Institute of Parasitology, Biology Centre CAS
Folia Parasitologica 2024, 71: 002
doi: 10.14411/fp.2024.002

Folia Parasitologica
http://folia.paru.cas.cz

Research Article OPEN ACCESS

Comprehensive molecular characterisation of the complete mitogenome of *Ergasilus tumidus* and phylogenetic relationships of Copepoda inferred from mitogenomes

Yan Huang[1], **Jin-Mei Feng**[2], **Wei Liu**[1], **Bin-Lian Sun**[1], **Xi-Ji Shu**[1], **Wen-Xiang Li**[3], **Gui-Tang Wang**[3] and **Cong-Jie Hua**[1,2]

[1] Hubei Key Laboratory of Cognitive and Affective Disorders, Institute of Biomedical Sciences, School of Medicine, Jianghan University, Wuhan, China;

[2] Department of Pathogenic Biology, School of Medicine, Jianghan University, Wuhan, China;

[3] State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan China

Abstract: Although parasitic copepods of the genus *Ergasilus* von Nordmann, 1832 are globally distributed parasites of fish, their phylogenetic relationships with other Copepoda are not clear, and the characteristics of their mitochondrial genomes (mitogenomes) are not thoroughly understood. The objective of this study was to address these knowledge gaps by sequencing the complete mitogenome of *Ergasilus tumidus* Markevich, 1940. The complete mitogenome (GenBank Acc. No. OQ596537) was 14,431 bp long and it comprised 13 protein-coding genes (PCGs), 22 tRNAs, two tRNAs, and two control regions (CRs). Phylogenetic analyses, conducted using concatenated nucleotide and amino acid sequences of 13 protein-coding genes, produced two partially incongruent topologies. While the order Calanoida was consistently resolved as the sister lineage to the other three orders, topological instability was observed in the relationships of the orders Cyclopoida, Siphonostomatoida and Harpacticoida. Siphonostomatoida clustered with Cyclopoida in the nucleotide-based phylogeny, but with Harpacticoida in the amino acid-based phylogeny. The latter topology conforms to the widely accepted relationships, but we speculate that the former topology is more likely to be the correct one. Our study provides a complete mitogenome sequence of *E. tumidus*, which helps us better understand the molecular evolution of the genus *Ergasilus*. Additionally, we suggest a different perspective on the controversial phylogenetic relationships among Siphonostomatoida, Cyclopoida and Harpacticoida, diverging from previously accepted views.

Keywords: *Ergasilus*, gene order, mitogenome, phylogeny

This article contains supporting tables (Tables S1–S3; Figures S1–S6) online at http://folia.paru.cas.cz/suppl/2024-71-002.pdf

Copepods inhabit diverse freshwater environments, including but not limited to lakes, subterranean waters, glacial meltwater pools, hot springs, hypersaline lakes, and phytotelmata (Boxshall and Jaume 2000). As a result, copepods have a global distribution and hold significant importance in shaping the structure of planktonic, benthic and groundwater ecosystems. Furthermore, they are capable of adapting to semi-terrestrial habitats such as moist moss and leaf litter found in humid forests (Boxshall and Defaye 2007). Their ecological roles in aquatic ecosystems are diverse. While some copepods offer benefits in aquaculture, others pose considerable risks, as parasitic copepod infections can lead to economic losses when farmed fish of commercial importance suffer severe infections (Piasecki et al. 2004, Mathews et al. 2018). Within the realm of copepod parasites that affect fish, the family Ergasilidae encompasses species with strong pathogenic potential (Hadfield and Smit 2019), Several of these species have been linked to significant mortality rates among cultured fish in freshwater and brackish environments (Delgado et al. 2011).

Within the family Ergasilidae, the genus *Ergasilus* von Nordmann, 1832 is considered to be the most diverse, with 163 described species, along with a possibility of an additional 34 unnamed members existing within species complexes across the globe (Walter 2024). As such, it holds a significant level of ergasilid diversity. Remarkably, a single fish can be infected with several thousand *Ergasilus* parasites, which primarily attach themselves to the fish gills. These species of *Ergasilus* can transmit various diseases, making them one of the most important parasite groups for aquaculture (Lester and Hayward 2006). Their global distribution further contributes to their importance in the field (Kabata 1979). Due to both their immense significance for the aquaculture industry and the extensive size of the

Address for correspondence: Cong-Jie Hua, Wuhan Institutes of Biomedical Sciences, School of Medicine, Jianghan University, Wuhan, 430056, China.
E-mail: huacongjie@jhun.edu.cn; ORCID: 0000-0002-4344-1813

genus, *Ergasilus* has been the subject of more taxonomic studies than any other ergasilid taxon. Surprisingly, despite this attention, no studies on mitogenomes of species of *Ergasilus* have been conducted thus far.

The mitochondrion, an essential organelle found in eukaryotic cells, possesses its own genome, distinct from the nuclear genome (Boore 1999). Initially regarded as the perfect molecular marker for investigating the evolutionary history of life on Earth, mitogenomes offer multiple comparative advantages, such as unilinear inheritance and the absence of recombination (Avise et al. 1987). Typically, mitogenomes are relatively small, maternally inherited, and demonstrate minimal levels of recombination (Rand 1994). Consequently, they have been extensively utilised as molecular markers for species identification, phylogenetic inference and research on population structure (Shao and Barker 2007).

Owing to the lack of mitochondrial genomic data for *Ergasilus*, the aim of this study was to generate mitogenomic data and use them to obtain a deeper understanding of the genetic features of *Ergasilus* and the phylogenetic relationships of Copepoda.

MATERIALS AND METHODS

Specimen collection and identification

Specimens of *Ergasilus tumidus* Markevich, 1940 were collected from the gills of *Rhodeus* sp. obtained from East Lake in Wuhan, Hubei Province, China, on 27 April 2021. The parasites were identified according to their morphological traits (Kuang and Qian 1991), and their identity was further verified using the 18S rDNA sequence, which was amplified with primers reported in Song et al. (2007). Before DNA extraction, the specimens were conserved in absolute alcohol and kept at -20°C at the Jianghan University Institute of Biology and Medical Sciences, located in Wuhan, China. The complete genomic DNA was extracted from a solitary individual (a whole specimen) utilising the SDS/Proteinase K method and the TIANamp Genomic DNA kit, following the prescribed guidelines provided by the manufacturer (Tiangen Biotech, Beijing, China).

Mitogenome amplification, assembly, annotation and characterisation

We designed primers (see Supplementary Table S1) targeting conserved regions of mitochondrial genes identified in previously published copepod mitogenomes. These fragments, once amplified and sequenced, were subsequently used for the design of specific primers to amplify the complete mitogenomes (Supplementary Table S1). The DNA extracted from *E. tumidus* was forwarded to Sangon Biotech Co., Ltd (Shanghai, China) for Sanger sequencing analysis. The mitogenome was assembled from raw data using Geneious (Kearse et al. 2012), and the assembled mitogenome was annotated using the MITOS WebServer (Bernt et al. 2013).

Subsequently, various tools such as BLAST and ORF Finder, provided by the NCBI, and MAFFT 7.0 (Katoh and Standley 2013), were utilised to fine-tune the annotation of protein-coding genes (PCGs), taking into consideration orthologues from related species (Peng et al. 2010). The boundaries of the two rRNA genes were provisionally identified by comparing their sequences with those of published copepod mitochondrial rRNAs. The structural maps of the mitogenome were generated using OGDRAW 1.3.1 (Lohse et al. 2007). To predict the secondary structures of tRNAs, we employed tRNAscan-SE online Search Server (http://lowelab.ucsc.edu/tRNAscan-SE/) (Lowe and Eddy 1997), followed by manual verification. Additionally, the RNA Structure tool (http://rna.urmc.rochester.edu/RNAstructureWeb/) (Reuter and Mathews 2010) was utilised to predict the secondary structures of *rrnS* and *rrnL*.

Base composition, codon usage, relative synonymous codon usage (RSCU), and amino acid content were calculated using MEGA v.7.0 (Tamura et al. 2013). The ggplot2 plugin in PhyloSuite v.1.2.3 (Xiang et al. 2023) was utilised to generate the RSCU figure. Base composition skews were computed using the formulas: AT-skew = [A - T] / [A + T] and GC-skew = [G - C] / [G + C] via MEGA v.7.0 (Perna and Kocher 1995). Three-dimensional scatter plots depicting AT-skew, GC-skew and AT% were created using OriginPro v.9.0 (Abdullah and Khairurrijal 2009). Sliding window analysis, with a window size of 200 bp and a step size of 20 bp, was performed using DnaSP v.5 (Librado and Rozas 2009) to estimate nucleotide divergence (Pi) among copepod mitogenome's PCGs, rRNAs, and tRNAs.

Additionally, Kimura-2-parameter (K2P) genetic distances of mt PCGs (including substitutions, i.e., transitions and transversions) were calculated using MEGA X (Kumar et al. 2018). Selection pressures on the 13 PCGs were assessed by computing Ka (non-synonymous mutation rate) and Ks (synonymous mutation rate) values through KaKs_Calculator 2.0 (Wang et al. 2013). All software packages were used according to their respective manuals, and a model averaging method was employed to calculate Ka, Ks and Ka/Ks. This method encompassed 14 different models for calculation and determined the average values of Ka, Ks and Ka/Ks based on these models. Sequence motifs in the control region (CR) were identified by employing TandemRepeats Finder (Benson 1999).

Phylogenetic analyses and gene order analyses

Phylogenetic analyses were performed utilising the newly sequenced mitogenome of *E. tumidus*, along with 33 other mitogenomes (criterion: mitogenomes had to have more than 10,000 bp; Table 1). The nucleotide and amino acid datasets employed for phylogenetic analyses consisted of the amino acid alignment of 13 protein-coding genes (PCGs). Some mitogenomes, namely those of *Calanus finmarchicus*, *Calanus glacialis*, *Calanus sinicus*, *Caligus clemensi*, *Caligus rogercresseyi*, *Lovenula raynerae*, and *Schizopera knabeni*, were either incomplete or incompletely annotated as they lacked certain PCGs or tRNA genes. PhyloSuite was utilised to extract the data from GenBank files, including the nucleotide sequences of PCGs, rRNAs, and tRNAs. Following that the nucleotide sequences of PCGs were translated into their respective amino acid sequences using PhyloSuite.

Gene alignments were performed using MAFFT 7.0 (Katoh and Standley 2013), which is integrated within PhyloSuite, using the normal alignment mode for all genes. After eliminating ambiguously aligned regions with Gblocks (Talavera and Castresana 2007), the alignments were concatenated using PhyloSuite. PartitionFinder 2 (Lanfear et al. 2017) was utilised to determine the most suitable evolutionary models for conducting maximum likelihood (ML) and Bayesian inference (BI) phylogenetic anal-

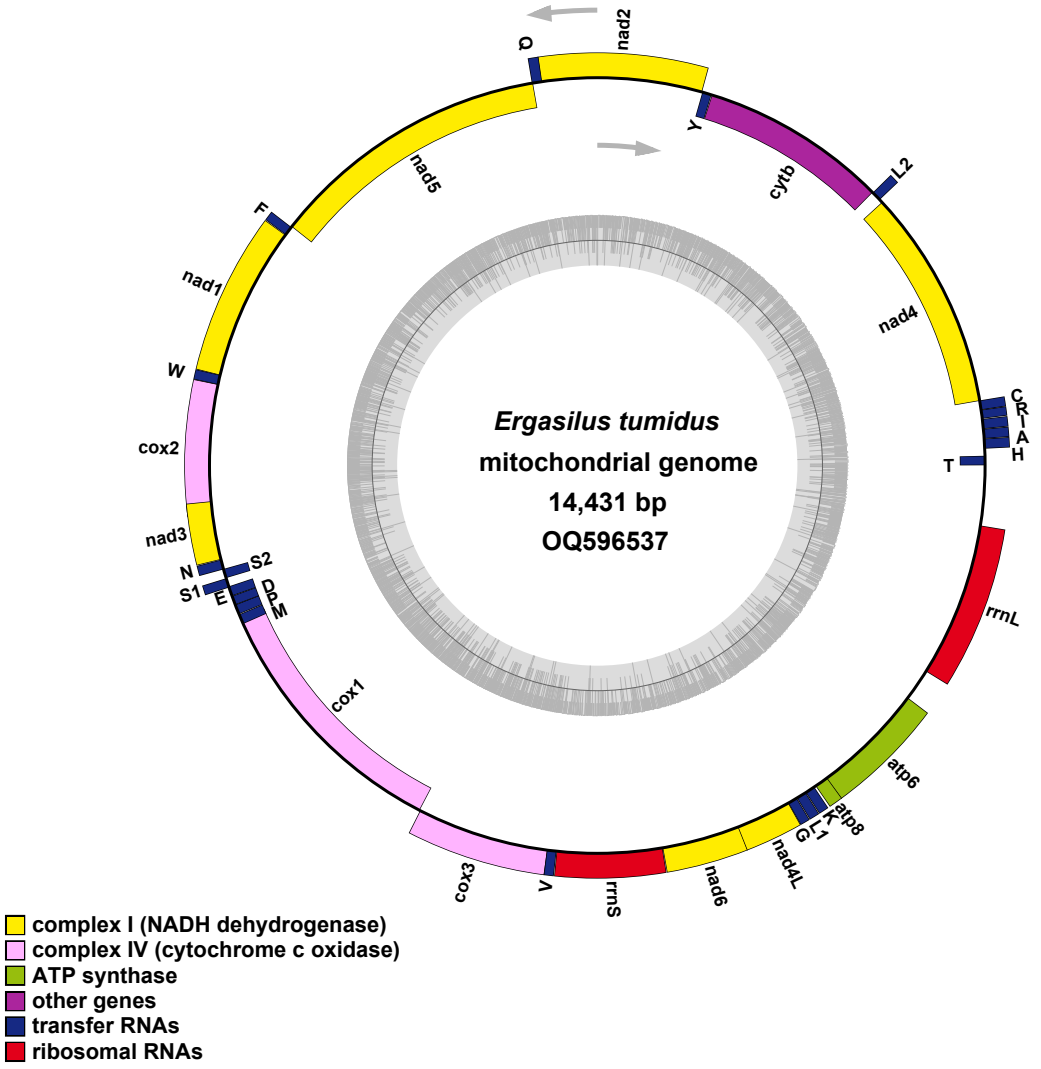

Fig. 1. The structure of the mitogenome of *Ergasilus tumidus* Markevich, 1940.

yses. These analyses were performed using two PhyloSuite plugins: IQ-TREE (Nguyen et al. 2015) and MrBayes (Ronquist et al. 2012), respectively. *Vulcanolepas fijiensis*, *Metaplax longipes*, *Scyllarides squammosus*, *Austinogebia edulis*, *Typhlatya galapagensis*, *Munidopsis verrilli*, *Longpotamon parvum*, *Daphnia pulex*, and *Artemia tibetiana* were used as outgroups. Topology support in the ML analysis was inferred by calculating bootstrap support from 1,000 replicates. For the BI analysis, two independent runs were executed, each comprising four chains that ran concurrently for 2,000,000 generations. Sampling was conducted at intervals of 315 steps, with a burn-in rate of 25%. MLGO (Hu et al. 2014) was used to reconstruct the phylogeny using the mitogenomic gene order data, with 1,000 bootstrap replicates, and an input file generated by PhyloSuite. Phylograms and gene orders were visualised in iTOL (Letunic and Bork 2019) and annotated using files generated by PhyloSuite.

RESULTS

Nucleotide composition and genome organisation

The circular mitogenome of *Ergasilus tumidus* was complete and spanned 14,431 base pairs (bp.) (Fig. 1). It consisted of 37 genes, which included 13 PCGs, 22 tRNA genes and two rRNA genes. Additionally, it had two control regions (CRs). Among these genes, 25 (including 10 PCGs and 15 tRNAs) were located on the heavy strand (H-strand), while the remaining ten genes (comprising three PCGs and seven tRNAs) were encoded on the light strand (L-strand).

Notably, *E. tumidus* shared the same gene order as *Sinergasilus undulatus* (Supplementary Fig. S1). The mitogenome of *E. tumidus* exhibited a high AT content of 65.7% (Supplementary Table S2). The AT-skew of *E. tumidus* (0.006) was higher than the average AT-skew observed in copepod mitogenomes (-0.0208). In contrast, the GC-skew (0.0170) was lower than the copepods' average.

The scatterplots depicting three-dimensional relationships among AT content, AT-skew and GC-skew in mitogenomes of copepods are shown in Supplementary Fig. S2. The range of AT-skew varied from -0.170 in *Tigriopus japonicus* Mori, 1938 to 0.097 in *S. undulatus*. Similarly, the range of GC-skew spanned from -0.235 in *Lamproglena chinensis* Yü, 1937 to 0.268 in *T. japonicus*. Most species belonging to the orders Siphonostomatoida and Calanoida exhibited sim-

Table 1. Detailed sequence information of mitochondrial genomes used in the present phylogenetice analysis.

Order	Species	Total size (bp)	PCG size (bp)	tRNA size (bp)	rRNA size (bp)	GenBank
Calanoida	*Calanus sinicus*	20,460	11,133	1,268	1,793	GU355641.1
	Calanus glacialis	20,674	8,829	1,082	1,794	MF422146.1
	Calanus finmarchicus	17,682	6,618	1,031	664	MG001887.1
	Calanus hyperboreus	17,910	11,103	1,379	1,728	NC_019627.1
	Calanus simillimus	27,876	11,091	1,373	1,837	NC_063666.1
	Lovenula raynerae	14,365	11,034	1,384	1,073	MH710604.1
	Phyllodiaptomus tunguidus	16,446	11,076	1,357	1,683	NC_046743.1
	Labidocera rotunda	16,564	11,022	1,381	1,729	NC_064109.1
	Eurytemora affinis	18,573	10,977	1,339	1,730	NC_046694.1
Cyclopoida	*Paracyclopina nana*	15,981	10,908	1,461	1,793	EU877959.1
	Ergasilus tumidus	14,431	11,127	1,307	1,559	OQ596537.1
	Lamproglena orientalis	28,462	21,720	2,558	3,286	OQ411235.1
	Lamproglena chinensis	13,933	10,857	1,286	1,566	OQ411234.1
	Lernaea cyprinacea	14,656	10,827	1,287	1,539	NC_025239.1
	Sinergasilus polycolpus	14,000	10,986	1,314	1,303	NC_028085.1
	Sinergasilus undulatus	14,239	11,001	1,328	1,576	NC_054173.1
Harpacticoida	*Tigriopus japonicus*	14,628	10,959	1,370	1,614	AB060648.1
	Tigriopus kingsejongensis	14,940	10,920	1,362	1,600	MK598762.1
	Tigriopus californicus	14,578	10,962	1,306	1,587	NC_008831.2
	Schizopera knabeni	10,649	8,151	1,014	1,554	KF667527.1
	Amphiascoides atopus	13,831	10,920	1,250	1,540	NC_023783.1
Siphonostomatoida	*Caligus rogercresseyi*	13,468	6,552	1,266	1,651	HQ157565.1
	Caligus clemensi	13,440	7,593	1,251	1,650	HQ157566.1
	Lepeophtheirus salmonis	16,148	10,812	1,314	1,661	NC_056769.1
	Pennella sp.	14,620	10,869	1,333	1,655	ON161759.1
Outgroup	*Vulcanolepas fijiensis*	17,374	11,085	1,441	2,046	MN061491.1
	Metaplax longipes	16,305	11,190	1,497	2,253	NC_040976.1
	Scyllarides squammosus	15,644	11,169	1,482	2,202	NC_044425.1
	Austinogebia edulis	15,761	11,163	1,465	2,309	JN897376.1
	Typhlatya galapagensis	16,430	11,154	1,444	2,054	KX844711.1
	Munidopsis verrilli	17,636	11,124	1,472	2,156	MH717896.1
	Longpotamon parvum	19,637	11,157	1,526	2,139	MN737134.1
	Daphnia pulex	15,333	11,070	1,453	2,067	NC_000844.1
	Artemia tibetiana	15,826	10,560	1,388	1,872	JQ975177.1

ilar AT content and AT/GC-skews, resulting in a relatively close distribution on the three-dimensional scatter plot. On the contrary, the values of the AT content, AT-skew and GC-skew in the orders Cyclopoida and Harpacticoida displayed a relatively scattered distribution on the plots.

Protein-coding genes

The concatenated protein-coding genes (PCGs) of *E. tumidus* had a total length of 11,127 base pairs. With the exception of *cytb* and *cox2*, which employed TAC and TTG start codons, respectively, all other genes employed ATA or ATG as start codons. The termination codons used by all PCGs were TAA and TAG (Table 2).

The RSCU values of copepode mitogenomes are provided in Supplementary Fig. S3 and summarised in Supplementary Table S3. Copepods exhibited varying frequencies of synonymous codon usage (Supplementary Fig. S4), with the codon UUA being the most frequently utilised, followed by CCU, UCU, and AGA. Among the 25 investigated mitogenomes, leucine (Leu) was the most commonly used amino acid, accounting for an average of 16.41% usage, followed by serine (Ser; 11.00%), phenylalanine (Phe; 9.77%), and isoleucine (Ile; 6.32%). On the other hand, arginine (Arg) had the lowest percentage at 1.79%. It is worth noting that there were no significant differences observed in amino acid usage among the species.

Nucleotide diversity and evolutionary rate analysis

A sliding window analysis was conducted using concatenated alignments of 13 PCGs, two rRNAs, and 22 tRNAs (Fig. 2). The analysis revealed a significant range of nucleotide diversity among copepods, as indicated by the Pi values for the 200 bp windows, which ranged from 0.302 to 0.555 (Fig. 2). Among the PCGs, *cox1* (0.302), *cytb* (0.374), and *cox3* (0.378) showed relatively low sequence variability, whereas *nad4L* (0.555), *nad2* (0.551), *nad6* (0.548), and *nad4* (0.520) exhibited higher sequence variability. This pattern was further supported by the Kimura-2-parameter (K2P) genetic distance analysis, where *nad4L* exhibited the highest value among the 13 PCGs in the 18 copepod mitogenomes, followed by *nad2* and *nad6*. The non-synonymous/synonymous (Ka/Ks) ratio analysis confirmed these findings.

The ratios of Ka and Ks substitutions for PCGs are depicted in Supplementary Fig. S5. All of the Ka/Ks ratios were below 1, with *nad4L* displaying the highest value (0.799), followed by six genes (*nad6*, *nad2*, *atp8*, *nad3*, *nad4*, *atp6*) with Ka/Ks ratios ranging from 0.191 to 0.613. On the other hand, *cox1*, *cox2*, *cox3*, and *cytb* exhibited low Ka/Ks ratios, varying from 0.037 (*cox1*) to 0.187 (*nad5*). The values indicated that all PCGs have undergone purifying selection, particularly *cox1*, *cox2*, *cox3*, and *cytb*.

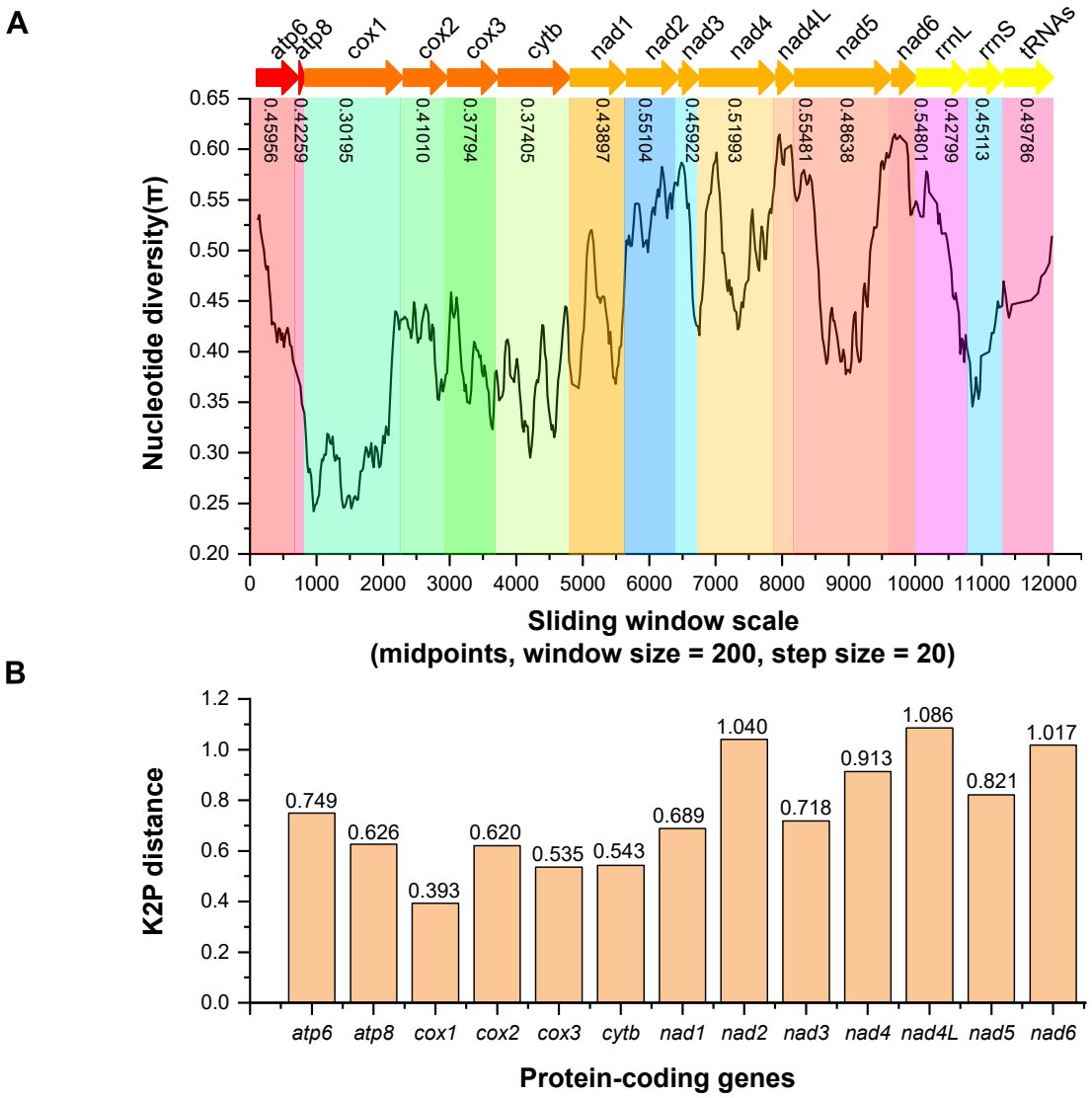

Fig. 2. Evolutionary analyses of 18 copepod mitogenomes. **A** – sliding window analysis of the alignment of 13 PCGs, two rRNAs, and 22 tRNAs. The black line represents nucleotide variation in a window of 200 bp (step size = 20 bp, with the values inserted at its midpoint). Gene boundaries are indicated by coloured segments, with the mean variation ratio per gene shown beneath each gene. **B** – the Kimura-2-parameter distance (K2P) among 13 PCGs of 18 copepod mitogenomes.

Transfer RNAs, ribosomal RNAs and CRs

In *E. tumidus*, the cumulative length of tRNAs was 1,307 bp, while individual tRNAs ranged in length from 54 to 63 bp. Most tRNAs exhibited the typical clover-leaf structure, except for *trnS1*, *trnS2*, and *trnR*, which lacked the dihydrouridine (DHU) arm, and *trnC*, which lacked the TψC (T) arm (Supplementary Fig. S6). There were a total of 11 mismatched base pairs observed in the tRNAs of *E. tumidus*. Among them, six were U-G pairs located in the amino acid acceptor (AA) arm, the anticodon (AC) arm, and the T arm. Four U-U pairs were found in the AA arm, while a single A-G pair was identified in the AA arm. The length of rRNAs in *E. tumidus* was 1,559 bp, with an AT content of 71.7%.

The CR of *E. tumidus* was positioned between *atp6* and *rrnL*, spanning 189 bp, with an AT content of 73.1%. The second CR was situated downstream of *rrnL*, covering 357 bp with an AT content of 62.2%. No tandem repeats were detected in either of these two CRs.

Gene order analyses

In the copepod dataset that we had at disposal (excluding incomplete mitogenomes), the families Ergasilidae and Lernaeidae of the Cyclopoida demonstrated the most conserved gene arrangement among the studied species. In comparison to the ground pattern of copepods, the only gene rearrangements identified in these two families were *trnS2* rearrangement from its original position between *trnS1* and *trnN* to a position between *trnT* and *atp6* in *Sinergasilus polycolpus*, and *trnL1* rearrangement from its original position between *nad1* and *nad4* to a position between *trnH* and *trnS1* in *Lernaea cyprinacea* Linnaeus, 1758. Notably, *trnH* and *trnS1* switched strands in *L. cyprinacea* in comparison to the other two lernaeid species. All other copepod species exhibited unique gene arrange-

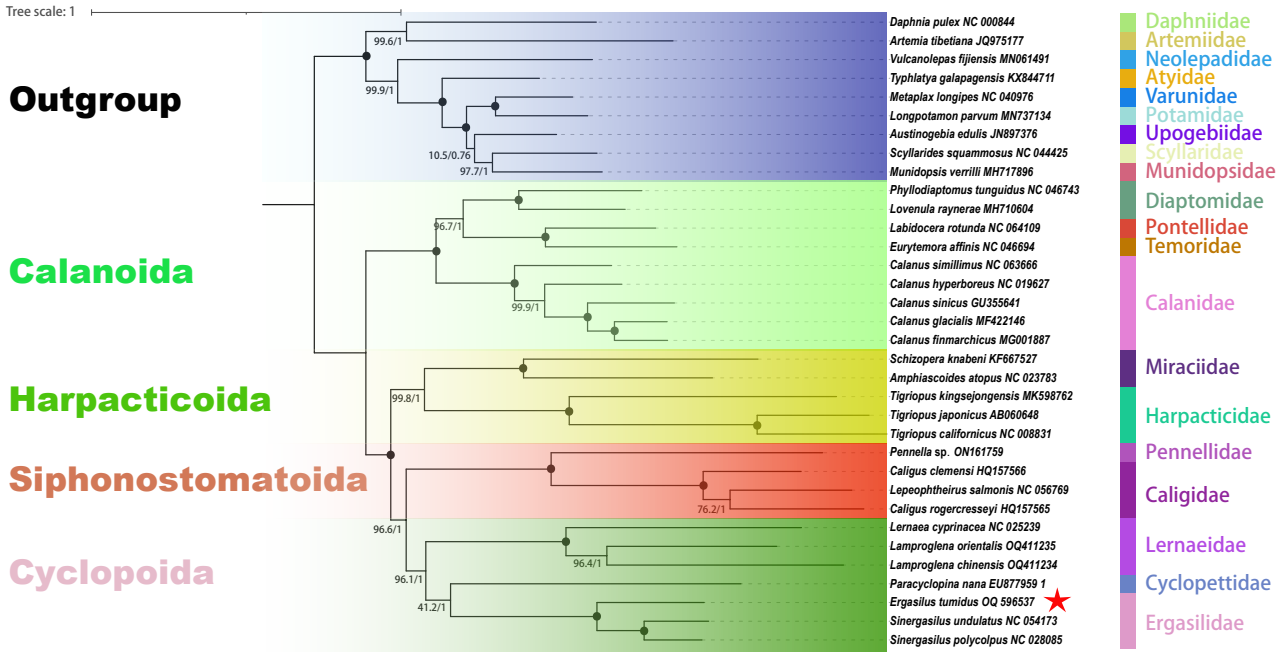

Fig. 3. Phylogenetic relationships of 34 copepods inferred using nucleotide sequences of 13 mitogenomic protein-coding genes. Statistical support values (bootstrap/posterior probability) for ML/BI analyses are shown at the nodes. A red pentagram highlights *E. tumidus*, the species newly sequenced for this study. Circles indicate ML/BI = 100/1.0; other values are given above the nodes. The GenBank accession numbers of the 34 mitogenome sequences are listed in Table 1.

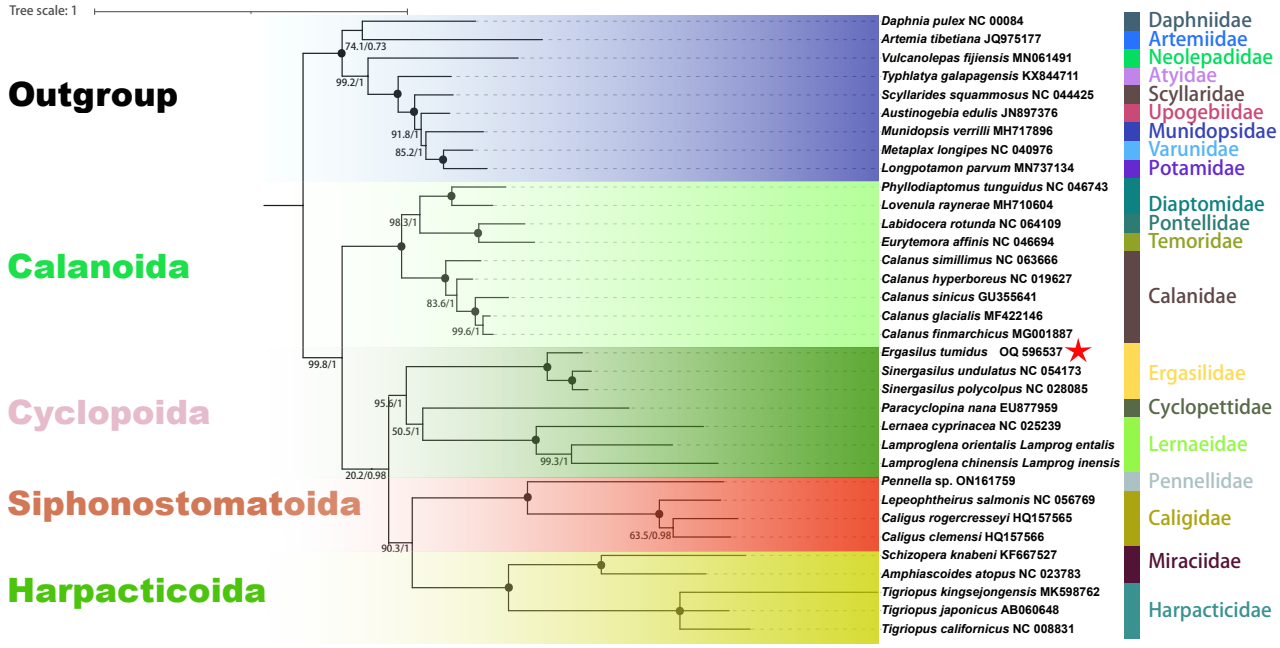

Fig. 4. Phylogenetic relationships of 34 copepods inferred using amino acid sequences of 13 mitogenomic protein-coding genes. Statistical support values (bootstrap/posterior probability) for ML/BI analyses are shown at the nodes. A red pentagram highlights *E. tumidus*, the species newly sequenced for this study. Circles indicate ML/BI = 100/1.0; other values are given above the nodes. The GenBank accession numbers of the 34 mitogenomes are listed in Table 1.

ments, with differences primarily limited to the positioning of tRNA genes, but also several rearranged PCGs.

Phylogenetic relationships

Bayesian inference (BI) and maximum-likelihood (ML) analyses based on the nucleotide dataset and amino acid dataset produced incongruent phylogenetic trees (Figs. 3 and 4). The four orders included in the datasets, Calanoida, Harpacticoida, Siphonostomatoida and Cyclopoida, were monophyletic in both analyses, with the posterior probability (pp) = 1 and bootstrap values (bv) in the range between 99.6% and 100% in BI and ML analyses. In the nucleotide-based tree, the order Calanoida formed a sister clade to the remaining three orders, with high support (pp =

Table 2. Organisation of the mitochondrial genomes of *Ergasilus tumidus* Markevich, 1940

Gene	Strand	Position (bp)	Length (bp)	Intergenic nucleotides	Start/Stop codon
trnT	H	1–58	58	0	
trnH	L	100–159	60	41	
trnA	L	159–217	59	-1	
trnI	L	217–276	60	-1	
trnR	L	278–331	54	1	
trnC	L	330–387	58	-2	
nad4	H	384–1715	1,332	-4	ATA/TAG
trnL1	L	1734–1793	60	18	
cytb	H	1793–2920	1,128	-1	ATC/TAA
trnY	H	2921–2978	58	0	
nad2	L	2980–3942	963	1	ATA/TAA
trnQ	L	3969–3997	62	-7	
nad5	H	3990–5693	1,704	-8	ATA/TAA
trnF	L	5695–5752	58	1	
nad1	L	5753–6676	924	0	ATA/TAA
trnW	L	6675–6733	59	-2	
cox2	L	6731–7504	774	-3	TTG/TAG
nad3	L	7428–7778	351	-77	ATA/TAA
trnN	L	7779–7838	60	0	
trnS1	H	7835–7892	58	-4	
trnS2	L	7893–7950	58	0	
trnE	H	7943–8004	62	-8	
trnD	H	8003–8060	58	-2	
trnP	H	8058–8119	62	-3	
trnM	H	8120–8182	63	0	
cox1	H	9148–9731	1,584	-35	ATA/TAA
cox3	L	9734–10528	795	2	ATG/TAA
trnV	L	10524–10581	58	-5	
rrnS	L	10582–11211	630	0	
nad6	L	11212–11682	471	0	ATA/TAG
nad4L	L	11675–12007	333	-8	ATG/TAA
trnG	L	12006–12064	59	-2	
trnL2	L	12065–12126	62	0	
trnK	L	12127–12187	61	0	
atp8	L	12197–12280	84	9	ATA/TAA
atp6	L	12273–12956	684	-8	ATG/TAA
NCR1		12957–13145	188	0	
rrnL	L	13146–14074	929	189	
NCR2		14075–14431	356	0	

1.00 and bv = 100%), while Siphonostomatoida and Cyclopoida were sister lineages (pp = 1.00 and bv = 67.6%). In contrast, in the amino acids-based tree, Siphonostomatoida clustered with Harpacticoida (pp = 1.00 and bv = 90.3%) (Fig. 4). The clade of Harpacticoida, Siphonostomatoida and Cyclopoida was well supported (pp = 1.00 and bv = 99.9%) in the nucleotide-based tree, while the support was low in the ML analysis of the amino acids dataset (20.2%).

Within the order Calanoida, four families were included in the dataset: Diaptomidae, Pontellidae, Temoridae and Calanidae. All four families were monophyletic (pp = 1.00 and bv = 100%). In the nucleotide-based tree, Calanidae formed the sister group to the remaining three families. In the order Harpacticoida, Miraciidae (pp = 1.00 and bv = 100%) and Harpacticidae (pp = 1.00 and bv = 100%) were monophyletic, while Pennellidae (pp = 1.00 and bv = 100%) and Caligidae (pp = 1.00 and bv = 100%) were polyphyletic in the order Siphonostomatoida. In the order Cyclopoida, all three included families were monophyletic. Regarding the amino acid dataset-based tree, all families in all four represented orders were monophyletic with high support values. *Ergosilus tumidus* clustered with *S. undulatus* and *S. polycolpus* in all analyses, which confirmed their close relationships.

DISCUSSION

Gene translocations

Mitogenome rearrangements are very common in copepods (Machida et al. 2004, Jung et al. 2006, Kilpert and Podsiadlowski 2006), so they have been recognised as one of the two (Isopoda being the other one) crustacean lineages exhibiting the highest level of mitogenomic architectural plasticity (Jakovlić et al. 2021). The gene arrangement of the mitogenome of *Ergasilus tumidus*, sequenced in this study, also deviates from the typical gene arrangement pattern observed in arthropods (Staton et al. 1997). However, the gene order of *E. tumidus* is identical to that of *Sinergasilus undulatus* and very similar to that of *Sinergasilus polycolpus*, with only one transposition event: a position change of *trnS2*. The gene orders of *E. tumidus* and *S. polycolpus* reveal that ergasilid species do not exhibit a synapomorphic gene order, implying that this family appears to undergoing relatively rapid architectural rearrangement rates.

To assess the phylogenetic value of gene order, we employed the MLGO algorithm to reconstruct a phylogeny based on gene orders. However, the reconstruction using MLGO yielded a topology that differed from the trees constructed based on mitochondrial sequence data (Supplementary Fig. S7). Only the monophyly of Pontellidae + Temoridae was well supported by all three phylogenetic trees. It is worth noting that the paraphyly of Harpacticidae + Ergasilidae + Miraciidae produced in the gene order-based phylogeny partially supported a closer phylogenetic relationship between Cyclopoida and Harpacticoida, which was a result not observed in the phylogenetic tree constructed based on mitogenome sequences. If tRNAs were not considered, there were four gene arrangement patterns among these three families, indicating that their phylogenetic relationship is not close. Thus, tRNA genes are commonly considered to be highly mobile and they are often excluded from gene order-based phylogenetic analyses (Kim et al. 2017). Moreover, a previous study has shown that gene orders can generate artifactual relationships when used for phylogenetic reconstruction (Zhang et al. 2019). Therefore, caution should be exercised when utilising gene order-based phylogenetics in phylogenetic analyses.

Phylogenetic relationships

In the phylogenies based on nucleotide sequences inferred in this study, the order-level relationships differ from the widely accepted phylogenetic hypothesis proposed by a majority of morphological and molecular phylogenetic studies, wherein Harpacticoida is closely affiliated with Siphonostomatoida, rather than with Cyclopoida (Huys et al. 2006, 2007, Minxiao et al. 2011, Tung et al. 2014). As for the topologies produced by amino acid sequences, the order-level relationships inferred from these phylogenomic

analyses are congruent with the widely accepted relationships, with Harpacticoida closely related to Siphonostomatoida instead of Cyclopoida (Huys et al. 2006, 2007, Minxiao et al. 2011, Tung et al. 2014).

The remaining topological features are congruent with the trees based on nucleotide sequences inferred in this study. The relationships among Cyclopoida, Siphonostomatoida and Harpacticoida were highly supported in the BI analysis, but poorly supported in the ML analysis (bv = 20%). However, BI is known to overestimate support for nodes (Suzuki et al. 2002), so we cannot treat the inferred relationships among Siphonostomatoida, Cyclopoida and Harpacticoida as reliable (Huelsenbeck and Hillis 1993). Bootstrapping produces more reliable results than Bayesian posterior probabilities (Suzuki et al. 2002), and nucleotide-based trees exhibited higher bv (and pp values) than amino acids-based trees. On this basis, we infer that nucleotide-based dataset may have produced more reliable topologies than the amino acids-based dataset.

Our results indicate that Siphonostomatoida clusters with Cyclopoida, and that Harpacticoida is a sister lineage to that clade. The phylogenetic status of Cyclopoida and Harpacticoida is controversial in the history of copepod classification (Huys et al. 2007, Minxiao et al. 2011, Tung et al. 2014, Eyun 2017, Bernot et al. 2021). However, one previous study produced a topology that was congruent with ours (Eyun 2017). It is important to note that this study was based on a relatively large number of genes (24 orthologous nuclear protein-coding genes) and it tested the performance of several different models. In comparison, the studies producing the Harpacticoida-Siphonostomatoida sister group relationship mostly relied on Bayesian analyses of nucleotide sequences and amino acid sequences of 12 PCGs, as well as Bayesian and Maximum Likelihood analyses based on 18S rDNA (Eyun 2017).

In addition, a previous study found evidence of host-induced morphological variation in some copepod species, which indicates that reliance on morphological traits alone may produce taxonomic artifacts in copepods (Hua et al. 2019). On this basis, we propose that the topologies grouping Siphonostomatoida and Cyclopoida together may more closely reflect the true evolutionary sequence of events. However, since these three orders encompass almost all parasitic copepods (Bernot et al. 2021), more molecular data are needed to evaluate the phylogenetic relationships among them.

Acknowledgements. The authors are grateful to Ivan Jakovlić for polishing the language and logical structure. We would also like to thank two anonymous reviewers, Jan Stefka and Tomas Scholz for investing time and expertise into reviewing our manuscript.

Authors' contribution. GTW, BLS, XJS and CJH were involved in designing the study. YH, JMF and WL conducted the experiments. YH and CJH performed the data analysis. YH and CJH also contributed to writing the paper. All authors have revised the manuscript critically for important intellectual content and read and approved the final manuscript.

REFERENCES

Abdullah M., Khairurrijal K. 2009: A simple method for determining surface porosity based on SEM images using OriginPro software. Indones. J. Phys. 20: 37–40.

Avise J.C., Arnold J., Ball R.M., Bermingham E., Lamb T., Neigel J.E., Reeb C.A., Saunders N.C. 1987: Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489–522.

Benson G. 1999: Tandem repeats finder: a program to analyze DNA sequences. Nucl. Acids Res. 27: 573–580.

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

Bernt M., Donath A., Jühling F., Externbrink F., Florentz C., Fritzsch G., Pütz J., Middendorf M., Stadler P.F. 2013: MITOS: improved *de novo* metazoan mitochondrial genome annotation. Mol. Phylogen. Evol. 69: 313–319.

Boore J.L. 1999: Animal mitochondrial genomes. Nucl. Acids Res. 27: 1767–1780.

Boxshall G.A., Defaye D. 2007: Global diversity of copepods (Crustacea: Copepoda) in freshwater. Hydrobiologia 595: 195–207.

Boxshall G.A., Jaume D. 2000: Making waves: the repeated colonization of fresh water by copepod crustaceans. Adv. Ecol. Res. 31: 61–79.

Delgado P.M., Delgado J.P.M., Arenas J.V., Orbe R.I. 2011: Massive infestation by *Perulernaea gamitanae* (Crustacea: Cyclopoida: Lernaidae) in juvenile gamitana, cultured in the Peruvian Amazon. Vet. Mex. 42: 59–64.

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

Hadfield K.A., Smit N.J. 2019: Parasitic Crustacea as vectors. In N.J. Smit, N.L.Bruce and K.A. Hadfield (Eds.), Parasitic Crustacea. State of Knowledge and Future Trends. Springer, Cham, pp. 331–342.

Hu F., Lin Y., Tang J. 2014: MLGO: phylogeny reconstruction and ancestral inference from gene-order data. BMC Bioinformatics 15: 354.

Hua C.J., Zhang D., Zou H., Li M., Jakovlić I., Wu S.G., Wang G.T., Li W.X. 2019: Morphology is not a reliable taxonomic tool for the genus *Lernaea*: molecular data and experimental infection reveal that *L. cyprinacea* and *L. cruciata* are conspecific. Parasit. Vectors 12: 579.

Huelsenbeck J.P., Hillis D.M. 1993: Success of phylogenetic methods in the four-taxon case. Syst. Biol. 42: 247–264.

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

Huys R., Llewellyn-Hughes J., Olson P.D., 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.

Jakovlić I., Zou H., Zhao X.-M., Zhang J., Wang G.-T., Zhang D. 2021: Evolutionary history of inversions in directional mutational pressures in crustacean mitochondrial genomes: implications for evolutionary studies. Mol. Phylogen. Evol. 164: 107288.

Jung S.-O., Lee Y.-M., Park T.-J., Park H.G., Hagiwara A., Leung K.M.Y., Dahms H.-U., Lee W., Lee J.-S. 2006: The

complete mitochondrial genome of the intertidal copepod *Tigriopus* sp. (Copepoda, Harpactidae) from Korea and phylogenetic considerations. J. Exp. Mar. Biol. Ecol. 333: 251–262.

Kabata Z. 1979: Parasitic Copepoda of British Fishes. The Ray Society, British Museum, London, 468 pp.

Katoh K., Standley D.M. 2013: MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30: 772–780.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C. 2012: Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647–1649.

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.

Kim J., Kern E., Kim T., Sim M., Kim J., Kim Y., Park C., Nadler S.A., Park J.-K. 2017: Phylogenetic analysis of two *Plectus* mitochondrial genomes (Nematoda: Plectida) supports a sister group relationship between Plectida and Rhabditida within Chromadorea. Mol. Phylogen. Evol. 107: 90–102.

Kuang P., Qian J. 1991: Economic Fauna of China: Parasitic Crustacea of Freshwater Fishes. Science Press, Beijing, 203 pp.

Kumar S., Stecher G., Li M., Knyaz C., Tamura K. 2018: MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol. Biol. Evol. 35: 1547.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2017: PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34: 772–773.

Lester R.J., Hayward C.J. 2006: Phylum Arthropoda. In P.T.K. Woo (Ed.), Fish Diseases and Disorders. Volume 1: Protozoan and Metazoan Infections. CABI, Wallingford, pp. 466–565.

Letunic I., Bork P. 2019: Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucl. Acids. Res. 47: W256–W259.

Librado P., Rozas J. 2009: DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.

Lohse M., Drechsel O., Bock R. 2007: OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr. Genet. 52: 267–274.

Lowe T.M., Eddy S.R. 1997: tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl. Acids Res. 25: 955–964.

Machida R.J., Miya M.U., Nishida M., Nishida S. 2004: Large-scale gene rearrangements in the mitochondrial genomes of two calanoid copepods *Eucalanus bungii* and *Neocalanus cristatus* (Crustacea), with notes on new versatile primers for the srRNA and COI genes. Gene 332: 71–78.

Mathews P.D., Patta A.C., Gama G.S., Mertins O. 2018: Infestation by *Ergasilus coatiarus* (Copepoda: Ergasilidae) in two Amazonian cichlids with new host record from Peru: an ectoparasites natural control approach. C. R. Biol. 341: 16–19.

Minxiao W., Song S., Chaolun L., 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.

Nguyen L.T., Schmidt H.A., Von Haeseler A., Minh B.Q. 2015: IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32: 268–274.

Peng G., Gao Q., Song Y., Zhao Q., Luo Y., Nie P. 2010: Mitochondrial genes of *Sinergasilus polycolpus* (Copepoda, Ergasilidae) parasitizing the gills of fish. Acta. Hydrobiol. Sinica 34: 177–183.

Perna N.T., Kocher T.D. 1995: Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41: 353–358.

Piasecki W., Goodwin A.E., Eiras J.C., Nowak B.F. 2004: Importance of Copepoda in freshwater aquaculture. Zool. Stud. 43: 193–205.

Rand D.M. 1994: Thermal habit, metabolic rate and the evolution of mitochondrial DNA. Trends. Ecol. Evol. 9: 125–131.

Reuter J.S., Mathews D.H. 2010: RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics 11: 1–9.

Ronquist F., Teslenko M., Van Der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012: MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61: 539–542.

Shao R., Barker S. 2007: Mitochondrial genomes of parasitic arthropods: implications for studies of population genetics and evolution. Parasitology 134: 153–167.

Song Y., Wang G. T., Yao W. J., Gao Q., Nie P. 2007: Phylogeny of freshwater parasitic copepods in the Ergasilidae (Copepoda: Poecilostomatoida) based on 18S and 28S rDNA sequences. Parasitol. Res. 102: 299–306.

Staton J.L., Daehler L.L., Brown W.M. 1997: Mitochondrial gene arrangement of the horseshoe crab *Limulus polyphemus* L.: conservation of major features among arthropod classes. Mol. Biol. Evol. 14: 867–874.

Suzuki Y., Glazko G.V., Nei M. 2002: Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA 99: 16138–16143.

Talavera G., Castresana J. 2007: Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56: 564–577.

Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. 2013: MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30: 2725–2729.

Tung C.-H., Cheng Y.-R., Lin C.-Y., Ho J.-S., Kuo C.-H., Yu J.-K., 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.B., G. (Ed.) 2024: World of Copepods Database. Access year: 2024. *Ergasilus* von Nordmann, 1832., www.marinespecies.org

Wang N., Xiang Y., Fang L., Wang Y., Xin H., Li S. 2013: Patterns of gene duplication and their contribution to expansion of gene families in grapevine. Plant. Mol. Biol. Rep. 31: 852–861.

Xiang C.Y., Gao F., Jakovlić I., Lei H.P., Hu Y., Zhang H., Zou H., Wang G.T., Zhang D. 2023: Using PhyloSuite for molecular phylogeny and tree-based analyses. iMeta 2: e87.

Zhang D., Li W.X., Zou H., Wu S.G., Li M., Jakovlić I., Zhang J., Chen R., Wang G. 2019: Homoplasy or plesiomorphy? Reconstruction of the evolutionary history of mitochondrial gene order rearrangements in the subphylum Neodermata. Int. J. Parasitol. 49: 819–829.

Received 29 July 2023 Accepted 5 January 2024 Published online 7 February 2024

Cite this article as: Huang Y., Feng J.-M., Liu W., Sun B.L., Shu X.-J., Li W.-X., Wang G.T., Hua C.-J. 2024: Comprehensive molecular characterisation of the complete mitogenome of *Ergasilus tumidus* and phylogenetic relationships of Copepoda inferred from mitogenomes. Folia Parasitol. 71: 002.