 17 *P. orientalis* individuals. The subsequent read cleaning procedure resulted in 38.5 Gbp of genomic data, consisting of 35.5 Gbp of paired-end, 349.0 Mbp of unpaired and 2.6 Gbp of merged data. Genomic data were assembled in 2.04 Gbp with 1.60 mega scaffolds. The BUSCO assessment resulted in a complete score of 87.7% for both complete single-copy and duplicate genes, with 9.6% fragmented and 2.7% missing genes. The quality of the assembly was generally poor, especially in terms of highly fragmented scaffolds; however, we could still extract sufficient single-copy orthologous genes from the BUSCO analysis for downstream phylogenetic inference.

Phylogenetic relationships among Copepoda orders

The protein sequences of 40 single-copy orthologous BUSCO genes were aligned, resulting in a concatenated alignment of 24 012 amino acid (AA) positions. All four tree inference methods yielded nearly consistent topologies among orders, including those outside Copepoda, with high nodal support (Fig. 1). Within Copepoda, the order Platycopioida has always emerged as the earliest branching group,

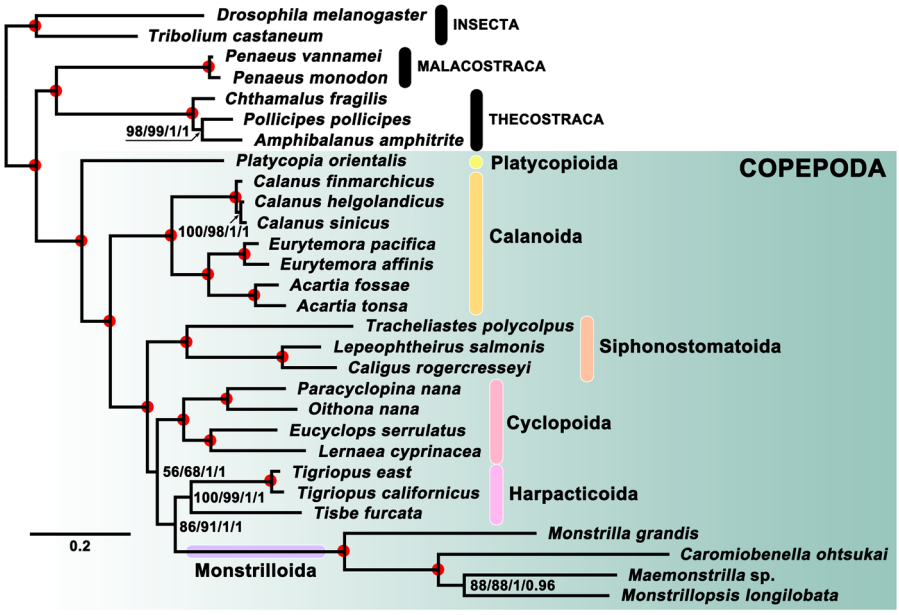

Fig. 1. Phylogenetic tree of 22 copepods and seven other arthropods inferred based on 40 single-copy orthologous BUSCO genes (i.e. Full Matrix), using RAxML-NG, IQ-TREE2, MrBayes and PhyloBayes. Nodal support values are provided in the following order: BP_R, BP_I, PP_M, PP_P. Red dots at nodes indicate maximum support in all analyses. Three different arthropod taxa (Insecta, Malacostraca and Thecostraca) were analysed, and the final tree was rooted using two non-crustacean insect species, fruit fly (*Drosophila melanogaster*) and red flour beetle (*Tribolium castaneum*), as outgroups. Scale bar represents amino acid substitutions per site.

and within the Neocopepoda, Calanoida diverged from the other four orders (Cyclopoida, Siphonostomatoida, Harpacticoida and Monstrilloida), forming distinct clades: Gymnoplea for Calanoida and Podoplea for the remaining groups. Within Podoplea, the four copepod orders branched sequentially in the order of Siphonostomatoida, Cyclopoida, Harpacticoida and Monstrilloida in the full tree inferred using all the data acquired here. Three additional analyses revealed similarly consistent topologies, with the exception of those based on the dataset excluded Harpacticoida (see Discussion). In general, the evolutionary patterns at higher taxonomic ranks were consistent with previous findings (Lozano-Fernandez et al., 2019; Bernot et al., 2021, 2023, 2025).

Alternative tree topology test

Despite the overall similarity of the present phylogeny to previous studies, the relationships among the podoplean groups still differ. Alternative tree topology tests were conducted to address this issue. A test based on the CopMatrix A revealed little support for the grouping of Siphonostomatoida and Cyclopoida (Jeon et al., 2024), marginally rejecting this hypothesis at the 0.05 significance level (Table 1). Conversely, another test based on the previous alignment of 25 nuclear protein-coding gene sequences (CopMatrix B) (Jeon et al., 2024) showed more support for the earlier Copepoda tree topology but also showed possible support for the present topology, with *P*-values exceeding 0.15 in all test cases (Table 1).

Discussion

This study is the first in which the phylogeny of Copepoda has been reaffirmed through a phylogenomic analysis with a particular focus on Platycopioida. Our findings strongly support the previously recognized copepod clades, confirming the validity of the two infraclasses (Progymnoplea and Neocopepoda) as well as the subdivisions into Gymnoplea and Podoplea within Neocopepoda (Huys and Boxshall, 1991; Ho, 1994; Ho et al., 2003). Consistent with other studies, Platycopioida was placed on the most basal branch of Copepoda, thus representing Progymnoplea. In addition, we recovered the close affinity between Harpacticoida and Monstrilloida, with a long branch leading to Monstrilloida (Jeon et al., 2024).

However, the results of the present study also revealed different relationships among the Podoplea lineages. In previous studies (Eyun, 2017; Jeon et al., 2024) in which similar molecular methods were followed, the results suggested a grouping or at the very least, a close affinity between Cyclopoida and Siphonostomatoida, prompting the authors to further propose a sister relationship to Harpacticoida (or to the Harpacticoida-Monstrilloida clade). By contrast, the results we report here indicate that

Table 1
Statistical result of alternative tree topology tests

Alignment dataset (reference)	Tree topology*	logL	ΔL	*p*-KH	*p*-SH	*p*-AU
CopMatrix A (this study)	TP1	−433 664.556	Best	0.956 (+)	1.000 (+)	0.954 (+)
	TP2	−433 711.167	46.611	0.044 (−)	0.044 (−)	0.046 (−)
CopMatrix B (Jeon et al., 2024)	TP1	−266 663.942	24.642	0.152 (+)	0.152 (+)	0.150 (+)
	TP2	−266 639.300	Best	0.848 (+)	1.000 (+)	0.850 (+)

Δ*L*, log*L* difference from the maximal log*L*. *p*-KH, *P*-value of one-sided Kishino–Hasegawa test (Kishino and Hasegawa, 1989). *p*-SH, *P*-value of Shimodaira–Hasegawa test (Shimodaira and Hasegawa, 1999). *p*-AU, *P*-value of approximately unbiased test (Shimodaira, 2002). Positive (+) signs denote the 95% confidence sets, whereas negative (−) signs indicate significant exclusion.

*Topology examined: TP1, (CAL,(SIP,(CYC,(HAR,MON)))), this study; TP2, (CAL,((SIP,CYC),(HAR,MON))), from Jeon et al. (2024); CAL = Calanoida; SIP = Siphonostomatoida; CYC = Cyclopoida; HAR = Harpacticoida; MON = Monstrilloida.

Siphonostomatoida is the earliest diverging branch, forming a sister relationship with all other podoplean taxa (Fig. 1). Despite the generally high nodal support for our phylogeny, placing Siphonostomatoida as the most basal branch of Podoplea is unusual based on current knowledge, particularly when compared to earlier morphology-based assessments (Ho, 1990, 1994; Huys and Boxshall, 1991; Ho et al., 2003) and other molecular analyses (Tung et al., 2014; Eyun, 2017; Hwang et al., 2019; Jeon et al., 2024).

To address the cause of this topological discrepancy, alternative tree topology tests were performed. Statistical comparisons between the present and previous podoplean subtree topologies TP1 and TP2 using two alignment datasets CopMatrix A and CopMatrix B indicated that the present CopMatrix A alignment has a very low probability of supporting the previous TP2 topology, which depicted groupings of Siphonostomatoida-Cyclopoida and Harpacticoida-Monstrilloida. This test marginally rejects the hypothesis at a significance level of 0.05. Conversely, another result suggested that the CopMatrix B alignment occasionally reproduces the present TP1 topology (Table 1). Securing sufficient genetic data is arguably one of the most critical factors for successful molecular phylogenetic analysis (Rosenberg and Kumar, 2001, 2003; Eyun, 2017). In this context, the present CopMatrix A, comprising longer sequences from 40 single-copy orthologous genes (with 18 794 patterns out of 24 012 alignment sites), is more likely to provide a robust phylogenetic signal for the Copepoda phylogeny than is the previous CopMatrix B (11 162 out of 16 976). This aligns with other recent phylogenomic studies (Lozano-Fernandez et al., 2019; Bernot et al., 2023, 2025) and synthetic tree reconstructions (Bernot et al., 2021) in which similar phylogenetic relationships among the given copepod orders were proposed. These studies, in turn, have generally been conducted using larger amounts of sequence data compared to the other studies mentioned previously. Examples are up to 939 460 (Bernot et al., 2025), 121 508 (Bernot et al., 2023) and 57 149 AAs (Lozano-Fernandez et al., 2019) vs. partial 18S rRNA sequence (Tung et al., 2014), 11 162 AAs (Jeon et al., 2024) and 13 mitochondrial protein-coding gene sequences (Hwang et al., 2019). These results underscore the benefit of using more genetic data to improve phylogenetic resolution. Although the overall increase in data may be relatively minor, the present analysis focused on the Copepoda phylogeny with fewer taxa outside of Copepoda, meaning that even such 'small increase' would have a greater impact on the analysis.

One notable difference between the CopMatrix A and CopMatrix B datasets was the completeness of the gene mining. Specifically, the CopMatrix B provided more extensive gene coverage per species than did the present CopMatrix A. This disparity was particularly evident in several species analysed here, which exhibited gene coverage below 50% (out of a total of 40 genes): *Calanus finmarchicus* (19 genes; Calanoida), *Eucyclops serrulatus* (19 genes; Cyclopoida), *Lernaea cyprinacea* (17 genes; Cyclopoida) and *Monstrilla grandis* (16 genes; Monstrilloida) (Table S1). To investigate the potential effects of these incomplete taxa, we conducted additional phylogenetic analyses using the Matrix A that excluded the four taxa with highly incomplete gene representations by employing IQ-TREE2 (for ML analysis) and PhyloBayes (for BI analysis). The resulting topologies were nearly identical to that of full tree (Fig. 2a), although the BI tree exhibited slightly different relationships between the three monstrilloid species (inset in Fig. 2a). Interestingly, Monstrilloida as a group maintained a close phylogenetic position to Harpacticoida even after the exclusion of *M. grandis*, which appears to be the most basal monstrilloid taxon (Jeon et al., 2024). This finding suggests that the removal of incomplete taxa does not substantially affect the overall ordinal relationships but may introduce some phylogenetic ambiguities at lower taxonomic levels. We suspect that the absence of *M. grandis*, whose phylogenetic position bridges Harpacticoida to other distantly related monstrilloids, likely resulted in an erroneous subtree for Monstrilloida by increasing branch length (increased from 0.34 AA substitutions per site to 0.51 in the ML analysis,

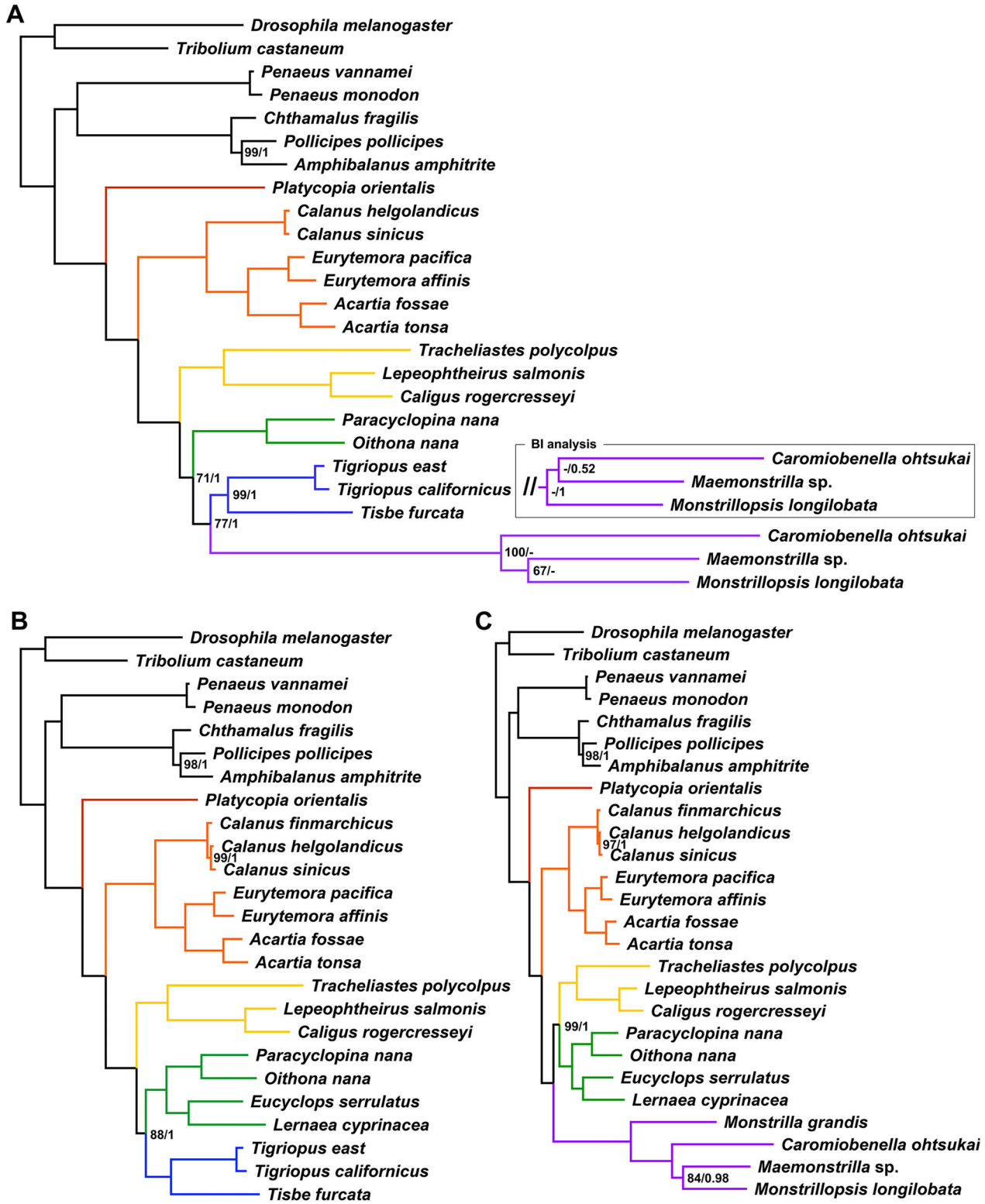

Fig. 2. Phylogenetic trees reconstructed based on (a) alignment excluding incomplete taxa (Matrix A), (b) alignment excluding Monstrilloida taxa (Matrix B) and (c) alignment excluding Harpacticoida taxa (Matrix C). The inset accompanying panel a depicts a different monstrilloid subtree resulting from the BI analysis. Each tree was inferred using IQ-TREE2 and PhyloBayes. Nodal support values are provided in the order of BP_I, PP_P. Nodes without support values indicate those with maximum support in both analyses. Branch colour indexes for copepod orders: red, Platycopioida; orange, Calanoida; yellow, Siphonostomatoida; green, Cyclopoida; blue, Harpacticoida; and purple, Monstrilloida.

and from 0.78 to 1.18 in the BI analysis), and thus exacerbating the long-branch attraction effect. Changes in nodal support values further reflect this issue: a significant decrease from 91% to 77% in BP_I for the split of Harpacticoida and Monstrilloida, and from 100% to 67% in BP_I for the split of *Maemonstrilla* sp. and *Monstrillopsis longilobata* Lee, Kim & Chang, 2016. In addition, a decrease in PP_P from 1.00 to 0.52 at the node of *Caromiobenella ohtsukai* Jeon, Lee & Soh, 2018 under a different topological pattern in the ML analysis further reflects the instability of the Monstrilloida subtree. Nevertheless, these observations support the idea that including taxa with incomplete gene coverage, even if the obtainable genetic data are relatively small, can improve the overall phylogenetic accuracy (Wiens and Tiu, 2012).

In addition to the aforementioned issue, the long branch leading to Monstrilloida may also have influenced the overall accuracy of the present phylogeny. To clarify this, we examined further topological changes by conducting ML and BI analyses on two datasets: one excluding Monstrilloida (Matrix B) and the other excluding Harpacticoida (Matrix C) instead. We anticipated that comparing these results would offer a clearer understanding of the possible long-branch attraction in the present phylogeny (Siddall and Whiting, 1999). The phylogeny derived from the Matrix B was identical to the full tree topology (Fig. 2b), whereas that derived from the Matrix C placed Monstrilloida at the most basal position within Podoplea (Fig. 2c). This contrast can be interpreted as evidence that the current phylogeny is being affected by the long-branch attraction involving the Monstrilloida clade and further emphasizes the importance of taxon sampling (Rannala et al., 1998; Heath et al., 2008). However, in reality, obtaining data from all taxa of interest is challenging, and this is even more pronounced for non-model organisms, such as Copepoda, owing to the general lack of available data in public databases (e.g. GenBank). Consequently, sequence mining from limited Copepoda sources can results in taxonomically biased datasets. This limitation seems to apply to the present analysis to some extent because each order was represented by a small number of taxa. Nonetheless, this study holds promise in that it describes other stable ordinal-level relationships, especially for Platycopioida. Further discrepancies from previous studies were observed in the pancrustacean-level phylogenetic comparisons. Previous studies have proposed several different topologies, including (Hexapoda, (Malacostraca, (Thecostraca, Copepoda))) (Oakley et al., 2013; Lozano-Fernandez et al., 2019), (Hexapoda, (Thecostraca, (Malacostraca, Copepoda))) (Schwentner et al., 2017), and ((Hexapoda, Copepoda), (Thecostraca, Malacostraca)) (Bernot et al., 2023). Although our analysis recovered the topology (Hexapoda, ((Malacostraca, Thecostraca), Copepoda)), consistent with that reported by Schwentner et al. (2018), the alternative topologies mentioned above cannot be definitively supported or rejected based on the present analysis, owing to the limited number of non-copepod taxa included in our dataset: Malacostraca was represented by two closely related species (*Penaeus monodon* and *P. vannamei*), Thecostraca by three species all belonging to the subclass Cirripedia and Hexapoda —one of the most diverse animal groups—by only two species. This limited representation may have led to altered relationships, similar to the case of copepod phylogeny discussed above. Despite such ambiguity, Hexapoda consistently formed a distinct clade, which led us to designate them as an outgroup in this study. While further analytical attempts or interpretations to reveal more accurate phylogenetic relationships may be beyond the scope of this work, we emphasize that a broader and denser taxon sampling is as important as obtaining sufficient genetic data in reconstructing reliable phylogenetic trees.

One potential criticism of the present justification is that the current phylogenetic evaluation relies heavily on comparisons with the results (i.e. CopMatrix B and TP2) of our previous study (Jeon et al., 2024). This reliance arises because other phylogenomic studies mostly deal with only a few copepod orders (typically one or two orders in addition to Calanoida) in large pancrustacean-level phylogenetic inferences (Minxiao et al., 2011; Oakley et al., 2013; Schwentner et al., 2018). Consequently, these copepod subtrees do not provide sufficient detail for a more thorough comparison of podoplean phylogeny. In other studies that were limited by insufficient taxon sampling and/or unbalanced sampling of copepods, the authors have often suggested completely different evolutionary hypotheses, portraying Calanoida (representative of Gymnoplea) as a more derived clade than other podopleans (Braga et al., 1999; Schwentner et al., 2017). These conflicting findings leave open further questions regarding the incongruence and uncertainty in copepod phylogeny.

Furthermore, other podoplean copepod orders, such as Gelyelloida, Misophrioida and Mormonilloida, remain unanalyzed. This implies that the true phylogenetic relationships in Copepoda may differ significantly from the present interpretation. Fortunately, recent studies have made continued efforts to include taxa that were overlooked in the past (e.g. Jeon et al., 2024; Bernot et al., 2025), providing new insights into the broader copepod phylogeny. Moving forward, increased taxon sampling, not only in terms of quantity but also with a balanced representation of taxa, will be essential, along with continued genome sequencing efforts for unexplored copepod orders. Although the current lack of genetic information has hindered the establishment of a

robust Copepoda phylogeny, the results reported here and necessary future work will bring us closer to understanding their true relationships and evolution, as well as advance the overall systematics of the group by addressing all the caveats discussed in this study.

Funding

This study was supported by the Basic Science Research Program funded by the Ministry of Education (RS-2023-00243296). This work was also supported by the Korea Institute of Marine Science & Technology Promotion (RS-2025-02215227, RS-2026-25543836 and RS-2026-25525656) funded by the Ministry of Oceans and Fisheries.

Conflict of interest

The authors declare that they have no competing interest related to this study.

Acknowledgements

We thank all members of the laboratory for their insightful comments and discussion.

Data availability

All data used in this study are available at https://eyunlab.cau.ac.kr/platycopioida. Raw sequencing data are deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession PRJNA1253919 (BioSample accession: SAMN48090785; SRA accession: SRR33263809).

References

Anufriieva, E.V., 2015. Do copepods inhabit hypersaline waters worldwide? A short review and discussion. Chinese J. Oceanol. Limnol. 33, 1354–1361.

Arbizu, P.M., 1997. *Sarsicopia polaris* gen. et sp. n., the first Platycopioida (Copepoda: Crustacea) from the Arctic Ocean, and its phylogenetic significance. Hydrobiologia 350, 35–47.

Bayly, I.A.E. and Boxshall, G.A., 2009. An all-conquering ecological journey: from the sea, calanoid copepods mastered brackish, fresh, and athalassic saline waters. Hydrobiologia 630, 39–47.

Bernot, J.P., Boxshall, G.A. and Crandall, K.A., 2021. A synthesis tree of the Copepoda: integrating phylogenetic and taxonomic data reveals multiple origins of parasitism. Peer J 9, e12034.

Bernot, J.P., Owen, C.L., Wolfe, J.M., Meland, K., Olesen, J. and Crandall, K.A., 2023. Major revisions in pancrustacean phylogeny and evidence of sensitivity to taxon sampling. Mol. Biol. Evol. 40, msad175.

Bernot, J.P., Khodami, S., Boyen, J., de Troch, M., Boxshall, G.A. and Martinez Arbizu, P., 2025. Copepod phylogenomics supports Canuelloida as a valid order separate from Harpacticoida. Mol. Phylogenet. Evol. 206, 108311.

Blanco-Bercial, L., Bradford-Grieve, J. and Bucklin, A., 2011. Molecular phylogeny of the Calanoida (Crustacea: Copepoda). Mol. Phylogenet. Evol. 59, 103–113.

Boxshall, G. and Halsey, S.H., 2004. An introduction to copepod diversity. The Ray Society, London.

Bradford-Grieve, J.M., Blanco-Bercial, L. and Boxshall, G.A., 2017. Revision of family Megacalanidae (Copepoda: Calanoida). Zootaxa 4229, 1–183.

Braga, E., Zardoya, R., Meyer, A. and Yen, J., 1999. Mitochondrial and nuclear rRNA based copepod phylogeny with emphasis on the Euchaetidae (Calanoida). Mar. Biol. 133, 79–90.

Capella-Gutiérrez, S., Silla-Martínez, J.M. and Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Chen, S., Zhou, Y., Chen, Y. and Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890.

Choi, H., An, Y.K., Lee, C.-J., Song, C.-u., Kim, E.-J., Lee, C.E., Cho, S.-J. and Eyun, S., 2025. Genome assembly, gene content, and plastic gene expression responses to salinity changes in the brackishwater clam (*Corbicula japonica*) from a dynamic estuarine environment. J. Hazard. Mater. 483, 136627.

Eyun, S., 2017. Phylogenomic analysis of Copepoda (Arthropda, Crustacea) reveals unexpected similarities with earlier proposed morphological phylogenies. BMC Evol. Biol. 17, 23.

Eyun, S., Soh, H.Y., Posavi, M., Mounro, J.B., Hughes, D.S.T., Murali, S.C., Qu, J., Dugan, S., Lee, S.L., Chao, H., Dinh, H., Han, Y., Doddapaneni, H.V., Worley, K.C., Muzny, D.M., Park, E.-O., Silva, J.C., Gibbs, R.A., Richards, S. and Lee, C.E., 2017. Evolutionary history of chemosensory-related gene families across the Arthropoda. Mol. Biol. Evol. 34, 1838–1862.

Fosshagen, A., 1972. Marine biological investigations in the Bahamas: 17. Platycopiidae (Copepoda, Calanoida) from the Bahamas. Sarsia 48, 51–60.

Fosshagen, A. and Iliffe, T.M., 1985. Two new genera of Calanoida and a new order of Copepoda, Platycopioida, from marine caves on Bermuda. Sarsia 70, 345–358.

Fosshagen, A. and Iliffe, T.M., 1988. A new genus of Platycopioida (Copepoda) from a marine cave on Bermuda. Hydrobiologia 167/168, 357–361.

Heath, T., Hedtke, S.M. and Hillis, D.M., 2008. Taxon sampling and the accuracy of phylogenetic analyses. J. Syst. Evol. 46, 239–257.

Ho, J.-s., 1990. Phylogenetic analysis of copepod orders. J. Crustac. Biol. 10, 528–536.

Ho, J.-s., 1994. Copepod phylogeny: a reconsideration of Huys & Boxshall's 'parsimony versus homology'. Hydrobiologia 292/293, 31–39.

Ho, J.-s., Dojiri, M., Hendler, G. and Deets, G.B., 2003. A new species of Copepoda (Thaumatopsyllidae) symbiotic with a brittle star from California, U.S.A., and designation of a new order Thaumatopsylloida. J. Crustac. Biol. 23, 582–594.

Huys, R. and Boxshall, G., 1991. Copepod evolution. The Ray Society, London.

Huys, R., Llewellyn-Hughes, J., Olson, P.D. and Nagasawa, K., 2006. Small subunit rDNA and Bayesian inference reveal *Pectenophilus ornatus* (Copepoda *incertae sedis*) as highly transformed Mytilicolidae, and support assignment of Chondracanthidae and Xarifiidae to Lichomolgoidea (Cyclopoida). Biol. J. Linn. Soc. 87, 403–425.

Huys, R., Llewellyn-Hughes, J., Conroy-Dalton, S., Olson, P.D., Spinks, J.N. and Johnston, D.A., 2007. Extraordinary host switching in siphonostomatoid copepods and the demise of the Monstrilloida: Integrating molecular data, ontogeny and antennulary morphology. Mol. Phylogenet. Evol. 43, 368–378.

Hwang, D.-S., Choi, B.-S., Lee, M.-C., Han, J., Kim, S. and Lee, J.-s., 2019. Complete mitochondrial genome of the Antarctic

copepod *Tigriopus kingsejongenesis* (Harpacticoida, Harpacticidae). Mitochondrial DNA Part B 4, 1470–1471.
Jeon, D., Lee, W. and Soh, H.Y., 2018. A new genus and two new species of monstrilloid copepods (Copepoda: Monstrillidae): integrating morphological, molecular phylogenetic, and ecological evidence. J. Crustac. Biol. 38, 45–65.
Jeon, M.-S., Jeong, D.M., Doh, H., Kang, H.A., Jung, H. and Eyun, S., 2023. A practical comparison of the next-generation sequencing platform and assemblers using yeast genome. Life Sci. Alliance 6, e202201744.
Jeon, D., Song, C.-u., Jeong, H.G., Ohtsuka, S., Lee, W., Soh, H.Y. and Eyun, S., 2024. An integrated phylogenomic approach for potential host-associated evolution of monstrilloid copepods. Oceanogr. Mar. Biol. 62, 350–375.
Jersabek, C.D., Brancelj, A., Stoch, F. and Schabetsberger, R., 2001. Distribution and ecology of copepods in mountainous regions of the Eastern Alps. Hydrobiologia 453/454, 309–324.
Jung, H., Ventura, T., Chung, J.S., Kim, W.-J., Nam, B.-H., Kong, H.J., Kim, Y.-O., Jeon, M.-S. and Eyun, S.-i., 2020. Twelve quick steps for genome assembly and annotation in the classroom. PLoS Comput. Biol. 16, e1008325.
Kabata, Z., 1979. Parasitic Copepoda of British fishes. The Ray Society, London.
Katoh, K. and Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Ki, J.-s., Lee, K.-w., Park, H.G., Chullasorn, S., Dahms, H.-U. and Lee, J.-s., 2009. Phylogeography of the copepod *Tigriopus japonicus* along the Northwest Pacific rim. J. Plankton Res. 31, 209–221.
Kishino, H. and Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J. Mol. Evol. 29, 170–179.
Kozlov, A.M., Darriba, D., Flouri, T., Morel, B. and Stamatakis, A., 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35, 4453–4455.
Krajíček, M., Fott, J., Miracle, M.R., Ventura, M., Sommaruga, R., Kirschner, P. and Černý, M., 2016. The genus *Cyclops* (Copepoda, Cyclopoida) in Europe. Zool. Scr. 45, 671–682.
Lang, K., 1948. Copepoda 'Notodelphyoida' from the Swedish west-coast with an outline on the systematics of the copepods. Ark. Zool. 40A, 1–36.
Lartillot, N. and Philippe, H., 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. 21, 1095–1109.
Lartillot, N., Brinkmann, H. and Philippe, H., 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol. Biol. 7, S4.
Lozano-Fernandez, J., Giacomelli, M., Fleming, J.F., Chen, A., Vinther, J., Thomsen, P.F., Glenner, H., Palero, F., Legg, D.A., Iliffe, T.M., Pisani, D. and Olesen, J., 2019. Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling. Genome Biol. Evol. 11, 2055–2070.
Machida, R.J. and Tsuda, A., 2010. Dissimilarity of species and forms of planktonic *Neocalanus* copepods using mitochondrial *COI*, *12S*, nuclear *ITS*, and *28S* gene sequences. PLoS One 5, e10278.
Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A. and Zdobnov, E.M., 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol. Biol. Evol. 38, 4647–4654.
Mikhailov, K.V. and Ivanenko, V.N., 2021. Low support values and lack of reproducibility of molecular phylogenetic analysis of Copepoda orders. Arthropoda Sel. 30, 39–42.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. and Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534.
Minxiao, W., Song, S., Chaolun, L. and Xin, S., 2011. Distinctive mitochondrial genome of calanoid copepod *Calanus sinicus* with multiple large non-coding regions and reshuffled gene order: useful molecular markers for phylogenetic and population studies. BMC Genomics 12, 73.
Oakley, T.H., Wolfe, J.M., Lindgren, A.R. and Zaharoff, A.K., 2013. Phylotranscriptomics to bring the understudied into the fold: monophyletic Ostracoda, fossil placement, and pancrustacean phylogeny. Mol. Biol. Evol. 30, 215–233.
Ohtsuka, S. and Boxshall, G.A., 1994. *Platycopia orientalis*, new species (Copepoda: Platycopioida), from the North Pacific, with descriptions of copepodid stages. J. Crustac. Biol. 14, 151–167.
Ohtsuka, S., Soh, H.Y. and Ueda, H., 1998. *Platycopia compacta* n. sp., the second species of Platycopioida (Crustacea: Copepoda) in the Indo-Pacific region, with remarks on development, feeding, swimming, and zoogeography. Zoolog. Sci. 15, 415–424.
Prjibelski, A.D., Antipov, D., Meleshko, D., Lapidus, A. and Korobeynikov, A., 2020. Using SPAdes *de novo* assembler. Curr. Protoc. 70, e102.
Rannala, B., Huelsenbeck, J.P., Yang, Z. and Nielsen, R., 1998. Taxon sampling and the accuracy of large phylogenies. Syst. Biol. 47, 702–710.
Reid, J.W., 1986. Some usually overlooked cryptic copepod habitats. Syllogeous 58, 594–598.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. and Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Rosenberg, M.S. and Kumar, S., 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. Proc. Natl. Acad. Sci. U. S. A. 98, 10751–10756.
Rosenberg, M.S. and Kumar, S., 2003. Taxon sampling, Bioinformatics, and phylogenomics. Syst. Biol. 52, 119–124.
Schwentner, M., Combosch, D.J., Nelson, J.P. and Giribet, G., 2017. A phylogenomic solution to the origin of insects by resolving crustacean-hexapod relationships. Curr. Biol. 27, 1818–1824.e1815.
Schwentner, M., Richter, S., Rogers, D.C. and Giribet, G., 2018. Tetraconatan phylogeny with special focus on Malacostraca and Branchiopoda: highlighting the strength of taxon-specific matrices in phylogenomics. Proc. R. Soc. B Biol. Sci. 285, 20181524.
Selden, P.A., Huys, R., Stephenson, M.H., Heward, A.P. and Taylor, P.N., 2010. Crustaceans from bitumen clast in Carboniferous glacial diamictite extend fossil record of copepods. Nat. Commun. 1, 50.
Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.
Shimodaira, H. and Hasegawa, M., 1999. Multiple comparisons of log-Likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.
Siddall, M.E. and Whiting, M.F., 1999. Long-branch abstractions. Cladistics 15, 9–24.
Suárez-Morales, E., Martínez, A., Martinez Arbizu, P., Khodami, S. and Mercado-Salas, N.F., 2026. A new family of the order Monstrilloida (Copepoda) from deep waters of the North Atlantic supported by morphological and genetic evidence. PeerJ 14, e21176.
Takenaka, Y., Yamaguchi, A., Tsuruoka, N., Torimura, M., Gojobori, T. and Shigeri, Y., 2012. Evolution of bioluminescence in marine planktonic copepods. Mol. Biol. Evol. 29, 1669–1681.
Tung, C.-H., Cheng, Y.-r., Lin, C.-y., Ho, J.-s., Kuo, C.-h., Yu, J.-k. and Su, Y.-H., 2014. A new copepod with transformed body plan and unique phylogenetic position parasitic in the acorn worm *Ptychodera flava*. Biol. Bull. 226, 69–80.
Walter, T.C. and Boxshall, G., 2024. World of Copepods database. Available at: http://www.marinespecies.org/copepoda (accessed December 11, 2024).
Wiens, J.J. and Tiu, J., 2012. Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling. PLoS One 7, e42925.
Willett, C.S., 2012. Quantifying the elevation of mitochondrial DNA evolutionary substitution rates over nuclear rates in the intertidal copepod *Tigriopus californicus*. J. Mol. Evol. 74, 310–318.

Wolff, T., 1959. The hadal community, an introduction. Deep-Sea Res. 6, 95–124.
Wood, D.E., Lu, J. and Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 257.

Supporting Information

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

Fig. S1. MCMC trace plots for PhyloBayes-MPI analyses of full dataset. Trace plots for continuous parameters (logL, statent, length, alpha, Nmode and rrent) comparing two independent MCMC chains for the full dataset (~3810 cycles). All chains rapidly reached stationarity within the first 100 iterations and showed tight overlap throughout the remaining cycles.

Fig. S2. MCMC trace plots for PhyloBayes-MPI analyses of dataset excluding incomplete taxa. Trace plots for continuous parameters (logL, statent, length, alpha, Nmode and rrent) comparing two independent MCMC chains for the dataset excluding incomplete taxa (~5641 cycles). All chains rapidly reached stationarity within the first 100 iterations and showed tight overlap throughout the remaining cycles.

Fig. S3. MCMC trace plots for PhyloBayes-MPI analyses of dataset excluding monstrilloid taxa. Trace plots for continuous parameters (logL, statent, length, alpha, Nmode and rrent) comparing two independent MCMC chains for the dataset excluding monstrilloid taxa (~5433 cycles). All chains rapidly reached stationarity within the first 100 iterations and showed tight overlap throughout the remaining cycles.

Fig. S4. MCMC trace plots for PhyloBayes-MPI analyses of dataset excluding harpacticoid taxa. Trace plots for continuous parameters (logL, statent, length, alpha, Nmode and rrent) comparing two independent MCMC chains for the dataset excluding harpacticoid taxa (~3867 cycles). All chains rapidly reached stationarity within the first 100 iterations and showed tight overlap throughout the remaining cycles.

Movie S1. The movie of swimming *Platycopia orientalis*. https://youtu.be/NXRD6ihGxFo. These movies are available at: https://eyunlab.cau.ac.kr/platycopioida.

Movie S2. The movie of swimming *Platycopia orientalis*. https://youtu.be/4bIXOkQZrG4. These movies are available at: https://eyunlab.cau.ac.kr/platycopioida.

Movie S3. The movie of swimming platycopioids. https://youtu.be/qLBIOj33Mac. These movies are available at: https://eyunlab.cau.ac.kr/platycopioida.

Table S1. The forty complete single-copy BUSCO genes used for this study.

Table S2. Information on PhyloBayes analysis for different datasets.