The first complete mitochondrial genome of a parasitic isopod supports Epicaridea Latreille, 1825 as a suborder and reveals the less conservative genome of isopods

Article *in* Systematic Parasitology · June 2018

DOI: 10.1007/s11230-018-9792-2

CITATIONS
11

READS
303

4 authors, including:

An Jianmei
Shanxi Normal University
52 PUBLICATIONS **181** CITATIONS
SEE PROFILE

Christopher B. Boyko
Hofstra University
153 PUBLICATIONS **2,540** CITATIONS
SEE PROFILE

Syst Parasitol
https://doi.org/10.1007/s11230-018-9792-2

The first complete mitochondrial genome of a parasitic isopod supports Epicaridea Latreille, 1825 as a suborder and reveals the less conservative genome of isopods

Jialu Yu · Jianmei An · Yue Li · Christopher B. Boyko

Received: 30 July 2017 / Accepted: 14 March 2018

Abstract The complete mitochondrial genome sequence of the holoparasitic isopod *Gyge ovalis* (Shiino, 1939) has been determined. The mitogenome is 14,268 bp in length and contains 34 genes: 13 protein-coding genes, two ribosomal RNA, 19 tRNA and a control region. Three tRNA genes (*trn*E, *trn*I and *trn*S1) are missing. Most of the tRNA genes show secondary structures which derive from the usual cloverleaf pattern except for *trn*C which is characterised by the loss of the DHU-arm. Compared to the isopod ground pattern and *Eurydice pulchra* Leach, 1815 (suborder Cymothoida Wägele, 1989), the genome of *G. ovalis* shows few differences, with changes only around the control region. However, the genome of *G. ovalis* is very different from that of non-cymothoidan isopods and reveals that the gene order evolution in isopods is less conservative compared to other crustaceans. Phylogenic trees were constructed using maxiumum likelihood and Bayesian inference analyses based on 13 protein-coding genes. The results do not support the placement of *G. ovalis* with *E. pulchra* and *Bathynomus* sp. in the same suborder; rather, *G. ovalis* appears to have a closer relationship to *Ligia oceanica* (Linnaeus, 1767), but this result suggests a need for more data and further analysis. Nevertheless, these results cast doubt that Epicaridea Latreille, 1825 can be placed as an infraorder within the suborder Cymothoida, and Epicaridea appears to also deserve subordinal rank. Further development of robust phylogenetic relationships across Isopoda Latreille, 1817 will require more genetic data from a greater diversity of taxa belonging to all isopod suborders.

This article is part of the Topical Collection Arthropoda.

J. Yu · J. An (✉) · Y. Li
School of Life Science, Shanxi Normal University, Linfen 041000, People's Republic of China
e-mail: anjianmei@hotmail.com

C. B. Boyko
Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA

Introduction

Mitochondrial genomes, with their maternal inheritance, clear orthology of genes and absence of introns, are often used to analyze phylogenetic relationships of metazoans. The changes in gene order prove to be extremely reliable phylogenetic characters because the probability that homoplastic translocations occur in closely related taxa is very low (Kilpert et al., 2012). The metazoan mitochondrial (mt) genome is a circular double stranded DNA molecule of about 12–20 kb length, and typically contains 37 genes: 13 protein-coding genes, two ribosomal genes and 22 transfer RNA genes (Wolstenholme, 1992).

Isopoda Latreille, 1817, with more than 10,000 species (WoRMS, 2018), is the largest taxon within

Published online: 11 April 2018

Table 1 Isopods with complete or nearly complete mitochondrial genomes known

Species	Suborder	No. of known mt genes	Unknown or missing genes	Genome length (bp)	Reference
Asellus aquaticus (Linnaeus, 1758)	Asellota	33[a]	*trn*E, *trn*I, *trn*W, *trn*S1	13,639	Kilpert et al. (2012)
Janira maculosa Leach, 1814	Asellota	25[a]	*trn*R, *trn*E, *trn*I, *trn*W, *trn*S1, *trn*F, *trn*K, *trn*N, *atp*6, *atp*8, *nad*1, *nad*5	9,871	Kilpert et al. (2012)
Eurydice pulchra Leach, 1815	Cymothoida	31[a]	*trn*R, *trn*E, *trn*I, *trn*L2, *trn*F, *trn*S1	13,858	Kilpert et al. (2012)
Bathynomus sp.	Cymothoida	33[a]	*trn*H, *trn*I, *trn*L1, *trn*S1	14,965	Shen et al. (2017)
Gyge ovalis (Shiino, 1939)	Cymothoida	34[b]	*trn*E, *trn*I, *trn*S1	14,268	Present study
Limnoria quadripunctata Holthuis, 1949	Limnoriidea	39[b]; *trn*E, *trn*V (repeated)	0	16,515	Lloyd et al. (2015)
Armadillidium vulgare (Latreille, 1804)	Oniscidea	28[a]	*trn*R, *trn*E, *trn*I, *trn*L1, *trn*S1, *trn*L2, *trn*K, *trn*N, *trn*A	13,858	Marcadé et al. (2007)
Ligia oceanica (Linnaeus, 1767)	Oniscidea	36[b]	*trn*R	15,289	Kilpert & Podsiadlowski (2006)
Cylisticus convexus (De Geer, 1778)	Oniscidea	32[a]	*trn*E, *trn*G, *trn*F, *trn*V, *trn*L2	14,154	Chandler et al. (2015)
Trachelipus rathkii (Brandt, 1833)	Oniscidea	29[a]	*trn*T, *trn*Y, *trn*A, *trn*E, *trn*G, *trn*K, *trn*L2, *trn*S1	14,129	Chandler et al. (2015)
Eophreatoicus sp.	Phreatoicidea	37[b]	0	14,994	Kilpert & Podsiadlowski (2010)
Sphaeroma serratum (Fabricius, 1787)	Sphaeromatidea	30[a]	*trn*R, *trn*E, *trn*C, *trn*I, *trn*L1, *trn*W, *trn*S1	13,467	Kilpert et al. (2012)
Glyptonotus cf. *antarcticus* Eights, 1852	Valvifera	33[a]	*trn*E, *trn*I, *trn*W, *trn*S1	13,809	Kilpert et al. (2012)
Idotea balthica (Pallas, 1772)	Valvifera	32[a]	*trn*N, *trn*E, *trn*I, *trn*S1, *trn*W	14,247	Podsiadlowski & Bartolomaeus (2006)

[a]Incomplete. [b]Complete

Malacostraca Latreille, 1802: Peracarida Calman, 1904 and contains species that show an amazing ecological diversity and morphological plasticity. Although the mt genome sequences of 7,436 species of organisms are currently available on GenBank, a mere 435 come from arthropods, 272 of which are from crustaceans and 230 from malacostracans. Only four isopods have their complete mt genomes reported: the semi-terrestrial *Ligia oceanica* (Linnaeus, 1767); the Australian freshwater *Eophreatoicus* sp.; the wood-boring *Limnoria quadripunctata* Holthuis, 1949; and the giant isopod *Bathynomus* sp. (see Table 1). The differences in the identity and number of missing mt genes support the idea that each of these isopod species belongs to a distinct lineage as implied by their current taxonomic placement in separate suborders. An additional nine isopod species with incomplete mt genome sequences are known (Table 1). Marcadé et al. (2007) found the mt genome of *Armadillidium vulgare* (Latreille, 1804) to show the strongest differences from other isopods as this genome is *c*.42 kb long and consists of two molecules: a circular, *c*.28 kb, dimer formed by two 14 kb monomers and a linear 14 kb monomer.

Epicaridea Latreille, 1825 is a group of uniquely holoparasitic isopods that are found on crustacean hosts in both the larval and adult stages and this group currently contains 815 species (Williams & Boyko, 2016). The history of this group dates back to at least the Jurassic (Boyko et al., 2013). *Gyge ovalis* (Shiino,

1939) is a common parasite of *Upogebia* spp. (Malacostraca: Decapoda: Upogebiidae) in Asia, with records known from Japan (Shiino, 1939), Korea, Taiwan, and Shandong Province, China (An et al., 2009; see also Markham, 2004 for a taxonomic history of the species).

This parasitic mode of life, with its specialised and often results in reduced morphology, has made it difficult to determine the phylogenetic relationships of species within Epicaridea and of epicarideans within Isopoda (Shi et al., 2014). Molecular data offer an alternative approach to analyze their phylogenetic relationships. The phylogenetic relationships of the 11 suborders of isopods are highly variable (Dreyer & Wägele, 2002; Brandt & Poore, 2003; Shen et al., 2017). Epicaridea has traditionally been regarded as a distinct parasitic suborder based on morphological characters (Monod, 1922; Strömberg, 1972). However, Dreyer & Wägele (2001, 2002) considered epicarideans as derived from fish parasites in the Cymothoida Wägele, 1989 based on a combination of molecular (SSU rDNA) and morphological characters. Boyko et al. (2013) showed that Epicaridea was monophyletic with respect to Cymothooidea Leach, 1814 but did not consider the relationships of other taxa within Isopoda.

Although the cytochrome *c* oxidase subunit 1 (*cox*1) gene has been sequenced for *Bopyroides hippolytes* (Kröyer, 1838) (see Costa et al., 2007), *Orthione griffenis* Markham, 2004 (see Hong et al., 2015) and *Athelges paguri* (Rathke, 1843) (see Raupach et al., 2015), no complete mt genome of any epicaridean has yet been reported. Here we provide data on the mt genome of *G. ovalis*, the first complete mt genome of any parasitic isopod, and only the fifth complete mt genome sequence for isopods in general. We also evaluate hypotheses on the relationships between several suborders of Isopoda based on protein-coding genes.

Materials and methods

Sample and DNA extraction

A specimen of *G. ovalis* was collected by J. An and L. Yue on 13 July 2013 along with its host, *Upogebia major* (De Haan), from Yangma Island (37.45°N, 121.61°E), Yantai, Shandong Province, China. The specimen was preserved in absolute ethanol (> 99.7%) until DNA extraction (November 2014). Eggs extracted from ovigerous females were used for DNA extraction because DNA is easier to obtain from eggs than adult specimens, the quantity of DNA from eggs is higher, and there is no contamination from host DNA as is often found when DNA is extracted from adult parasites. Total genomic DNA was isolated from eggs with a genome DNA rapid extraction kit (Aidlab Biotechnologies Co., Ltd). The voucher specimen is deposited in Shanxi Normal University under the collection number CIET201307130006.

PCR primers and amplicon analysis

A complete list of primers is provided in Table 2. All primers were purchased from Sangon Biotech (Shanghai) Co., Ltd. Primers designed with DNAStar to match generally conserved regions of target mtDNA were used to amplify short fragments from 16S, 12S, *cox*1, *atp*6, *cox*3, *nad*4, *cytb*, and *nad*1. Specific primers were designed based on these conserved region sequences and used to amplify the remaining mtDNA sequence in several PCR reactions. The PCR reaction was carried out with LA Taq polymerase for 35 cycles at 94°C for 30 s, annealed at 50°C for 30 s, followed by extension at 72°C for 1min per 1 kb. The final $MgCl_2$ concentration in the PCR reaction was 2 mM. PCR products were cloned into pMD18-T vector (Takara Biomedical Technology, Beijing, China) and then sequenced, or sequenced directly by the dideoxynucleotide procedure using an ABI 3730 automatic sequencer. The reaction was carried out in a solution containing 0.5 µl TaKaRa LA*Taq* (5 U/µl), 5 µl 10× LA *Taq* Buffer II (Mg^{2+}Plus), 8 µl dNTP Mixture (2.5 mM each), 60 ng template DNA, Primer 1 (final concentration 0.2–1.0 µM), Primer 2 (final concentration 0.2–1.0 µM), and ddH_2O up to 50 µl. Every part of the mt genome was ascertained at least twice to prevent sequencing errors. Ambiguous parts were carefully inspected referring to the chromatograms and repeatedly sequenced if necessary. The PCR products were examined by agarose gel electrophoresis (1% agarose in TBE, stained with ethidium bromide) under UV transillumination. PCR amplicons were subsequently sequenced or stored at -20°C. Sequences were assembled by DNAstar software and adjusted manually to generate the complete sequence of mt DNA.

Table 2 List of PCR primers used to determine *Gyge ovalis* (Shiino, 1939) mitochondrial genome sequence

Primer	Orientation	Sequence (5′-3′)
*G. ovalis nad*5-16S	Forward	CCGCGAGCTATAGCAGCACC
	Reverse	CCTCTCCAGAGATCCTATCGG
G. ovalis 16S	Forward	GATAGAAACCAACCTGGCTC
	Reverse	CCTAACACCTCTTAGGTCCAG
G. ovalis 16S-*nad*2	Forward	GGTTTTACTCTTGACATAGCG
	Reverse	GTTATGTTACCTTTCAAGTG
*G. ovalis nad*2-*cox*1	Forward	TGCTTCCTCCCAACCCTTAG
	Reverse	GAGCAATACTAAGTGAGAG
*G. ovalis cox*1	Forward	CTCCTGATATGGCATTCCCAC
	Reverse	CCTTTACACCTGTAGGAACTGC
*G. ovalis cox*1-*cox*2	Forward	GTCTTTTGGGTACTTGGGAATGG
	Reverse	CTTCAATCAATCTAGCCTG
*G. ovalis cox*2	Forward	GTTCTAGTCAGGGTTGCCC
	Reverse	CCTCAATATCCACTCTCTC
*G. ovalis cox*2-*nad*1	Forward	GAGTTAAGGCGGATGCTATTCCC
	Reverse	CCAGACTATCTCCTATGAAG
*G. ovalis nad*1	Forward	GTCAGGCTATACATAGAATAGG
	Reverse	CTAGGGTATATTCAGATCCG
*G. ovalis nad*1-12S	Forward	AAGACAGGACCCACGATG
	Reverse	GTAGGTTTTGGAATACCGGG
G. ovalis 12S	Forward	TACACCTTGGGCAGAGATC
	Reverse	GTACTCTTACTATGTTACGAC
G. ovalis 12S-*cytb*	Forward	CAGCATAACTGGACTACACTATG
	Reverse	TGGAGCTTGTCCTGTGGAAG
G. ovalis cytb	Forward	GAAAACATATTACGGTGTG
	Reverse	CAGTAGGATTCATCATCATAC
G. ovalis cytb	Forward	GATCGTAGGATAGCATAAGC
	Reverse	AACCTGAATCTCAGGAATCC
*G. ovalis cytb-nad*5	Forward	CACTGAACCAGGTCTTGTCC
	Reverse	CCTCTAACAAAGAAGCAAACCG

Gene annotation and sequence analysis

Transfer RNAs were automatically identified by two computer programs, tRNAscan-SE 1.21 (Lowe & Eddy, 1997) and MITOS (http://mitos.bioinf.uni-leipzig.de/index.py). Gene identification was determined by searching for ORFs between tRNAs and by comparing the corresponding sequences using BLASTn and BLASTx. Protein-coding genes and RNA subunit genes were identified by BLAST search of GenBank databases and by comparison to the mt genome of *Idotea balthica* (Pallas, 1772) (GenBank: DQ442915; Podsiadlowski & Bartolomaeus, 2006). Protein-coding genes were specified by aligning them to multiple amino acid sequences of other crustaceans. The complete annotated mt genome sequence of *G. ovalis* was submitted to the GenBank database under the accession number KY038353. Nucleotide frequencies of the protein-coding genes, rRNA genes, as well as the total genome, were determined with BioEdit 7.0 (Hall, 1999).

Phylogenetic analysis of mt gene sequences

The twelve available isopod species with determined mtDNA sequences were used for phylogenetic analyses, except for *A. vulgare* for which the NADH dehydrogenase subunit translation

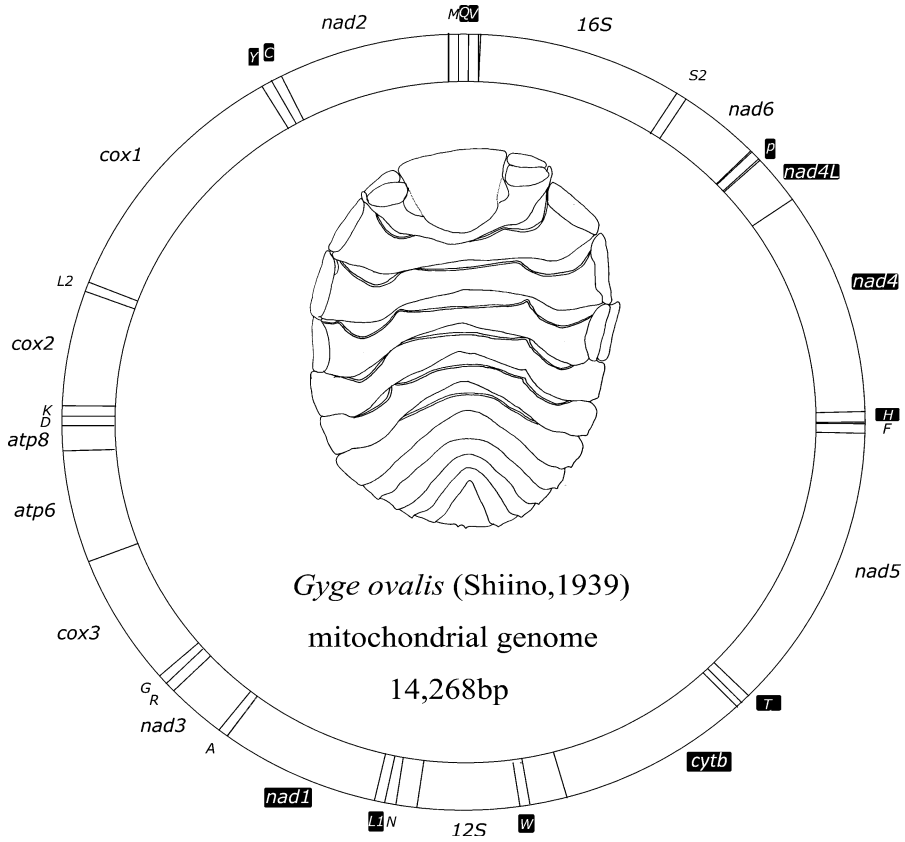

Fig. 1 Map of the mitochondrial genome of *Gyge ovalis* (Shiino, 1939). Transfer RNAs are represented by their one-letter amino acid codes. Inverted (white on black) legends indicate genes located on (–) strand

mechanism is unknown. Following Shen et al. (2017), we used six decapod species as outgroups: *Austinograea rodriguezensis* Tsuchida & Hashimoto, 2002, *Alvinocaris longirostris* Kikuchi & Ohta, 1995, *Geothelphusa dehaani* (White, 1847), *Halocaridina rubra* (Holthuis, 1963), *Panulirus japonicus* (von Siebold, 1824) and *Shinkaia crosnieri* Baba & Williams, 1998. Nucleotide sequence alignments were built for all 13 protein-coding genes individually using MEGA 5.1 (Tamura et al., 2004); they were then concatenated to a single alignment. We used Modeltest (Posada & Crandall, 1998) and PhyML 3.0 (Guindon & Gascuel, 2003) webserver to determine the optimal nucleotide sequence model for phylogenetic analysis (evaluated according to the Akaike information criterion). The GTR + G + I (general time reversible model with a proportion of invariant sites and gamma-distributed rate variation across sites) model was found to be most appropriate.

The Maximum likelihood analysis was performed used the PhyML 3.0 webserver, and a bootstrap test was used to analyze the repeated test 1,000 times.

The Bayesian inference analysis, assuming a fixed nucleotide sequence GTR + G + I model, was performed using MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003). The starting tree was randomly chosen and four Monte Carlo Markov chains were run for 10,000,000 generations with trees sampled every 1,000 generations. The initial 25% of the trees were discarded as 'burn-in'. The best tree and 100 bootstrap replicates were analyzed.

Results

The complete mt genome of *G. ovalis* (~~KY038353~~) has a size of 14,268 bp (Fig. 1, Table 3). It is most likely organised in a single circular DNA molecule. The

Table 3 Gene content of the complete mitochondrial genome of *Gyge ovalis* (Shiino, 1939)

Gene	Strand	GenBank position	Size	Start codon	Stop codon	Intergenic nucleotides
tRNA-His	−	1–64	64			0
*nad*4	−	65–1,394	1,330	A	CAT	-7
*nad*4l	−	1,388–1,684	297	TTA	TAT	9
tRNA-Pro	−	1,694–1,752	59			2
*nad*6	+	1,755–2,237	483	ATA	TAG	-2
tRNA-Ser(UCN)	+	2,236–2,297	62			0
16S	+	2,298–3,471	1,174			0
tRNA-Val	−	3,472–3,532	61			0
tRNA-Gln	−	3,533–3,590	58			6
tRNA-Met	+	3,597–3,659	63			0
*nad*2	+	3,660–4,638	979	ATT	T	2
tRNA-Tyr	−	4,641–4,694	54			0
tRNA-Cys	−	4,695–4,754	60			0
*cox*1	+	4,755–6,294	1,540	ATT	T	0
tRNA-Leu(UUR)	+	6,295–6,354	60			0
*cox*2	+	6,355–7,033	679	GTG	T	0
tRNA-Lys	+	7,034–7,094	61			-2
tRNA-Asp	+	7,03–7,151	59			0
*atp*8	+	7,152–7,302	151	TTG	T	-12
*atp*6	+	7,291–7,962	672	ATG	TAA	-1
*cox*3	+	7,962–8,745	784	ATG	T	0
tRNA-Gly	+	8,746–8,802	57			0
tRNA-Arg	+	8,803–8,861	59			0
*nad*3	+	8,862–9,210	349	ATA	T	0
tRNA-Ala	+	9,211–9,269	59			3
*nad*1	−	9,273–10,194	922	A	AAT	0
tRNA-Leu(CUN)	−	10,195–10,255	61			-4
tRNA-Asn	+	10,252–10,314	63			0
CR1	+	10,315–10,435	121			0
12S		10,436–11,019	584			0
tRNA-Trp	−	11,020–11,079	60			0
CR2		11,080–11,290	211			0
cytb	−	11,291–12,403	1,113	TTA	TAT	33
tRNA-Thr	−	12,437–12,494	58			1
*nad*5	+	12,496–14,209	1,714	TAT	T	0
tRNA-Phe	+	14,210–14,268	59			

sequence annotation revealed 34 mt genes (13 protein-coding genes, 19 tRNA genes and 2 rRNA genes).

Nineteen tRNA genes of the mt genome of *G. ovalis* were determined, and secondary structures in non-coding regions of the genes *trn*E, *trn*I and *trn*S1 were not found. Among the 19 genes, 15 were identified using tRNA scan-SE 1.21, and the other four tRNA genes (*trn*C, *trn*R, *trn*N and *trn*F) were found with the online software MITOS. These tRNA genes are distributed throughout the genome and can be found on both strands (Table 4). The putative secondary structures of all identified tRNAs are

Table 4 Nucleotide frequencies, AT content, AT and GC skew for mitochondrial genes of *Gyge ovalis* (Shiino, 1939)

Gene (strand)	Nucleotide frequencies				% AT	ATskew	GCskew	+/− strand	+/− strand
	A	C	G	T				*E. pulchra*	Isopoda ground pattern
*atp*6 (+)	0.222	0.210	0.228	0.341	0.563	-0.211	0.041	+	+
*atp*8 (+)	0.265	0.219	0.152	0.364	0.639	-0.162	-0.181	+	+
*cox*1 (+)	0.226	0.190	0.229	0.356	0.582	-0.223	0.093	−	+
*cox*2 (+)	0.236	0.175	0.260	0.330	0.566	-0.166	0.195	+	+
*cox*3 (+)	0.228	0.193	0.240	0.339	0.568	-0.196	0.109	+	+
cytb (−)	0.306	0.165	0.262	0.266	0.572	0.070	0.227	−	−
*nad*1 (−)	0.341	0.215	0.196	0.248	0.589	0.158	-0.046	−	−
*nad*2 (+)	0.252	0.144	0.252	0.351	0.604	-0.164	0.273	−	+
*nad*3 (+)	0.218	0.209	0.235	0.338	0.556	-0.270	0.059	+	+
*nad*4 (−)	0.346	0.168	0.222	0.264	0.610	0.134	0.138	−	−
*nad*4l (−)	0.350	0.185	0.179	0.286	0.636	0.101	-0.016	−	−
*nad*5 (+)	0.222	0.153	0.247	0.378	0.600	-0.260	0.235	+	+
*nad*6 (+)	0.201	0.164	0.259	0.377	0.578	-0.304	0.225	+	+
*rrn*L (+)	0.263	0.186	0.193	0.359	0.622	-0.162	0.018	−	−
*rrn*S (+)	0.307	0.194	0.241	0.259	0.565	0.085	0.108	+	+
Total	0.266	0.185	0.226	0.324	0.590	-0.105	0.099		

shown in Fig. 2. The majority of tRNAs have a common t-shaped or clover-leaf secondary structure. In tRNA-Cys, this pattern is modified and the DHU-arm is absent.

Nucleotide frequencies of all mt genes of *G. ovalis* are listed in Table 4. The average A+T content of the protein-coding genes of the *G. ovalis* mt genome is approximately 59%.

The gene order of the mt monomer of *G. ovalis* is shown in Fig. 3. There is one major non-coding region of 211 bp length located between *trn*W and *cytb*. We detected a hairpin structure in the mt control region of *G. ovalis* (Fig. 4).

The topological structures of the trees generated by maximum likelihood and Bayesian inference analysis were almost identical (Figs. 5, 6). *Gyge ovalis* formed a clade with *L. oceanica* (Oniscidea) that does not group with the other oniscidean taxa. *Bathynomus* sp. (Cymothoida) and *S. serratum* (Sphaeromatidea) formed one clade and two species of Oniscidea Latreille, 1802, *Cylisticus convexus* (De Geer, 1778) and *Trachelipus rathkii* (Brandt, 1833), also formed a clade. Valvifera G. O. Sars, 1883 was well supported, represented here by *G.* cf. *antarcticus* and *I. balthica*.

Discussion

Genome organization

Compared to the standard isopod genome, three partial tRNA genes (*trn*E, *trn*I and *trn*S1) are missing in *G. ovalis*, which is different from the isopod mt genomes of *Eophreatoicus* sp. and *L. quadripunctata*. Kilpert & Podsiadlowski (2006) reported that the *trn*R of the *L. oceanica* mt genome was missing, and Marcadé et al. (2007) found that nine tRNA genes (*trn*A, *trn*N, *trn*R, *trn*E, *trn*I, *trn*L1, *trn*L2, *trn*K and *trn*S1) of the *A. vulgare* mt genome were missing. It therefore appears that the loss of one or more tRNA genes is a common occurrence in isopods. Shao et al. (2009) showed that in sucking lice (Insecta: Anoplura) the mt genome is split into smaller minichromosomes, but this phenomenon has not been demonstrated in crustaceans nor is there such evidence in the present study. Another hypothesis for the loss of tRNA genes was proposed by Raimond et al. (1999), who suggested that some tRNA may be organised in linear fragments but none have been found in *G. ovalis*.

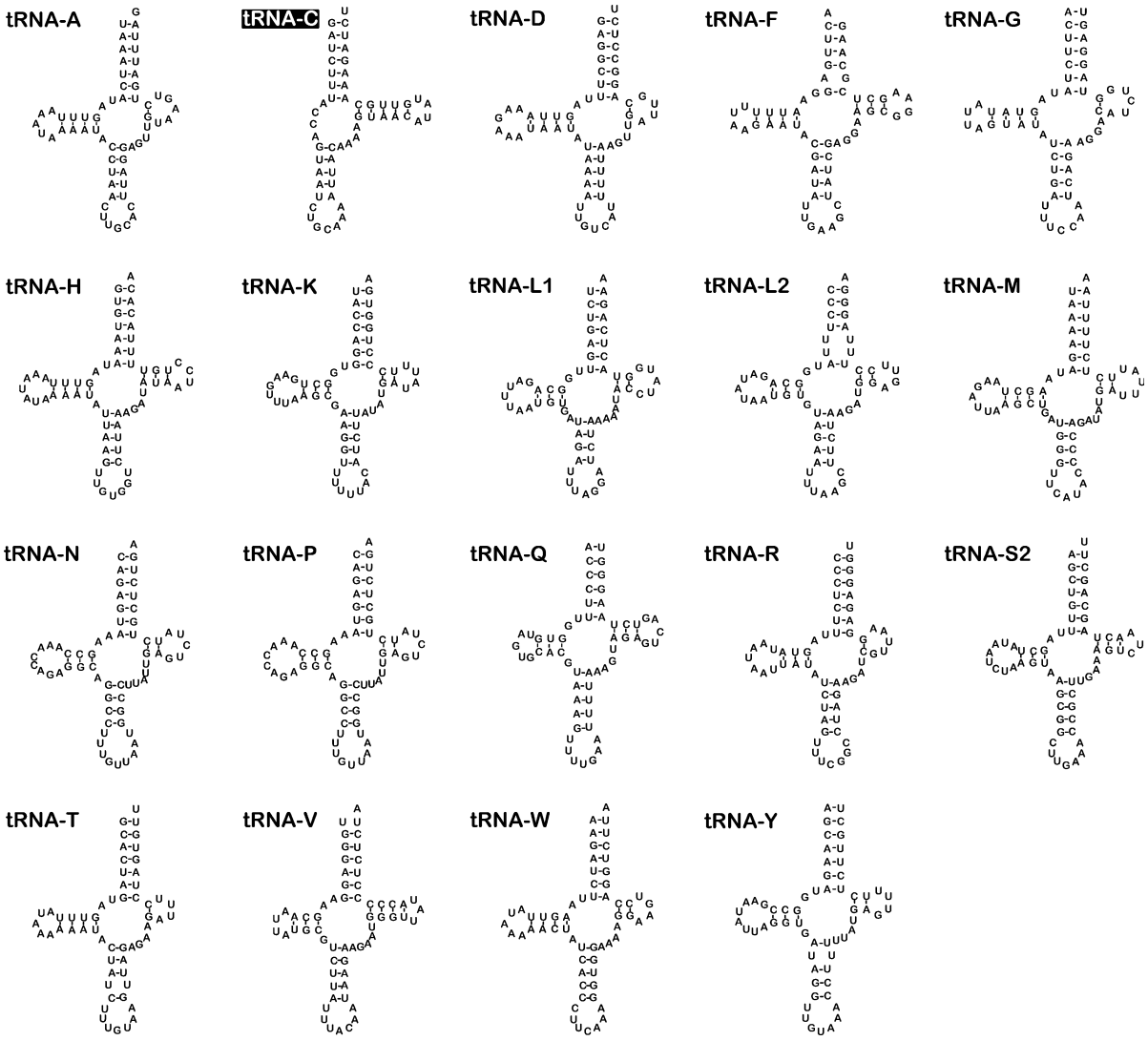

Fig. 2 Putative secondary structure of the mitochondrial tRNAs in *Gyge ovalis* (Shiino, 1939). Most tRNAs feature a standard clover-leaf structure. Exception: DHU-arm is non-existent in tRNA-C

Transfer RNAs

Modified secondary structures that are all characterised by the loss of the DHU-arm were also found in tRNA-Cys of *Eophreatoicus* sp., *L. oceanica*, *Glyptonotus* cf. *antarcticus* Eights, 1852, *A. vulgare*, *Janira maculosa* Leach, 1814 and *Asellus aquaticus* (Linnaeus, 1758). Kilpert & Podsiadlowski (2010) stated that the missing DHU-arm in tRNA-Cys might be a putative autapomorphy of Isopoda. It appears that tRNA-Cys is less conservative than other tRNAs.

Protein-coding genes

The highest A+T content value reported to date from any isopod is 71.4% for *A. vulgare*, while the lowest A+T value, 54.4%, was reported from *Sphaeroma serratum* (Fabricius, 1787) (see Kilpert et al., 2012). The A+T content of *G. ovalis* is therefore on the relatively low end of the range. Hassanin (2006) thought that a bias of nucleotide frequencies is generally found in the mt genome of arthropods, and the (+) strand contains more cytosine and adenine,

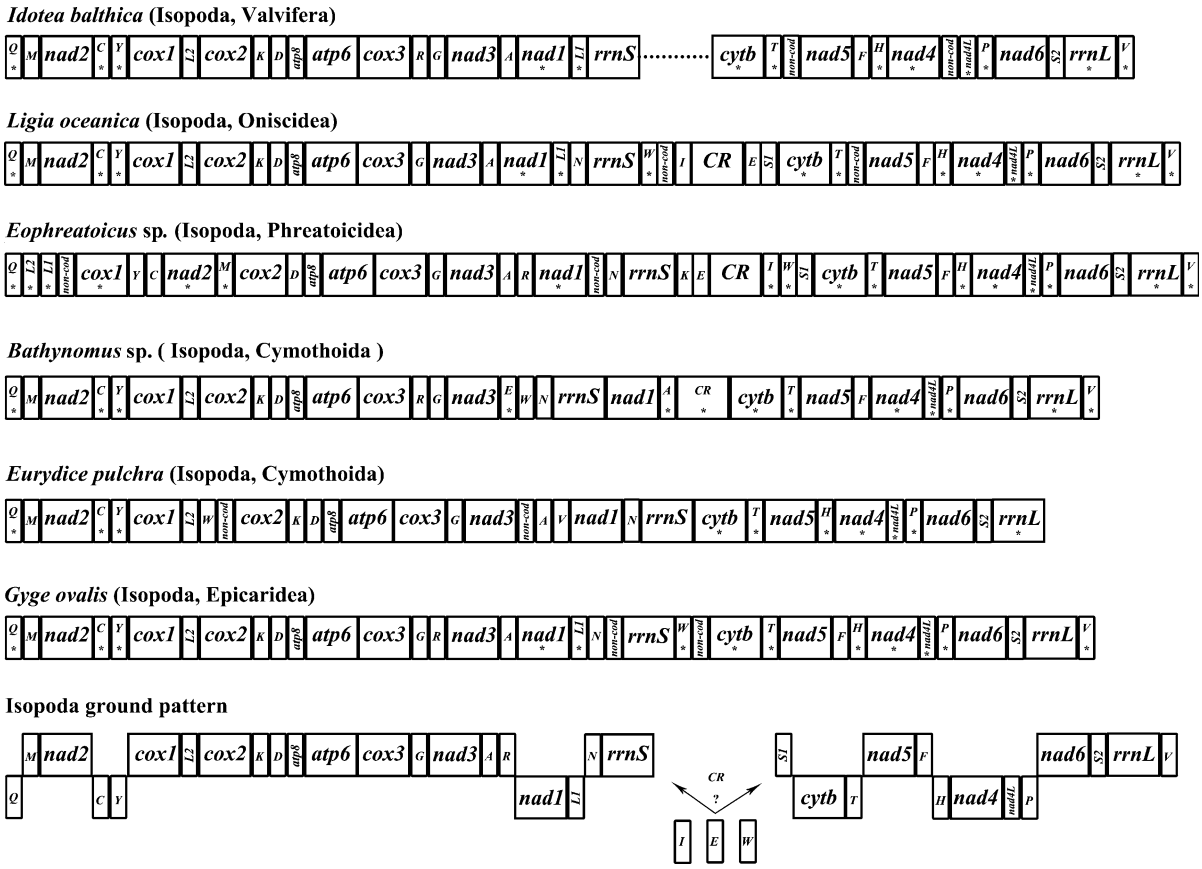

Fig. 3 Schematic diagram of the mitochondrial gene order in *Gyge ovalis* (Shiino, 1939), relative to other available isopod mitochondrial genomes and the Isopoda ground pattern

while the (−) strand consequently is richer in guanine and thymine. Nevertheless, most isopod species show a clearly positive GC skew for the (+) strand genes and a negative one for the (−) strand genes. This is in contrast to most other malacostracan crustaceans and is probably due to an inversion of the mt control region, or at least of the replication origin. However, *G. ovalis* does not fully follow the positive GC skew for (+) strand genes and negative GC skew for (−) strand genes as, for example, the GC skew for the (+) strand gene *atp*8 was -0.181, whereas the (−) strand genes *cytb* and *nad*4 have a GC skew of 0.227 and 0.138, respectively. Other isopod species have different values as, for example, there is no clear bias in *G.* cf. *antarcticus* or *J. maculosa*, where the GC skews are nearly 0. However, a noticeable bias was detected for *A. aquaticus*, where GC skews are negative for (+) strand genes and positive for (−) strand genes. This is in clear contrast to the situation in the majority of examined species of isopods (Kilpert et al., 2012). In many cases gene boundaries of protein-coding genes could be easily identified, as they conform to the ordinary start codons (ATG, ATA and ATT) and stop codons (TAA and TAG) of arthropod mt genomes. In addition, the *cox*2 gene start codon is GTG and the *atp*8 gene start codon is TTG; the two gene start codons are relatively rare but are typical invertebrate codons. Six genes (*nad*4, *nad*2, *cox*1, *cox*2, *atp*8, *nad*5) terminate with T. All these genes are typical for invertebrate mt DNA (Sielaff et al., 2016; Wolstenholme, 1992). The unfinished T codon was not counted separately, as we presumed that it would be completed (TAA) by posttranscriptional polyadenylation (Ojala et al., 1981; Schuster & Stern, 2009).

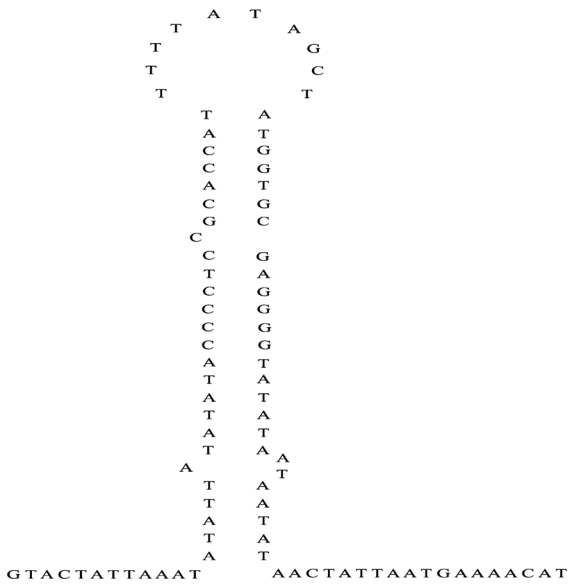

Fig. 4 Hairpin structure in the mitochondrial control region of *Gyge ovalis* (Shiino, 1939)

Gene translocations

Comparison with the putative isopod ground pattern mt genome (Kilpert et al., 2012) revealed several gene rearrangements in *G. ovalis*. Between *trn*G and *trn*A, the genes are found as follows: *trn*G, *trn*R, *nad*3, *trn*A. The gene order in the isopod ground pattern is *trn*G, *nad*3, *trn*A and *trn*R. Compared with *I. balthica*, the order in *G. ovalis* is different for some tRNA gene loci, but *trn*G and *trn*R always retain the same relative position because they are replicated together. *Eurydice pulchra* Leach, 1815, the closest related species to *G. ovalis* in most analyses, shows essentially the same gene order, in addition to a partial tRNA being missing or translocated. *Trn*V is located between *trn*A and *nad*1 in *E. pulchra* whereas in *G. ovalis* it is between *rrn*L and *trn*Q; *trn*F is missing in *E. pulchra* whereas in *G. ovalis trn*F maintains the same position relative to *nad*5; all known isopods are identical for this character. However, there are major changes in the

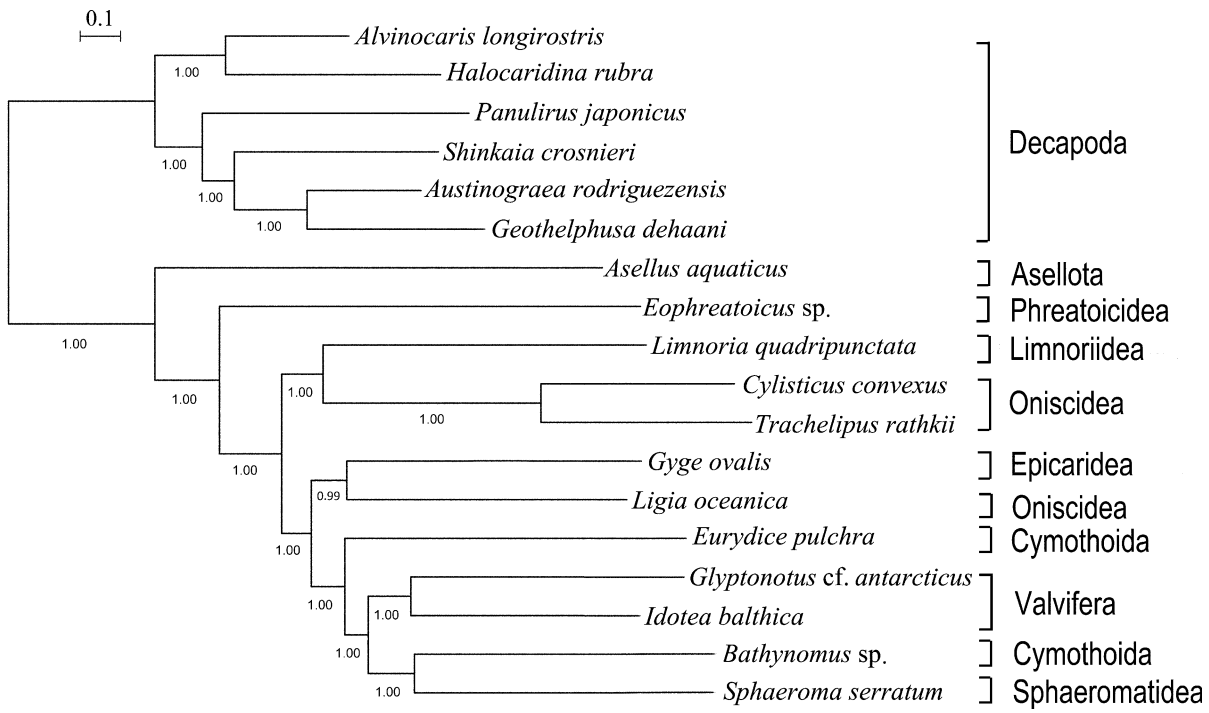

Fig. 5 Bayesian analysis tree showing relationships among isopods based on nucleotide data for 13 protein-coding genes. Six decapod species served as the outgroup. Numbers next to nodes indicate bootstrap support. The scale-bar indicates the number of substitutions per site

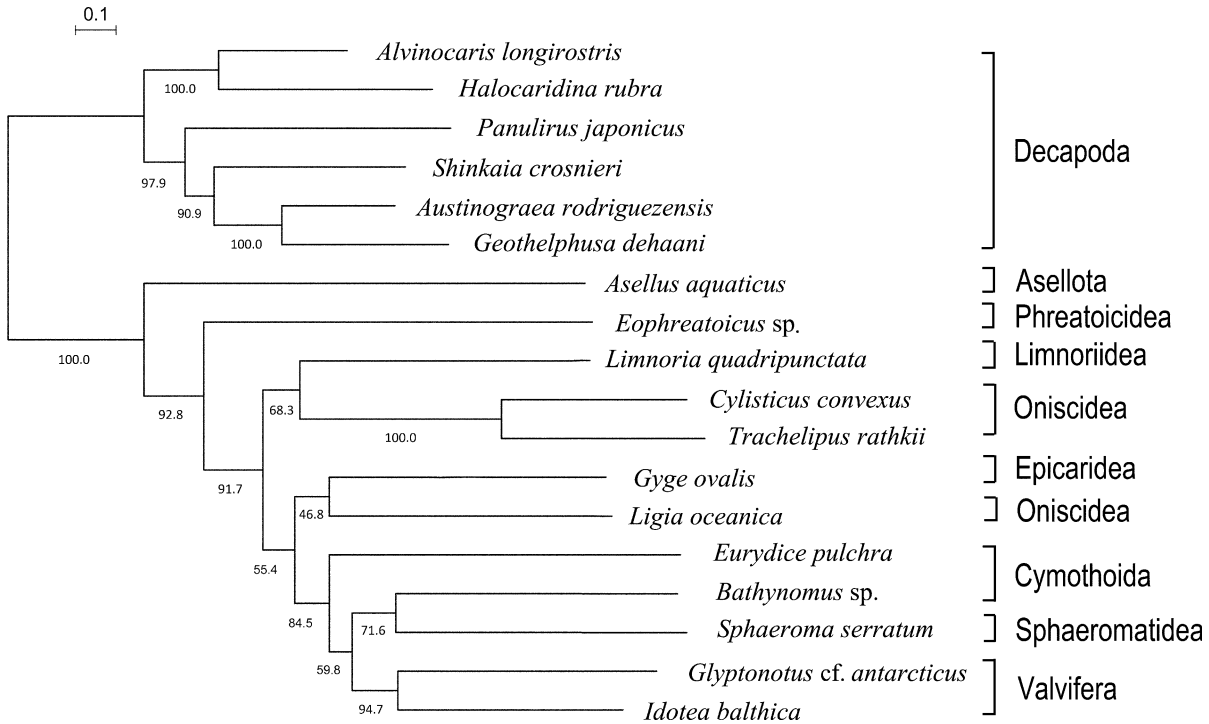

Fig. 6 Maximum likelihood tree showing relationships among isopods based on nucleotide data for 13 protein-coding genes. Six decapod species served as the outgroup. Numbers next to nodes indicate bootstrap support. The scale-bar indicates number of substitutions per site

gene arrangement of *G. ovalis* compared to the gene order of other isopods, including the translocation of protein-coding genes. Compared to *Eophreatoicus* sp., the gene arrangement of *G. ovalis* between *trn*Q and *cox*2 is remarkable (Fig. 3). The gene order of *Eophreatoicus* sp. is *trn*Q, L2, L1, *cox*1, *trn*Y, *trn*C, *nad*2, *trn*M, *cox*2, but that of *G. ovalis* is *trn*Q, *trn*M, *nad*2, *trn*C, *trn*Y, *cox*1, L2, *cox*2. Furthermore, for *Eophreatoicus* sp. it is very likely that, after *trn*L(UUR) was duplicated, one tRNA gene changed its identity to *trn*L(CUN) *via* random point mutation of the anticodon triplet. Subsequently, the original tRNA gene degraded, so that the *trn*L(CUN) gene is now located next to *trn*L(UUR). This phenomenon, however, is not found in *G. ovalis.*

Control region

The non-coding region of 211 bp in *G. ovalis* is assumed to be the mt control region, probably containing the origins of replication and regulatory elements for transcription. It seems to be relatively short in length, compared to the control region of other isopods. The hairpin structure detected is thought to be involved in the replication process (Zhang et al., 1995).

Phylogenetic analyses

The topological structures for both the ML and Bayesian trees (Figs. 5, 6) are, with the exception for the position of *G. ovalis*, essentially the same as the ML and Bayesian trees recovered by Shen et al. (2017). Shen et al. (2017) found that *L. oceanica* did not group with other species of Oniscidea. They also showed that *L. quadripunctata* (which they mistakenly labeled as a member of Cymothoida) did not form a clade with *E. pulchra* and *Bathynomus* sp. but still grouped with three of the four oniscidean taxa (minus *L. oceanica*). This result is also shown in our analyses, although we did not use the sequence for *A. vulgare* for the reasons given previously. Kilpert et al. (2012) found *E. pulchra* and *S. serratum* formed a clade, while *G.* cf. *antarcticus* and *I. balthica* also formed a

clade. It is clear from these results that the phylogeny of isopods, as inferred from mitogenomic data, is still in a state of flux. However, several results are apparent. First, *L. oceanica* consistently does not group with the other oniscidean taxa and this suggests that there is either something amiss with the *L. oceanica* sequence or there is difficulty in aligning the sequence. It would be good to replicate the mt data for *L. oceanica* to determine if the original sequence contains problems. Additionally, based on the present study, the monophyly of Cymothoida need to be reconsidered as the *G. ovalis* sequence does not appear inside any clade containing other cymothoid taxa. Therefore, we suggest that Epicaridea may be better placed as a suborder within Isopoda, not as an infraorder within Cymothoida. Our results do not support Dreyer & Wägel's (2001) hypothesis that the parasitic Epicaridea (which they called Bopyridae) and Cymothoidae are sister groups within Cymothoida.

Concluding remarks

Epicaridea and Cymothoidae are often regarded as closely related taxa using 18S SSU rDNA data (Dreyer & Wägele, 2002) and external morphological characters (Brandt & Poore, 2003). Dreyer & Wägele (2001) hypothesised that the decapod-parasitising bopyrids (epicarideans) evolved from species that parasitised fishes (Cymothoidae) and should be considered a family-level taxon (Bopyridae) within the superfamily Cymothooidea. Brandt & Poore (2003) thought the Epicaridea should be reduced to two superfamilies, the Bopyroidea and Cryptoniscoidea, within the suborder Cymothoida. Boyko et al. (2013) used 18S rDNA to study the relationships of epicarideans and their data supported retaining the Epicaridea as a taxon within the Cymothoida but distinct from the Cymothooidea.

In the present study, the genome of *G. ovalis*, a member of the Epicaridea, shows missing *trn*I and *trn*S1 genes; this is also found in *E. pulchra* (Cymothooidea) although the genome of the latter species is not complete. However, comparison with the other four isopods having known complete mt genomes shows that the genome of *G. ovalis* has changes in gene arrangement and translocation of protein-coding genes compared to *E. pulchra*. The present study shows that epicarideans may be closely related to Cymothooidea but there are not yet enough mt data to favor the hypothesis of Dreyer & Wägele (2001) over that of Brandt & Poore (2003) or *vice versa*. Indeed, our ML and Bayesian inference analyses show a lack of clarity as to the relationships between the Epicaridea, Cymothooidea, Valvifera and Spharomatidea, as well as possible issues with the placement of *Ligia* Fabricius, 1798. Our analyses indicate that *G. ovalis* is not part of the same suborder as *E. pulchra* and *Bathynomus* sp. More data are needed to determine if this phylogenetic pattern may be influenced by the apparently less conservative gene order evolution of isopods compared to other crustaceans. More complete mt genome sequences for isopods, and an effort to generate repeat datasets for those taxa that appear in the tree in places that defy conventional explanations (e.g. *Ligia*) will also clarify the relationships of the isopod suborders.

Acknowledgements Our thanks to two anonymous reviewers whose comments greatly improved the manuscript.

Funding This study was supported by the National Science Foundation of China (No. 31471970) and Fund Program for the Scientific Activities of Selected Returned Overseas Professionals in Shanxi Province (2016).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All applicable institutional, national and international guidelines for the care and use of animals were followed.

References

An, J., Williams, J. D., & Yu, H. (2009). The Bopyridae (Crustacea: Isopoda) parasitic on thalassinideans (Crustacea: Decapoda) from China. *Proceedings of the Biological Society of Washington, 122,* 225–246.

Boyko, C. B., Moss, J., Williams, J. D., & Shields, J. D. (2013). A molecular phylogeny of Bopyroidea and Cryptoniscoidea (Crustacea: Isopoda). *Systematics & Biodiversity, 11,* 495–506.

Brandt, A., & Poore, G. C. B. (2003). Higher classification of the flabelliferan and related Isopoda based on a reappraisal of relationships. *Invertebrate Systematics, 17,* 893–923.

Chandler, C. H., Badawi, M., Moumen, B., Grève, P., & Cordaux, R. (2015). Multiple conserved heteroplasmic sites in tRNA genes in the mitochondrial genomes of terrestrial isopods (Oniscidea). *G3: Genes, Genomes, Genetics*, *5*, 1317–1322.

Costa, F. O., Dewaard, J. R., Boutillier, J., Ratnasingham, S., Dooh, R. T., Hajibabaei, M., et al. (2007). Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries & Aquatic Sciences, 64,* 272–295.

Dreyer, H., & Wägele, J. W. (2001). Parasites of crustaceans (Isopoda: Bopyridae) evolved from fish parasites: Molecular and morphological evidence. *Zoology, 103,* 157–178.

Dreyer, H., & Wägele, J. W. (2002). The Scutocoxifera tax. nov. and the information content of nuclear ssu rDNA sequences for reconstruction of isopod phylogeny (Peracarida: Isopoda). *Journal of Crustacean Biology, 22,* 217–234.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology, 52,* 696–704.

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series, 41,* 95–98.

Hassanin, A. (2006). Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. *Molecular Phylogenetics & Evolution, 38,* 100–116.

Hong, J. S., Lee, C. L., & Min, G. S. (2015). Occurrence of *Orthione griffenis* Markham, 2004 (Bopyridae: Isopoda: Crustacea) parasite of the mud shrimp (*Upogebia major*) in Korean waters and its ecological implications toward marine bioinvasion. *Journal of Crustacean Biology, 35,* 605–615.

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.

Kilpert, F., & Podsiadlowski, L. (2010). The Australian freshwater isopod (Phreatoicidea: Isopoda) allows insights into the early mitogenomic evolution of isopods. *Comparative Biochemistry & Physiology Part D Genomics & Proteomics, 5,* 36–44.

Kilpert, F., Held, C., & Podsiadlowski, L. (2012). Multiple rearrangements in mitochondrial genomes of Isopoda and phylogenetic implications. *Molecular Phylogenetics & Evolution, 64,* 106–117.

Lloyd, R. E., Streeter, S. D., Foster, P. G., Littlewood, D. T. J., Huntley, J., Beckham, G. T., et al. (2015). The complete mitochondrial genome of *Limnoria quadripunctata* Holthuis (Isopoda: Limnoriidae). *Mitochondrial DNA, 26,* 825–826.

Lowe, T. M., & Eddy, S. R. (1997). TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research, 25,* 955–964.

Marcadé, I., Cordaux, R., Doublet, V., Debenest, C., Bouchon, D., & Raimond, R. (2007). Structure and evolution of the atypical mitochondrial genome of *Armadillidium vulgare* (Isopoda: Crustacea). *Journal of Molecular Evolution, 65,* 651–659.

Markham, J. C. (2004). New species and records of Bopyridae (Crustacea: Isopoda) infesting species of the genus *Upogebia* (Crustacea: Decapoda: Upogebiidae): the genera *Orthione* Markham, 1988, and *Gyge* Cornalia & Panceri, 1861. *Proceedings of the Biological Society of Washington, 117,* 186–198.

Monod, T. (1922). Sur un essai de classification rationelle des Isopodes. *Bulletin de la Société Zoologique de Paris, 47,* 134–140.

Ojala, D., Montoya, J., & Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. *Nature, 290,* 470–474.

Podsiadlowski, L., & Bartolomaeus, T. (2006). Major rearrangements characterize the mitochondrial genome of the isopod *Idotea balthica* (Crustacea: Peracarida). *Molecular Phylogenetics & Evolution, 40,* 893–899.

Posada, D., & Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics, 14,* 817–818.

Raimond, R., Marcadé, I., Bouchon, D., Rigaud, T., Bossy, J. P., & Souty-Grosset, C. (1999). Organization of the large mitochondrial genome in the isopod *Armadillidium vulgare*. *Genetics, 151,* 203–210.

Raupach, M. J., Barco, A., Steinke, D., Beermann, J., Laakmann, S., Mohrbeck, I., et al. (2015). The application of DNA barcodes for the identification of marine crustaceans from the North Sea and adjacent regions. *PLoS One, 10,* e0139421.

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19,* 1572–1574.

Schuster, G., & Stern, D. (2009). RNA polyadenylation and decay in mitochondria and chloroplasts. *Progress in Molecular Biology & Translational Science, 85,* 393–422.

Shao, R. F., Kirkness, E. F., & Barker, S. C. (2009). The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. *Genome Research, 19,* 904–912.

Shen, Y., Kou, Q., Zhong, Z., Li, X., He, L., He, S., et al. (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 & Evolution, 7,* 1869–1881.

Shi, E., Zhang, X., & An, J. (2014). Cladistic analysis of the phylogenetic relationships among species of family Bopyridae (Crustacea, Isopoda) based on morphological characters. *Sichuan Journal of Zoology, 33,* 665–672.

Shiino, S. M. (1939). Bopyrids from Kyûsyû and Ryûkyû. *Records of Oceanographic Works in Japan, 10,* 79–99.

Sielaff, M., Schmidt, H., Struck, T. H., Rosenkranz, D., Welch, D. B. M., Hankeln, T., et al. (2016). Phylogeny of Syndermata (syn. Rotifera): Mitochondrial gene order verifies epizoic Seisonidea as sister to endoparasitic Acanthocephala within monophyletic Hemirotifera. *Molecular Phylogenetics & Evolution, 96,* 79–92.

Strömberg, J.-O. (1972). Isopod phylogeny. Aspects based on embryological, morphological and palaeontological evidence. *Contributions of the Zoological Institute of the University of Lund, 1972,* 1–112.

Tamura, K., Nei, M., & Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 11030–11035.

Willams, J. D., & Boyko, C. B. (2016). Introduction to the Symposium. Parasites and pests in motion: Biology, biodiversity and climate change. *Integrative & Comparative Biology, 56,* 556–560.

Wolstenholme, D. R. (1992). Animal mitochondrial DNA: structure and evolution. *International Review of Cytology, 141,* 173–216.

WoRMS Editorial Board (2018). World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 5 March 2018.

Zhang, D. X., Szymura, J. M., & Hewitt, G. M. (1995). Evolution and structural conservation of the control region of insect mitochondrial DNA. *Journal of Molecular Evolution, 40,* 382–391.