The complete mitochondrial genome of *Cymothoa indica* has a highly rearranged gene order and clusters at the very base of the Isopoda clade

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

1 Key Laboratory of Aquaculture Disease Control, Ministry of Agriculture, and State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, P. R. China, 2 Bio-Transduction Lab, Biolake, Wuhan, P. R. China, 3 University of Chinese Academy of Sciences, Beijing, P. R. China, 4 Department of Zoology, College of Science, King Saud University, Riyadh, Saudi Arabia, 5 Department of Zoology, GC University, Faisalabad, Pakistan

* gtwang@ihb.ac.cn

Abstract

As a result of great diversity in life histories and a large number of described species, taxonomic and phylogenetic uncertainty permeates the entire crustacean order of Isopoda. Large molecular datasets capable of providing sufficiently high phylogenetic resolution, such as mitochondrial genomes (mitogenomes), are needed to infer their evolutionary history with confidence, but isopod mitogenomes remain remarkably poorly represented in public databases. We sequenced the complete mitogenome of *Cymothoa indica*, a species belonging to a family from which no mitochondrial genome was sequenced yet, Cymothoidae. The mitogenome (circular, 14484 bp, A+T = 63.8%) is highly compact, appears to be missing two tRNA genes (*trnI* and *trnE*), and exhibits a unique gene order with a large number of rearrangements. High compactness and the existence of palindromes indicate that the mechanism behind these rearrangements might be associated with linearization events in its evolutionary history, similar to those proposed for isopods from the *Armadillidium* genus (Oniscidea). Isopods might present an important model system to study the proposed discontinuity in the dynamics of mitochondrial genomic architecture evolution. Phylogenetic analyses (Bayesian Inference and Maximum Likelihood) conducted using nucleotide sequences of all mitochondrial genes resolved Oniscidea and Cymothoida suborders as paraphyletic. *Cymothoa indica* was resolved as a sister group (basal) to all remaining isopods, which challenges the accepted isopod phylogeny, where Cymothoida are the most derived, and Phreatoicidea the most basal isopod group. There is growing evidence that Cymothoida suborder might be split into two evolutionary distant clades, with parasitic species being the most basal split in the Isopoda clade, but a much larger amount of molecular resources carrying a high phylogenetic resolution will be needed to infer the remarkably complex evolutionary history of this group of animals with confidence.

Introduction

Isopoda is an exceptionally speciose (>10,000) order of crustaceans that mostly but not exclusively inhabit aquatic habitats [1,2]. As a result of great diversity in life histories and a large number of described species, taxonomic and phylogenetic uncertainty permeate the entire order [1,3–7]. As studies relying on morphology and small molecular datasets (such as single gene datasets) did not manage to resolve their phylogeny, this is an indication that phylogenetic resolution provided by the commonly used markers (such as *cox1* or *18S*) is too low to unequivocally resolve the evolutionary history of Isopoda. Therefore, availability of molecular resources carrying a high phylogenetic resolution is indispensable for identification and evolutionary history studies of Isopoda.

Mitochondrial genomes (mitogenomes) carry a large amount of data, which makes them capable of providing much higher resolution than traditionally used morphological and (single-gene) molecular markers, so mitochondrial phylogenomics is increasingly used to address phylogenetic and taxonomic controversies [8–10]. Furthermore, isopod mitogenomes generally exhibit a large number of gene order rearrangements [4,11], and some groups of isopods even possess a non-standard, linearised, mitogenome organisation and unique tRNA-encoding mechanisms [12–16]. As the evolution of mitogenomic architecture appears to be highly discontinuous [17,18], with some major animal taxa exhibiting a highly conserved mitochondrial architecture (most vertebrates being a good example), and other taxa exhibiting a rapidly-evolving architecture [17–21], we hypothesise that isopods might present an important model system to study the complex dynamics of the evolution of mitochondrial genomic architecture. However, isopod mitogenomes remain remarkably poorly represented in the GenBank, with only five complete and 19 partial mitogenomes currently (Apr, 2018) available for the entire order (20 species in total), which presents a major obstacle to their application.

Among the non-represented taxa (taxa from which no mitochondrial genome was sequenced yet) is the entire large (≈366 species and ≈42 genera [22]) family Cymothoidae Leach 1818 (suborder Cymothoida, superfamily Cymothooidea). Species belonging to this family are largely obligate parasites of fishes that feed on host tissues and fluids at least at some stage of their life [1,22,23]. Cymothoid isopods are mostly protandrous hermaphrodites that have a biphasic life cycle: after the free-swimming micropredatory stage, they attach permanently to fishes (or other crustaceans), upon which they change sex and morphology [24]. They exhibit a range of parasitic feeding strategies: on the external body surfaces, in the buccal and opercular cavities, or burrowing into the muscle of their fish hosts [1,22]. The buccal mode sometimes results in the intriguing phenomenon of parasitic castration [25].

Identification and taxonomy of cymothoid isopods are complicated by a number of factors, including morphological similarity, sequential hermaphroditism, sexual dimorphism (females up to three times larger than males), flexible host preference almost completely unrelated to phylogeny, habitat flexibility (sea, brackish and fresh water), global distribution of many species, etc. Along with limited molecular resources currently available, this causes frequent incorrect identifications and misuse of species names, so synonymies and paraphyly are widespread, which is reflected in widely varying estimates of the number of valid species and genera within the family Cymothoidae [1,3,22–24,26–28]. Additionally, although the monophyly of Cymothoida is believed to be supported by morphological data and rejected by molecular data [7], a relatively recent morphological study also failed to find evidence for the monophyly of Cymothoida [29], so it is increasingly likely that the suborder is indeed paraphyletic [5,11,29]. This phylogenetic and taxonomic uncertainty permeates the deep-level phylogeny of Isopoda as well: there are some indications that the position of Cymothoida, which was traditionally regarded as the most derived isopod taxon [4,7,26,30], may be relatively basal within the

Isopoda [6,7,27] (throughout the manuscript we use these two terms, basal and derived, to refer to common ancestors, not extant species). The few attempts to apply mitochondrial phylogenomics to study the evolutionary history of Isopoda resolved this suborder relatively 'centrally' within the Isopod clade, clustering with Sphaeromatidea and Valvifera [4–6,11]. As only two cymothoid mitogenomes are currently available, both belonging to free-living Cirolanidae species, interpretation and reliability of these findings are severely hampered by the low number of mitogenomes available.

To address this problem, we sequenced and characterised the first complete mitochondrial genome of a parasitic cymothoid isopod, *Cymothoa indica* Schiödte & Meinert 1884 (Cymothoidae), and conducted comparative mitogenomic and phylogenomic analyses. Our results indicate that the availability of this mitogenomic sequence has the potential to advance our understanding of the evolution of mitogenomic architecture and phylogeny of Isopoda, but the amount of available data remains too limited to draw conclusions with confidence.

Results and discussion

Genome architecture and characteristics

The complete mitochondrial genome of *C. indica* is a circular, 14,484 bp-long molecule, somewhat smaller than most of the remaining complete isopod mitogenomes (Worksheet A in S1 File). Although there is evidence for a linear mitogenomic organisation in *Armadillidium vulgare* [12], an isopod species belonging to a different suborder (Oniscidea), we did not find any indications of such organisation in *C. indica* (i.e., all fragments overlapped during the assembly). The mitogenome possesses the standard 13 protein-coding genes and two rRNA genes (*12S* and *16S*), but only 20 tRNA genes, as *trnI* and *trnE* could not be detected (Table 1). A 390 bp-long putative control region (CR) was found between *trnS* and *nad1* genes. The A+T content of the complete mitogenome (63.8%) is average for isopods (54 to 72%; Worksheet A in S1 File). It should be emphasised here again that only five complete mitogenomes were available for the entire order Isopoda when we conducted these analyses, so all comparative analyses in this study were hampered by the fact that the remaining 18 mitogenomes were partial, and should be interpreted with that limitation in mind.

Gene order

The ancestral arthropod architecture, almost identical to the one exhibited by *Limulus polyphemus* (Fig 1), has remained almost unchanged for over 400 million years in many crustacean lineages [31], but not among the isopods [32,33]. Gene order rearrangements in this group of animals were discussed in detail in previous studies [4,11,32], so here we only briefly discuss the idiosyncrasies of the new mitogenome. *Cymothoa indica* also exhibits a completely unique order with a large number of rearrangements (Fig 1). The only available isopod mitogenome exhibiting somewhat similar architecture is that of the only other available Cymothoida (family Cirolanidae) species, *Bathynomus* sp. [11]. In terms of PCG and rRNA arrangement, the two mitogenomes are almost identical, with the exception of the position of *nad1*: in *C. indica* it is on the minus strand, which corresponds to the putative pancrustacean ground pattern [4,34], but in *Bathynomus* sp. it is on the plus strand (Worksheet B in S1 File). However, the two mitogenomes also differ in the arrangement of a number of tRNAs (L2, A, V, E, W, H, L1, S1), and the location of the putative control region (CR). In terms of uniquely derived gene positions of single species in relation to the putative pancrustacean ancestral gene order discussed by Shen et al. [11], *C. indica* shares the unique position of *trnS-nad1* with *Bathynomus* sp. In comparison to the putative ancestral isopod architecture [4], *C. indica* differs in the arrangement of *nad1* and *12S* rRNA genes, which appear to have switched places, and a number of tRNAs: L2,

Table 1. Organisation of the mitochondrial genome of *Cymothoa indica*.

Gene	From	To	Length	IGR	Start	Stop	Anticodon	Strand
trnQ	1	62	62				TTG	-
trnM	67	129	63	4			CAT	+
nad2	130	1131	1002		ATT	TAG		+
trnC	1129	1182	54	-3			GCA	-
trnY	1186	1246	61	3			GTA	-
cox1	1245	2786	1542	-2	ATG	TAA		+
cox2	2830	3508	679	43	ATA	T		+
trnK	3509	3569	61				TTT	+
trnD	3567	3619	53	-3			GTC	+
atp8	3626	3781	156	6	ATA	TAA		+
atp6	3775	4450	676	-7	ATG	T--		+
cox3	4451	5236	786		ATG	TAA		+
trnR	5243	5295	53	6			TCG	+
trnG	5296	5353	58				TCC	+
nad3	5351	5701	351	-3	ATC	TAA		+
trnA	5702	5756	55				TGC	+
trnV	5758	5819	62	1			TAC	+
trnN	5816	5879	64	-4			GTT	+
rrnS	5880	6604	725					+
CR	6605	6994	390					
nad1	6995	7927	933		ATT	TAA		-
trnL1	7937	7996	60	9			TAG	-
trnL2	8043	8102	60	46			TAA	-
trnS	8100	8158	59	-3			TCT	-
trnW	8155	8215	61	-4			TCA	-
cytb	8220	9344	1125	4	TTG	TAG		-
trnT	9345	9398	54				TGT	-
nad5	9398	11095	1698	-1	ATA	TAA		+
trnF	11097	11154	58	1			GAA	+
trnH	11147	11206	60	-8			GTG	-
nad4	11181	12515	1335	-26	ATG	TAA		-
nad4L	12514	12781	268	-2	TGG	T		-
trnP	12797	12860	64	15			TGG	-
nad6	12863	13342	480	2	ATA	TAA		+
trnS	13342	13401	60	-1			TGA	+
rrnL	13402	14475	1074					-

CR is control region. IGR is intergenic region, where a negative value indicates an overlap.

R, V, S1, W, E and I (the last two are missing). Intriguingly, it shares large, completely conserved (in terms of gene order) segments of the mitogenome of 16/11 genes (*cytb* through *trnY*/*rrnL*) with *Asellus aquaticus*/*Eophreatoicus* sp. respectively, which may bear relevance for resolving the conflicting topologies produced by different datasets for Isopoda (see discussion in the 'Phylogeny' section for further details).

Mitogenomic gene order rearrangements are usually attributed to the TDRL (tandem duplication and subsequent evolutionary loss of duplicated genes) mechanism [19,35], but high compactness of isopod mitogenomes indicates that this is not the most parsimonious
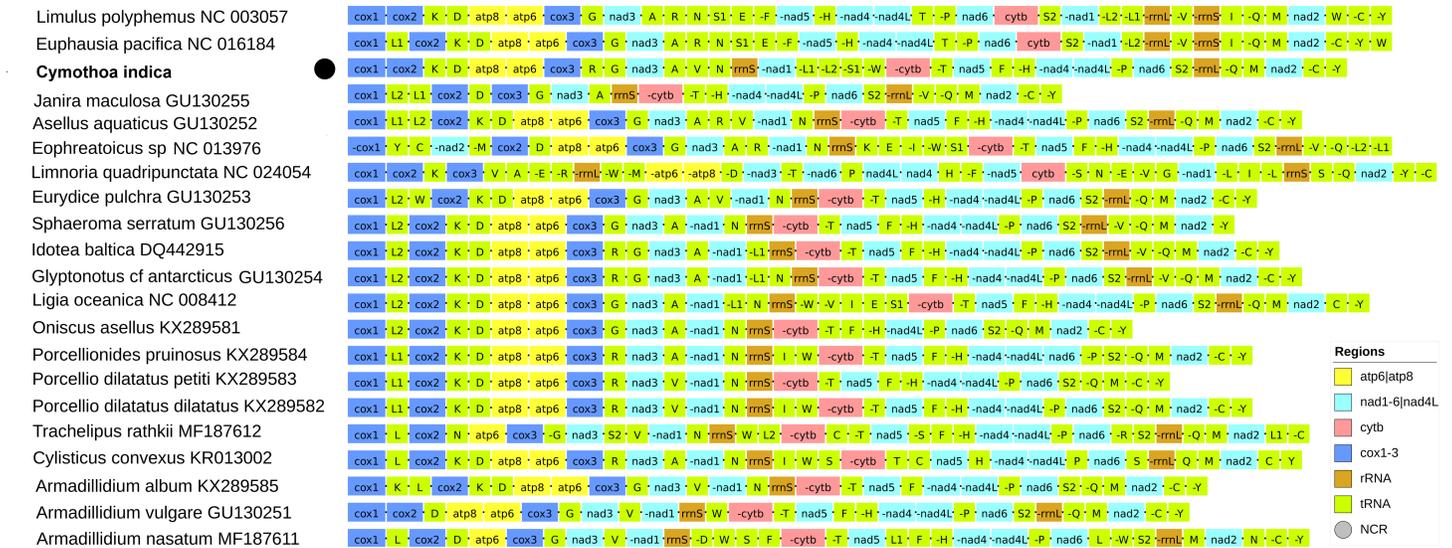
Fig 1. Gene order in isopod mitogenomes. The newly sequenced species, *Cymothoa indica*, is bolded and marked with a black dot. GenBank accession numbers are shown next to the species names, and gene legend is incorporated in the figure. Two outgroups, *Limulus polyphemus* and *Euphausia pacifica*, are also shown.

explanation in this group of animals, as TDRL events result in pseudogenes which are then usually [17,36] 'erased' from the mitogenome in the course of evolution [37]. In terrestrial isopods (Oniscidea: *Armadillidium*), a different mechanism was proposed: the atypical organisation of their mitogenomes, which are composed of linear monomers and circular dimers, might facilitate architecture conversions via creation of telomeric hairpins [14,15]. Doublet et al. [14] have characterised the (highly conserved) CR structure in isopod species that undergo (occasional) mitogenome linearization, but the putative CR of *C. indica* does not appear to possess any of those features. However, regardless of the different organisation, its CR sequence does contain a large number of palindromes (24 ≥ 6bp), which are required for the formation of hairpins. Therefore, we hypothesise that occasional linearization events in their evolutionary history might be the most parsimonious explanation for the large number of gene rearrangements observed in this species and Isopoda in general.

tRNAs

Whereas the majority of sequenced metazoan mitogenomes contain the full set of 22 tRNAs, the number of identified tRNA genes often varies in arthropod mitogenomes [16]. Most sequenced isopod mitogenomes appear to possess an incomplete set of tRNA genes [5,32]. As *trnI* and *trnE* appear to be missing from nine and 16 (respectively) of the available mitogenomes, *C. indica* is not an outlier in this aspect. Another tRNA gene commonly missing from isopod mitogenomes is *trnW* (13 species), but we annotated this gene upstream of *cytb*, as observed in some other isopods (Fig 1), and successfully folded it into a relatively standard cloverleaf structure. The loss of mitochondrial tRNA genes is usually compensated by the import of nuclear tRNAs [16,38], but some unique features described in oniscidean isopod mitogenomes raise suspicion that these missing tRNAs might be actually encoded in their mitogenomes: heteroplasmic mitochondrial DNA, which may allow for the presence of two tRNA genes with different anticodons at the same locus [15,39,40], and tRNA genes partially or fully overlapping with protein coding genes, that have been reported in the oniscidean *Armadillidium* genus [16]. Although we did find an unusually large overlap (26 bp between *nad4* and

trnH), we did not find any evidence for the existence of fully overlapping genes, nor for heteroplasmy (by examining electropherograms), so these features appear to remain limited to oniscidean isopods. A 46 bp-long non-coding fragment between *trnL1* and *L2* genes bears high similarity to *trnE* orthologues in other isopods (including the conserved anticodon), but it is missing the 3' end. We attempted to create a 19-bp overlap with the downstream *trnL2* gene, but the ARWEN tool [41] still did not manage to fold the gene into a cloverleaf structure. Therefore, although the conserved 5' end and anticodon make us suspect that this tRNA gene might be functional after undergoing post-translational editing [42], as we have no evidence for this, it remains merely a speculation.

NCRs and overlaps

Mitogenomes of isopods are believed to be very compact, with many overlaps between genes and short non-coding sequences [32]. In agreement with this, we found 12 intergenic regions ranging from 1 to 46 bp in size (excluding the putative control region). The largest was located between *cox1* and *cox2* genes, followed by 15 bp between *nad4L* and *trnP* genes, whereas the remaining nine were smaller than 10 bp (Table 1). Further evidence for this high compactness is a relatively large number (13) of gene overlaps observed, ranging from 1 bp to 26 bp. Eleven of these overlaps, including the largest, between *nad4* and *trnH*, involved at least one tRNA gene. This is expected and believed to be a consequence of lesser evolutionary constraints on tRNA genes [16]. The only overlaps involving two PCGs were those between *atp8*/*atp6* (7 bp) and *nad4*/*nad4L* (2 bp). However, there are several indications that *atp8* and *nad4L* are generally under relaxed evolutionary constraints: both are exceptionally small (in *C. indica*: *atp8* = 156bp, and *nad4l* = 268bp); *nad4* and *nad4L* often overlap in mitogenomes of many different groups of animals including isopods [32]; *atp8* is often even completely absent from mitogenomes [19]; and *atp6*/*atp8* overlaps of 4–12 bp were also reported in other isopods [5,11,32]. Therefore, we do not suspect an annotation artefact in either of the two overlaps. In conclusion, this mitogenome can also be characterised as highly compact, which bears relevance for inferring the evolutionary history of its architectural rearrangements.

Gene features

Most PCGs of *C. indica* exhibit sizes and start/stop codons standard for isopods (Worksheet C in S1 File). An exception is *cytb*, which (putatively) uses a non-standard TTG start codon. Alternatively, it might use a standard ATT start codon, but that would create a 12-bp overlap with the neighbouring *trnT* gene. The *atp6* gene exhibited a unique 6 bp-long insertion from positions 56 to 61, but this fragment of the sequence is generally poorly conserved so we don't deem this as suspicious. Whereas other isopod *atp6* genes end with a TAA stop codon, in the studied mitogenome this codon appears to have mutated into TAT (presuming this is not a sequencing artefact). This is not likely to affect its transcription, as it can still use T—as the stop codon [43]. In *cox1*, a frameshift mutation (or insertion) near the end of the gene appears to have caused a minor extension of the gene: 1542 bp. vs. 1531–1539 bp. in orthologs (*Limnoria quadripunctata* is an outlier with 1596 bp, but as all other genes are highly conserved in size, we suspect an annotation artefact here; Worksheet C in S1 File). Alternatively, it might use a non-standard stop codon, or span only 1524 bp. The 5' end of *nad5* gene is very divergent, so the start codon is questionable: we selected the standard ATA, creating a 1 bp overlap with the neighbouring tRNA, but if it uses one of alternative start codons, it might start 9, 15, 21, or 24 bp downstream with an ATC.

Several genes exhibit unusually broad size variability in isopods (Worksheet C in S1 File), but in most cases we suspect annotation artefacts, especially as most mitogenomes are incomplete. For example, *cytb* gene exhibits a huge variation in size in the available isopods, from 303 to 1206 bp, but outliers are mostly incomplete, unpublished or unverified mitogenomes, or from a study wherein a large number of mitogenomes were sequenced with the aim to focus on gene rearrangements [4]. An intriguing outlier is *Bathynomus sp.* [11], which exhibited a very divergent 3' end of the sequence, much longer that the rest of the orthologs. As the authors did not discuss this issue, we also suspect an annotation or a sequencing artefact here. Finally, *nad1* gene also exhibits a suspiciously wide size range (876 to 972 bp). Here again, *Bathynomus sp.* is among the outliers, but in this case its large size (969 bp) is very likely to be an artefact, as it can easily be shortened by 15 bp (to 954 bp) to use the same start codon as most other orthologs, ATA (S2 File). Other three outliers, *Oniscus asellus* (876 bp), *Porcellionides pruinosus* (882 bp), and *Armadillidium album* (972 bp) are unverified and come from the same unpublished study, so their unusual sizes would have to be independently confirmed.

Phylogeny

The two phylogenetic analyses, Bayesian Inference (BI) and Maximum Likelihood (ML), produced topologies (Figs 2 and 3) differing in three details: the position of *Limnoria quadripunctata* (sister-clade to Oniscidea in ML, a relatively basal position in BI), the relationship of *Eurydice pulchra* and *Sphaeroma serratum* (sister-clades in ML), and the length of the *Janira maculosa* branch (extremely long in ML). In other aspects, the results of the two analyses were congruent, with the BI topology exhibiting very high statistical support, and the ML topology a mix of mostly high and several lower values. The suborder Oniscidea was resolved as
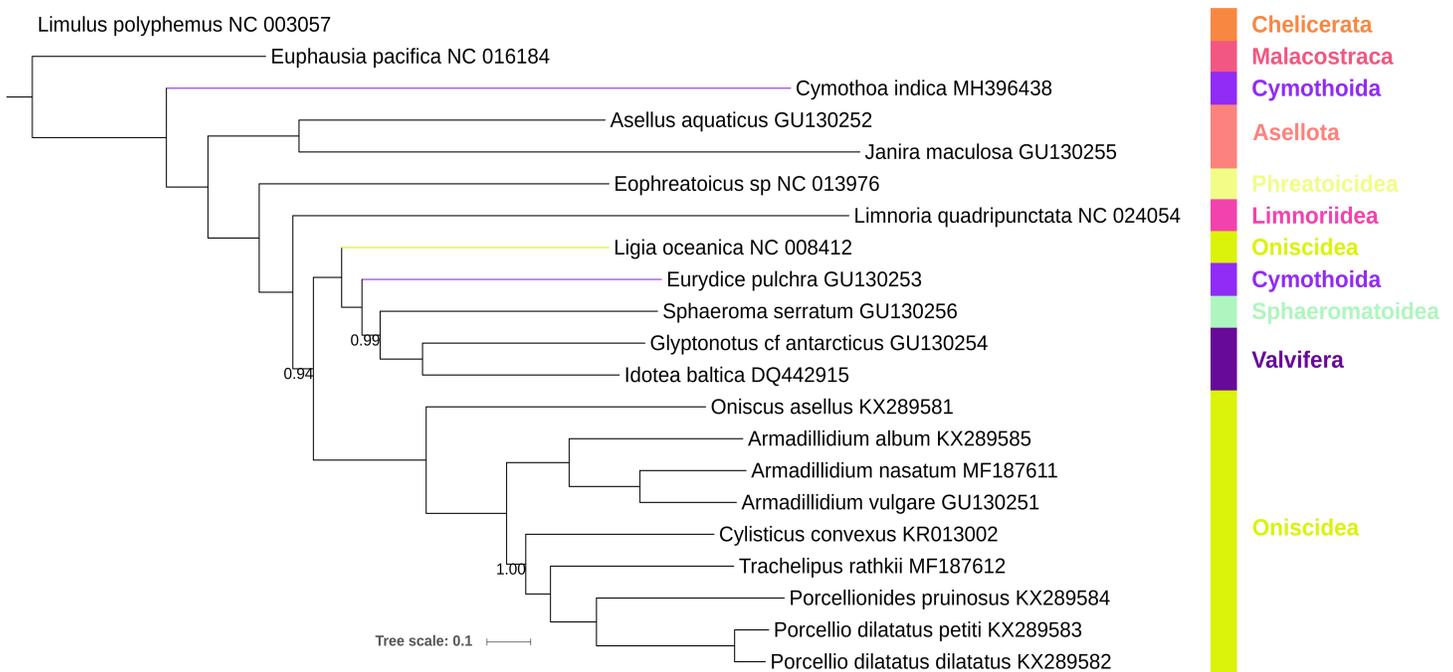
Fig 2. Mitochondrial phylogenomics of Isopoda: Bayesian inference analysis. The analysis was conducted using nucleotide sequences of all genes. *Limulus polyphemus* (branch cropped) and *Euphausia pacifica* are outgroups. Scale bar corresponds to the estimated number of substitutions per site. Bayesian posterior probability values (lower than 1.0) are shown next to corresponding nodes. GenBank accession numbers are shown next to species names. Taxonomic rank (suborder/superfamily) is shown to the right. Coloured branches highlight paraphyly.

https://doi.org/10.1371/journal.pone.0203089.g002
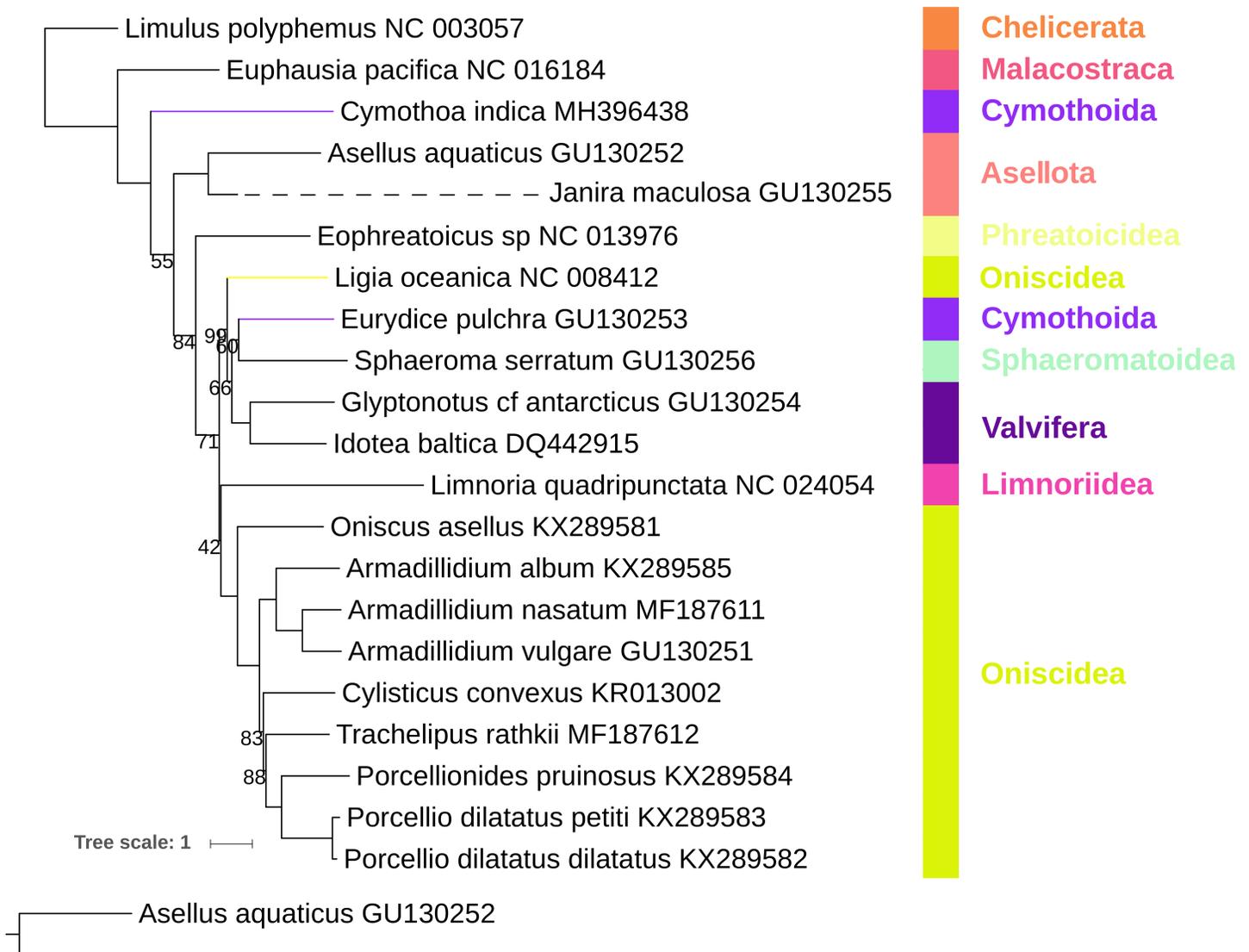
Fig 3. Mitochondrial phylogenomics of Isopoda: Maximum likelihood analysis. The analysis was conducted using nucleotide sequences of all genes. *Limulus polyphemus* and *Euphausia pacifica* are outgroups. Scale bar corresponds to the estimated number of substitutions per site. Bootstrap support values (lower than 100) are shown next to corresponding nodes. GenBank accession numbers are shown next to species names. Taxonomic rank (suborder/superfamily) is shown to the right. Coloured branches highlight paraphyly. *Janira maculosa* branch (dashed line) is shortened, with its original size shown below the phylogram.

https://doi.org/10.1371/journal.pone.0203089.g003

paraphyletic by *Ligia oceanica* [32] exhibiting a sister-clade relationship with a cluster of other taxa (*Eurydice pulchra*, Sphaeromatoidea, and Valvifera; see Worksheet A in S1 File for detailed taxonomy). Paraphyly of Oniscidea caused by this genus is a relatively well-established fact [6], and all recent mito-phylogenomic analyses resolved this species approximately in the same position [5,6,11]. Although it is tempting to interpret this as a sign that its position is resolved, we advise that additional mitogenomes belonging to this genus should be sequenced, to exclude the possibility of misidentification or unusual evolutionary rates, before any taxonomic changes are officially proposed. The *Limnoria* genus was described as 'rogue' a long time ago [7], and its position also remains unresolved to this day, including by our study. In recent mitochondrial phylogenomic studies it was mostly resolved as a sister-clade to Oniscidea (which corresponds to our ML topology) using different datasets (nucleotides and amino

acids) and algorithms (BI and ML) [5,11]. However, Lins et al. [6] resolved it as a sister-clade to all other isopoda (BI, amino acid dataset), and a combined mitonuclear dataset (*18S*, *28S*, and *cox1*) in the same study resolved it as a sister-clade to a derived Asellota clade (Asellota was paraphyletic in that analysis). Therefore, although most mitogenomic analyses resolve it as a sister-clade to Oniscidea, extreme variations in its position indicate that the evolution of this species is very peculiar, and that more molecular data of closely related species should be sequenced to identify the reasons underlying this instability. As the *J. maculosa* mitogenome is incomplete (<10,000 bp), we suspected that the long branch produced by the ML analysis may have been an artefact. To test this hypothesis, we removed the four genes (*atp6* and *8*, *nad1* and *5*) missing from *J. maculosa* from the entire dataset and re-conducted the ML analysis. This analysis produced a congruent topology (S1 Fig.) and resolved the issue, with the length of the branch comparable to the one produced by the BI analysis.

Most importantly, the newly-sequenced *C. indica* species was resolved as a sister group (basal) to all remaining isopods, which challenges the accepted isopod phylogeny, where Phreatoicidea (represented by *Eophreatoicus* sp.) is usually regarded as the most basal split within the Isopoda [4,7,30], and Cymothoida as the most derived group [4,7,26,30]. The most basal position of Phreatoicidea was also challenged by the position of Asellota clade (*A. aquaticus* and *J. maculosa*), which was derived in relation to Cymothoidae, but basal to Phreatoicidea and all other Isopoda. This topology (Asellota basal to Phreatoicidea) was produced by all other recent mitochondrial phylogenomic analyses as well [5,6,11] (with the exception of an amino acid dataset in [11]). Although all of these topologies support the close relationship of these two clades, and their relatively basal position within the Isopoda clade, none of them show support for the proposed monophyly of a combined asellotan and phreatoicidean clade [4,44]. The suborder Cymothoida was paraphyletic: the other cymothoid species included in the analysis, *Eurydice pulchra* (Cymothoida: Cirolanidae), clustered with Sphaeromatoidea and Valvifera (low support in ML). This position of *E. pulchra* is congruent with the topologies produced by most other mito-phylogenomic studies [4–6,11]. A topology highly congruent to ours, including the deep evolutionary split between free-living (Cymothoidae) and parasitic (Cirolanidae) cymothoid species, was produced before using a combined *18S*-morphology dataset [7]. Similar topologies were also produced by two Bayesian analyses of combined mitonuclear gene datasets [6,27], but this was not further discussed by the authors. Intriguingly, that same mitonuclear dataset resolved *E. pulchra* within the basal cymothoid clade [6]. The basal position of *C. indica* within the isopod clade is further indirectly supported by the similarity in gene order with *A. aquaticus* and *Eophreatoicus* sp. (*J. maculosa* is incomplete, so it is difficult to assess the level of similarity). Therefore, although the issue cannot be declared resolved yet, it appears that there is increasing evidence from different types of data (gene order, mitochondrial phylogenomics, mitochondrial and nuclear single-marker, and morphological data) for deep paraphyly of the suborder Cymothoida, and for the parasitic Cymothoidae being sister group to all other Isopoda.

Conclusions

As the absence of a sufficient number of sequenced mitogenomes is currently the foremost limiting factor to their broader application, we sequenced the first mitogenome of a species belonging to the large family Cymothoidae, *Cymothoa indica*. The results of our phylogenetic analyses, which resolved *C. indica* as the most basal split in the isopod clade, present a major challenge to the accepted deep phylogeny of Isopoda. There is growing evidence that Cymothoida might be split into two evolutionary distant clades, with parasitic species being the most basal split in the

isopod clade. However, the small number of isopod mitogenomes currently available (as well as the minor topological instability observed) prevents us from making any conclusions with confidence. Therefore, a much larger amount of molecular resources carrying a high phylogenetic resolution will be needed to infer the remarkably complex evolutionary history of this group of animals with confidence. Aside from the importance of mitochondrial genomes for taxonomy and phylogenetics, the highly rearranged and unique gene order found in *Cymothoa indica* is a further indication that isopods might present an important model system to study the discontinuous dynamics of the evolution of mitochondrial genomic architecture. Therefore, sequencing of further isopod mitogenomes is strongly urged.

Materials and methods

Samples, identification, and DNA extraction

Two adult specimens were collected on 10/07/2017 in Dayawan Town, Guangdong Province, China (22˚42'58"– 22˚42'56" N; 114˚32'16"– 114˚32'25" E) from the mouth of a euryhaline fish species, *Mugil cephalus* Linnaeus 1758. Live parasites were kept alive in 0.6% saline as long as possible (one day) to ensure that they were starved, and then stored in 75% ethanol at 4˚C. Specimens were morphologically identified under a dissecting microscope as described before [1,3,45,46]. After washing in distilled water, DNA was isolated from one specimen using Aidlab DNA extraction kit (Aidlab Biotechnologies, Beijing, China). This study has been reviewed and approved by the ethics committee of the Institute of Hydrobiology, Chinese Academy of Sciences. As the study involved an unregulated parasitic invertebrate, and as we obtained the samples from already dead fish bought on a local fish market, no permits were required to retrieve and process the samples.

Genome sequencing and assembly

Ten primer pairs used to amplify and sequence the entire mitogenome were designed to match conserved regions of mitochondrial genes and to overlap by approximately 100 bp (Table 2). Amplification reaction mixture and conditions used were described before [17]; briefly: 50μL with 5 U/μL of TaKaRa LA Taq polymerase (TaKaRa, Japan), 10×LATaq Buffer II, 2.5μM dNTP mixture, 0.2–1.0μM each primer, 60ng DNA template. Conditions: denaturation 98˚C/2min, and 40 cycles of 98˚C/10s, 50˚C/15s, 68˚C/1min per kb. When the product was not specific enough, PCR conditions were optimized by increasing the annealing temperature and decreasing the number of cycles. PCR products were sequenced using Sanger method and the same set of primers. All obtained fragments were quality-proofed by visually inspecting the electropherograms and BLASTed [47] to confirm that the amplicon is the actual target sequence. Mitogenome was assembled, annotated, and comparative analyses conducted, roughly as described previously [17,48]. Briefly: assembly was conducted manually using DNAstar v7.1 [49]. In each step we checked whether overlaps were identical, thereby making sure that the mitogenome is circular, and avoiding incorporation of *numt*s [50] into the sequence. The same software was also used to locate the putative ORFs for protein-coding genes. BLAST and BLASTx were used to compare the inferred ORFs with nucleotide and amino acid sequences of available orthologs, and manually determine the exact initiation and termination codon positions accordingly. tRNAs were annotated using tRNAscan [51] and ARWEN [41] tools, and the results checked manually. The annotation was recorded in a Word (Microsoft Office) document, and then parsed and extracted using an in-house MitoTool software [52]. The same software was also used for file conversions (including the creation of the GenBank file) and to generate tables with comparative mito-genomic statistics (including the file used to visualise gene orders in iTOL). Palindromes in

Table 2. Primers used for amplification and sequencing of the mitochondrial genome of *Cymothoa indica*.

Gene/region	Name	Sequence (5'-3')	Length
COX1	LYF1	GCTGGGATATTAGGTCTTAG	1436
	LYR1	GAGTGTTCGGAGGGAGGGAA	
COX1-COX2	LYF2	GACGTTATTCAGATTACCCTG	663
	LYR2	GGATAACAAGTTTGTTATCTG	
COX2	LYF3	CTGATGAAACTTTTTCATCAC	356
	LYR3	GAAACTATGATTTGCACCAC	
COX2-COX3	LYF4	GGACAATCCCATCACTTGGG	1462
	LYR4	TTAGGAGACAATCTTCTATG	
COX3	LYF5	GATGTCTCACGAGAAGCAAG	442
	LYR5	GAAAGCCATGAAAACCAGTAG	
COX3-16S	LYF6	CAATTATTCTTGGGATTAC	2402
	LYR6	GACCCTAAGAATTTGAAGATC	
16S	LYF7	TACGCTGTTATCCCTAGAG	828
	LYR7	CGTACCTTTAGCATTAGGG	
16S-CYTB	LYF8	GAAAAGAATTTCACATCTAAAG	5995
	LYR8	CCAAAAGGGTTTCTTGATCC	
CYTB	LYF9	GCAATCCCATATATCGGTTC	370
	LYR9	GAAAGTACCATTCAGGTTG	
CYTB-COX1	LYF10	CGATCATTTACCCTTATAGAC	2247
	LYR10	CGCCAATTATGATAGGTATAAC	

the CR were predicted using Palindrome analyser [53]. The mitogenome is available from the GenBank repository under the accession number MH396438.

Phylogenetic analyses

Three complete and 15 partial isopod sequences were retrieved from the GenBank for the phylogenetic analyses. A basal arthropod, *Limulus polyphemus* [31], and a Malacostraca species basal to Isopoda [32], *Euphausia pacifica* [54], were used as outgroups. To maximize the amount of phylogenetic signal, we conducted the analyses on a dataset containing all genes (PCGs and RNAs). Genes were extracted from GenBank files using MitoTool. Nucleotide sequences of protein-coding genes were aligned in batches (using codon-alignment mode) with MAFFT [55] integrated into another in-house software package—BioSuite [56]. As described before [57], RNAs were aligned by an algorithm that takes secondary structure information into account, Q-INS-i, incorporated into MAFFT-with-extensions software [58]. BioSuite was then used to concatenate the alignments, and another plug-in program in BioSuite, Gblocks [59,60], was used to remove ambiguously aligned regions [59,60]. As a result, although a majority of mitogenomes used for the analysis were incomplete, the final alignment (S3 File) had a relatively low proportion of gaps and undetermined characters (7.68%) and a high number of distinct alignment patterns (9,641). Best partitioning scheme and evolutionary models for partitions (GTR+I +G, GTR+G; S4 File) were selected using PartitionFinder [61], also implemented in BioSuite. Phylogenetic analyses were conducted using Bayesian Inference method implemented in MrBayes 3.2.6 [62] (default settings, two parallel runs, five million MCMC generations) and Maximum Likelihood method implemented in RAxML 8.1.21 [63] (1000 rapid bootstrap replicates). Phylograms and gene orders were visualized in iTOL [64]. WoRMS database [65] was used as the authority for the taxonomic nomenclature.

Supporting information

S1 File. Comparative analyses of isopod mitogenomes. Worksheet A: taxonomy and basic statistics for all available isopod mitogenomes. Worksheet B: ancestral gene orders. Worksheet C: gene statistics for all available isopod mitogenomes (+ outgroups): gene sizes, start and terminal codons. Species are represented by acronyms of their binomial scientific names (see footnote). The new sequence (*C. indica*) and outgroups are shaded grey.
(XLSX)

S2 File. Alignment of selected isopod *nad1* orthologs.
(FAS)

S3 File. Alignment of isopod mitogenomes used for phylogenetic analyses.
(FAS)

S4 File. PartitionFinder results: Partitioning and model selection.
(TXT)

S1 Fig. Maximum likelihood phylogenetic analysis conducted on a partial dataset. Four genes missing from *Janira maculosa* were removed from the entire dataset: *atp6*, *atp8*, *nad1* and *nad5*. See caption for Fig 3 for other details.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: Hong Zou, Wen-Xiang Li, Gui-Tang Wang.

Data curation: Hong Zou, Dong Zhang, Rong Chen, Shahid Mahboob, Khalid Abdullah Al-Ghanim, Fahad Al-Misned.

Formal analysis: Hong Zou, Ivan Jakovlić, Dong Zhang, Shahid Mahboob, Khalid Abdullah Al-Ghanim, Fahad Al-Misned.

Funding acquisition: Hong Zou, Shahid Mahboob, Wen-Xiang Li, Gui-Tang Wang.

Investigation: Hong Zou, Dong Zhang, Rong Chen, Shahid Mahboob, Khalid Abdullah Al-Ghanim, Fahad Al-Misned.

Methodology: Hong Zou, Ivan Jakovlić, Dong Zhang, Rong Chen.

Project administration: Hong Zou, Wen-Xiang Li, Gui-Tang Wang.

Resources: Hong Zou, Dong Zhang, Rong Chen, Shahid Mahboob, Khalid Abdullah Al-Ghanim, Fahad Al-Misned.

Software: Dong Zhang.

Supervision: Wen-Xiang Li, Gui-Tang Wang.

Validation: Hong Zou, Ivan Jakovlić, Dong Zhang, Rong Chen.

Visualization: Hong Zou, Ivan Jakovlić, Dong Zhang.

Writing – original draft: Hong Zou, Ivan Jakovlić.

Writing – review & editing: Hong Zou, Ivan Jakovlić, Dong Zhang, Rong Chen, Shahid Mahboob, Khalid Abdullah Al-Ghanim, Fahad Al-Misned, Wen-Xiang Li, Gui-Tang Wang.

References

1. Poore GCB, Bruce NL. Global Diversity of Marine Isopods (Except Asellota and Crustacean Symbionts). MacKenzie BR, editor. PLoS One. Public Library of Science; 2012; 7: e43529. https://doi.org/10.1371/journal.pone.0043529 PMID: 22952700

2. Wilson GDF. Global diversity of Isopod crustaceans (Crustacea; Isopoda) in freshwater. Hydrobiologia. 2008. pp. 231–240. https://doi.org/10.1007/s10750-007-9019-z

3. Martin MB, Bruce NL, Nowak BF. Review of the fish-parasitic genus Cymothoa Fabricius, 1793 (Crustacea: Isopoda: Cymothoidae) from Australia. Zootaxa. 2016; 4119: 1–72. https://doi.org/10.11646/zootaxa.4119.1.1 PMID: 27395199

4. Kilpert F, Held C, Podsiadlowski L. Multiple rearrangements in mitochondrial genomes of Isopoda and phylogenetic implications. Mol Phylogenet Evol. Elsevier Inc.; 2012; 64: 106–117. https://doi.org/10.1016/j.ympev.2012.03.013 PMID: 22491068

5. Yu J, An J, Li Y, Boyko CB. The first complete mitochondrial genome of a parasitic isopod supports Epicaridea Latreille, 1825 as a suborder and reveals the less conservative genome of isopods. Syst Parasitol. Springer Netherlands; 2018; https://doi.org/10.1007/s11230-018-9792-2 PMID: 29644508

6. Lins LSF, Ho SYW, Lo N. An evolutionary timescale for terrestrial isopods and a lack of molecular support for the monophyly of Oniscidea (Crustacea: Isopoda). Org Divers Evol. Organisms Diversity & Evolution; 2017; 17: 813–820. https://doi.org/10.1007/s13127-017-0346-2

7. Wilson GDF. The phylogenetic position of the Isopoda in the Peracarida (Crustacea: Malacostraca). Arthropod Syst Phylogeny. 2009; 67: 159–198.

8. Cameron SL. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu Rev Entomol. 2014; 59: 95–117. https://doi.org/10.1146/annurev-ento-011613-162007 PMID: 24160435

9. Lan T, Gill S, Bellemain E, Bischof R, Nawaz MA, Lindqvist C. Evolutionary history of enigmatic bears in the Tibetan Plateau-Himalaya region and the identity of the yeti. Proc R Soc B Biol Sci. The Royal Society; 2017; 284: 20171804. https://doi.org/10.1098/rspb.2017.1804 PMID: 29187630

10. Der Sarkissian C, Vilstrup JT, Schubert M, Seguin-Orlando A, Eme D, Weinstock J, et al. Mitochondrial genomes reveal the extinct Hippidion as an outgroup to all living equids. Biol Lett. 2015; 11: 20141058. https://doi.org/10.1098/rsbl.2014.1058 PMID: 25762573

11. Shen Y, Kou Q, Zhong Z, Li X, He L, He S, et al. 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. Ecol Evol. 2017; 7: 1869–1881. https://doi.org/10.1002/ece3.2737 PMID: 28331594

12. Raimond R, Marcadé I, Bouchon D, Rigaud T, Bossy JP, Souty-Grosset C. Organization of the large mitochondrial genome in the isopod Armadillidium vulgare. Genetics. 1999; 151: 203–210. Available: http://www.genetics.org/content/151/1/203 PMID: 9872960

13. Peccoud J, Chebbi MA, Cormier A, Moumen B, Gilbert C, Marcadé I, et al. Untangling heteroplasmy, structure, and evolution of an atypical mitochondrial genome by pacbio sequencing. Genetics. 2017; 207: 269–280. https://doi.org/10.1534/genetics.117.203380 PMID: 28679546

14. Doublet V, Helleu Q, Raimond R, Souty-Grosset C, Marcadé I. Inverted repeats and genome architecture conversions of terrestrial isopods mitochondrial DNA. J Mol Evol. 2013; 77: 107–118. https://doi.org/10.1007/s00239-013-9587-7 PMID: 24068302

15. Doublet V, Raimond R, Grandjean F, Lafitte A, Souty-Grosset C, Marcadé I, et al. Widespread atypical mitochondrial DNA structure in isopods (Crustacea, Peracarida) related to a constitutive heteroplasmy in terrestrial species. Genome. 2012; 55: 234–244. https://doi.org/10.1139/g2012-008 PMID: 22376074

16. Doublet V, Ubrig E, Alioua A, Bouchon D, Marcade I, Marechal-Drouard L. Large gene overlaps and tRNA processing in the compact mitochondrial genome of the crustacean Armadillidium vulgare. RNA Biol. 2015; 12: 1159–1168. https://doi.org/10.1080/15476286.2015.1090078 PMID: 26361137

17. Zou H, Jakovlić I, Chen R, Zhang D, Zhang J, Li W-X, et al. The complete mitochondrial genome of parasitic nematode Camallanus cotti: extreme discontinuity in the rate of mitogenomic architecture evolution within the Chromadorea class. BMC Genomics. 2017; 18: 840. https://doi.org/10.1186/s12864-017-4237-x PMID: 29096600

18. Smith DR, Keeling PJ. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proc Natl Acad Sci. 2015; 112: 10177–10184. https://doi.org/10.1073/pnas.1422049112 PMID: 25814499

19. Gissi C, Iannelli F, Pesole G. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity (Edinb). 2008; 101: 301–320. https://doi.org/10.1038/hdy.2008.62 PMID: 18612321

20. Smith DR. The past, present and future of mitochondrial genomics: Have we sequenced enough mtDNAs? Brief Funct Genomics. 2016; 15: 47–54. https://doi.org/10.1093/bfgp/elv027 PMID: 26117139

21. Ballard JWO, Whitlock MC. The incomplete natural history of mitochondria. Mol Ecol. 2004; 13: 729–744. https://doi.org/10.1046/j.1365-294X.2003.02063.x PMID: 15012752

22. Hata H, Sogabe A, Tada S, Nishimoto R, Nakano R, Kohya N, et al. Molecular phylogeny of obligate fish parasites of the family Cymothoidae (Isopoda, Crustacea): evolution of the attachment mode to host fish and the habitat shift from saline water to freshwater. Mar Biol. Springer Berlin Heidelberg; 2017;164. https://doi.org/10.1007/s00227-017-3138-5

23. Smit NJ, Bruce NL, Hadfield KA. Global diversity of fish parasitic isopod crustaceans of the family Cymothoidae. Int J Parasitol Parasites Wildl. Australian Society for Parasitology; 2014; 3: 188–197. https://doi.org/10.1016/j.ijppaw.2014.03.004 PMID: 25180163

24. Jones CM, Miller TL, Grutter AS, Cribb TH. Natatory-stage cymothoid isopods: Description, molecular identification and evolution of attachment. Int J Parasitol. 2008; 38: 477–491. https://doi.org/10.1016/j.ijpara.2007.07.013 PMID: 17880981

25. Fogelman RM, Kuris AM, Grutter AS. Parasitic castration of a vertebrate: Effect of the cymothoid isopod, Anilocra apogonae, on the five-lined cardinalfish, Cheilodipterus quinquelineatus. Int J Parasitol. Australian Society for Parasitology Inc.; 2009; 39: 577–583. https://doi.org/10.1016/j.ijpara.2008.10.013 PMID: 19041314

26. Brusca RC. A monograph on the Isopoda Cymothoidae (Crustacea) of the eastern Pacific. Zool J Linn Soc. 1981; 73: 117–199. https://doi.org/10.1111/j.1096-3642.1981.tb01592.x

27. Lins LSF, Ho SYW, Wilson GDF, Lo N. Evidence for Permo-Triassic colonization of the deep sea by isopods. Biol Lett. 2012; 8: 979–982. https://doi.org/10.1098/rsbl.2012.0774 PMID: 23054914

28. Joca LK, Leray VL, Zigler KS, Brusca RC. A new host and reproduction at a small size for the "snapper-choking isopod" Cymothoa excisa (Isopoda: Cymothoidae). J Crustac Biol. Oxford University Press; 2015; 35: 292–294. https://doi.org/10.1163/1937240X-00002312

29. Wilson GDF, Sims C a., Grutter AS. Toward a Taxonomy of the Gnathiidae (Isopoda) Using Juveniles: The External Anatomy of Gnathia aureamaculosa Zuphea Stages Using Scanning Electron Microscopy. J Crustac Biol. 2011; 31: 509–522. https://doi.org/10.1651/10-3432.1

30. Wetzer R. Mitochondrial genes and isopod phylogeny (Peracarida: Isopoda). J Crustac Biol. 2002; 22: 1–14. https://doi.org/10.2307/1549602

31. Lavrov DV, Boore JL, Brown WM. The complete mitochondrial DNA sequence of the horseshoe crab Limulus polyphemus. Mol Biol Evol. 2000; 17: 813–824. Available: http://www.ncbi.nlm.nih.gov/pubmed/10779542 https://doi.org/10.1093/oxfordjournals.molbev.a026360 PMID: 10779542

32. Kilpert F, Podsiadlowski L. 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. 2006; 7: 241. https://doi.org/10.1186/1471-2164-7-241 PMID: 16987408

33. Podsiadlowski L, Bartolomaeus T. Major rearrangements characterize the mitochondrial genome of the isopod Idotea baltica (Crustacea: Peracarida). Mol Phylogenet Evol. 2006; 40: 893–899. https://doi.org/10.1016/j.ympev.2006.04.008 PMID: 16737831

34. Boore JL, Lavrov DV., Brown WM. Gene translocation links insects and crustaceans. Nature. Nature Publishing Group; 1998; 392: 667–668. https://doi.org/10.1038/33577 PMID: 9565028

35. Boore JL. The Duplication/Random Loss Model for Gene Rearrangement Exemplified by Mitochondrial Genomes of Deuterostome Animals. In: Sankoff D, Nadeau JH, editors. Comparative Genomics. Springer Netherlands; 2000. pp. 133–147. https://doi.org/10.1007/978-94-011-4309-7_13

36. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 1999; 151: 1531–1545. PMID: 10101175

37. Lynch M, Koskella B, Schaack S. Mutation pressure and the evolution of organelle genomic architecture. Science. 2006. pp. 1727–1730. https://doi.org/10.1126/science.1118884 PMID: 16556832

38. Duchêne AM, Pujol C, Maréchal-Drouard L. Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet. 2009; 55: 1–18. https://doi.org/10.1007/s00294-008-0223-9 PMID: 19083240

39. Chandler CH, Badawi M, Moumen B, Grève P, Cordaux R. Multiple Conserved Heteroplasmic Sites in tRNA Genes in the Mitochondrial Genomes of Terrestrial Isopods (Oniscidea). G3 Genes|Genomes| Genetics. 2015; 5: 1317–1322. https://doi.org/10.1534/g3.115.018283 PMID: 25911226

40. Doublet V, Souty-Grosset C, Bouchon D, Cordaux R, Marcadé I. A thirty million year-old inherited heteroplasmy. PLoS One. 2008;3. https://doi.org/10.1371/journal.pone.0002938 PMID: 18698356

41. Laslett D, Canbäck B. ARWEN: A program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics. 2008; 24: 172–175. https://doi.org/10.1093/bioinformatics/btm573 PMID: 18033792

42. Segovia R, Pett W, Trewick S, Lavrov DV. Extensive and Evolutionarily Persistent Mitochondrial tRNA Editing in Velvet Worms (Phylum Onychophora). Mol Biol Evol. Oxford University Press; 2011; 28: 2873–2881. https://doi.org/10.1093/molbev/msr113 PMID: 21546355

43. Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature. 1981; 290: 470–474. https://doi.org/10.1038/290470a0 PMID: 7219536

44. Wilson GDF. Some of the deep-sea fauna is ancient. Crustaceana. 1999; 72: 1019–1030. https://doi.org/10.1163/156854099503915

45. Yu H, Li X. Study on the Cymothoidae from Chinese waters. Stud Mar Sin. 2003; 223–238 (In Chinese, with English Abstract).

46. Nagasawa K, Doi H. The spotfin burrfish (chilomycterus reticulatus), a new host record for cymothoa pulchra (isopoda, cymothoidae). Crustaceana. 2012; 85: 893–896. https://doi.org/10.1163/156854012X649900

47. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 1997; 25: 3389–3402. https://doi.org/10.1093/nar/25.17.3389 PMID: 9254694

48. Zhang D, Zou H, Wu SG, Li M, Jakovlić I, Zhang J, et al. Sequencing of the complete mitochondrial genome of a fish-parasitic flatworm Paratetraonchoides inermis (Platyhelminthes: Monogenea): tRNA gene arrangement reshuffling and implications for phylogeny. Parasit Vectors. 2017; 10: 462. https://doi.org/10.1186/s13071-017-2404-1 PMID: 29017532

49. Burland TG. DNASTAR's Lasergene sequence analysis software. In: Misener S, Krawetz SA, editors. Methods in Molecular Biology™. Totowa, NJ: Humana Press; 2000. pp. 71–91. https://doi.org/10.1385/1-59259-192-2:71

50. Hazkani-Covo E, Zeller RM, Martin W. Molecular poltergeists: Mitochondrial DNA copies (numts) in sequenced nuclear genomes. PLoS Genet. 2010; 6: e1000834. https://doi.org/10.1371/journal.pgen.1000834 PMID: 20168995

51. Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. 2005; 33: W686–9. https://doi.org/10.1093/nar/gki366 PMID: 15980563

52. Zhang D. MitoTool [Internet]. [cited 29 May 2018]. Available: https://github.com/dongzhang0725/BioSuite

53. Brázda V, Kolomazník J, Lýsek J, Hároníková L, Coufal J, Šťastný J. Palindrome analyser—A new web-based server for predicting and evaluating inverted repeats in nucleotide sequences. Biochem Biophys Res Commun. 2016; 478: 1739–1745. https://doi.org/10.1016/j.bbrc.2016.09.015 PMID: 27603574

54. Shen X, Wang H, Wang M, Liu B. The complete mitochondrial genome sequence of Euphausia pacifica (Malacostraca: Euphausiacea) reveals a novel gene order and unusual tandem repeats. Genome. 2011; 54: 911–22. https://doi.org/10.1139/g11-053 PMID: 22017501

55. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol. 2013; 30: 772–780. https://doi.org/10.1093/molbev/mst010 PMID: 23329690

56. Zhang D. BioSuite [Internet]. [cited 28 May 2018]. Available: https://github.com/dongzhang0725/BioSuite

57. Liu F-F, Li Y-P, Jakovlić I, Yuan X-Q. Tandem duplication of two tRNA genes in the mitochondrial genome of Tagiades vajuna (Lepidoptera: Hesperiidae). Eur J Entomol. 2017; 114: 407–415. https://doi.org/10.14411/eje.2017.052

58. Katoh K, Toh H. Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. BMC Bioinformatics. 2008; 9: 212. https://doi.org/10.1186/1471-2105-9-212 PMID: 18439255

59. Castresana J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. Mol Biol Evol. 2000; 17: 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334 PMID: 10742046

60. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007; 56: 564–577. https://doi.org/10.1080/10635150701472164 PMID: 17654362

61. Lanfear R, Calcott B, Ho SYW, Guindon S. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Mol Biol Evol. Oxford University Press; 2012; 29: 1695–1701. https://doi.org/10.1093/molbev/mss020 PMID: 22319168

62. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, et al. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012; 61: 539–542. https://doi.org/10.1093/sysbio/sys029 PMID: 22357727

63. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30: 1312–1313. https://doi.org/10.1093/bioinformatics/btu033 PMID: 24451623

64. Letunic I, Bork P. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. Bioinformatics. 2007; 23: 127–128. https://doi.org/10.1093/bioinformatics/btl529 PMID: 17050570

65. Horton T, Kroh A, Ahyong S, Bailly N, Boury-Esnault N, Brandão SN, et al. World Register of Marine Species (WoRMS). In: Available from http://www.marinespecies.org at VLIZ. [Internet]. WoRMS Editorial Board; 2018 [cited 2 May 2018]. doi:10.14284/170