Phylogeny of Terrestrial Isopods Based on the Complete Mitochondrial Genomes, Subvert the Monophyly of Oniscidea and Ligiidae up to New Subfamily Ligiaidea of Isopoda

Rui Zhang
Shanxi Normal University
Ruru Chen
Shanxi Normal University
Jianmei An (anjianmei@hotmail.com)
Shanxi Normal University
Carlos A. Santamaria
University of South Florida Sarasota-Manatee

Research article

Keywords: Oniscidea, Phylogeny, new subfamily, Ligiidae, Ligiaidea

DOI: https://doi.org/10.21203/rs.3.rs-25479/v1

Abstract

Background: Oniscidea is the only truly terrestrial taxon within the Crustacea, and vital to soil formation. However, the monophyly of suborder Oniscidea has been in dispute since 1995, with different studies disagreeing on whether the coastal Ligiidae are included within the suborder. To clarify the phylogenetic hypothesis of suborder Oniscidea, we sequenced the complete mitochondrial genomes of *Ligia exotica* (Roux, 1828) and *Mongoloniscus sinensis* (Dollfus, 1901).

Results: Like most metazoan, the complete mitogenomes of two species with circular double strands. The structure and characters of mitogenomes of these two species are analyzed. The constructed phylogenetic analyses show that Oniscidea is polyphyletic group, with *Ligia* being more closely related to marine isopods (Valvifera + Cymothoida + Sphaeromatidea).

Conclusions: We elevate the taxonomic status of the family Ligiidae to the suborder Ligiaidea which are with parallel rank with Oniscidea. Ligiaidea is much primitive than other exact terrestrial isopods. Crinocheta are strongly monophyly, family Agnaridae is more closely related to Porcellionidae rather than Armadillididae.

Background

Isopoda Latreille, 1817, with more than 10,300 species, is the largest taxon within Peracarida and contains species exhibiting amazing ecological diversity and morphological plasticity [1, 2]. The distribution range of the Isopoda ranges from the deep oceans to elevations of 4,700-m above sea level. Currently, the Isopoda is divided into several sub-orders, with all terrestrial species being placed in the Oniscidea [3, 4]. Members of this suborder are saprophagous residents of damp and gloomy environment, and most of them hide during the day and come out at night, due to their extreme sensitivity to temperature, humidity, light and other conditions in the environment. The suborder is currently separated into five lineages: Ligiidae, Tylidae, Meaoniscidae, Synocheta and Crinocheta [5, 6]. Of these lineages, the two formers appear to contain at least some amphibian genera, with *Ligia* Fabricius 1798 being of particular interest as its morphological, physiological and behavioral characteristics have led them to be considered intermediate forms between the marine isopod ancestors of and the modern land-forms of the Oniscidea [6, 7]. Indeed, Schmidt (2008) reported the more primitive species in the Oniscidea is the Ligiidae based on the morphological features of 3527 species, and the evolution of terrestrial isopods was from aquatic to terrestrial, not pass through the fresh water stage [6]. Thus, understanding the placement of Ligiidae within the Oniscidea is critical to understand the evolution of terrestrial isopods.

The Oniscidea have been regarded as a monophyletic group [5, 8, 9] based on shared morphological characters or phylogenetic analyses of one to two genes. Recent studies; however, have failed to capture the monophyly of the Oniscidea using various approaches and data types [10, 11, 12, 13, 14, 15]. Most recently, Dimitriou et al (2019) failed to recover the monophyly of the Oniscidea when carrying out phylogenetic reconstructions based on four highly conserved nuclear genes. In said study, *Ligia* species were reported to be most closely related to species within marine sub-orders instead of representing the

most basal split within a monophyletic Oniscidea clade, thus casting doubt not only in the placement of Ligiidae within the Oniscidea but also in our historical understanding of the evolutionary history of the terrestrial isopods. Considering additional support is needed for these findings, in this study, we analyze the mitochondrial genomes (hereafter mitogenomes) of two *Ligia* species, as well as an additional 22 other isopod species, to explore the evolutionary path of the Oniscidea. Our goal is to clarify the phylogenetic relationships amongst the Oniscidea.

Results

Genome organization and base composition

The mitochondrial genome of *M. sinensis* and *L. exotica* are circular, double-stranded DNA moledules, 16,018 and 14,978 bp in length respectively (Figs. 1 & 2). The complete mitogenome of *M. sinensis* contains 34 mt genes: 13 protein-coding genes (PCGs), 19 tRNA genes, and two rRNA genes (Fig. 1). It lacks three tRNA genes: TrnA, trnE, trnL1. The average A + T content of the *M. sinensis* mt genome is approximately 75.32%, which is higher than other isopods (typical range: 54.4%-71.2). 18 overlapping regions and 13 intergenic regions are found in the genome. (Table 1). Nucleotide frequencies of all mt genes of *M. sinensis* are listed (Table 2).

Table 1
Gene content of the *Mongoloniscus sinensis*. Mitogenome

Feature	Strand[a]	Position	Length (bp)	Initiation Codon	Stop Codon	Anticodon	Intergenic nucleotide
cob	-	1–1,107	1107	ATG	TAA		64
nad5	+	1,172-2,884	1713	ATT	TAA		*
trn-F	+	2,885-2,952	68			GAA	-33[b]
trn-H	-	2,920-2,984	65			GTG	5
nad4	-	2,990-4,333	1344	ATG	TAA		*
nad4l	-	4,334-4,612	279	ATT	TAA		-5
trn-P	-	4,608-4,665	58			TGG	-4
nad6	+	4,662-5,153	492	ATT	TAA		-2
trn-S2	+	5,152-5,202	51			TGA	-9
rrnL	-	5,194-6,336	1143				21
trn-Q	-	6,358-6,424	67			TTG	47
trn-M	+	6,472-6,529	58			CAT	-16
nad2	+	6,514-7,518	1005	ATG	TAG		-10
trn-C	-	7,509-7,562	54			GCA	-18

*Gene borders are defned based on borders with adjacent genes.

[a]Plus strand (+)/minus strand (−).

[b]Negatve values represent overlapping nucleotdes.

Feature	Strand[a]	Position	Length (bp)	Initiation Codon	Stop Codon	Anticodon	Intergenic nucleotide
trn-Y	-	7,545-7,600	56			GTA	4
cox1	+	7,605-9,140	1536	ATG	TAA		1
trn-L2	+	9,142-9,208	67			TAA	-10
cox2	+	9,199-9,870	672	ATC	TAA		-2
trn-K	+	9,869-9,926	58			TTT	-13
trn-D	+	9,914-9,975	62			GTC	-8
atp8	+	9,968 – 10,123	156	ATT	TAA		-13
atp6	+	10,111 – 10,785	675	ATG	TAA		*
cox3	+	10,786 – 11,613	828	ATG	TAA		-41
trn-R	+	11,573 – 11,633	61			TCG	21
nad3	+	11,655 – 12,011	357	ATT	TAA		-41
trn-V	+	11,971 – 12,025	55			TAC	-9
nad1	-	12,017 – 12,940	924	ATA	TAA		1
trn-N	+	12,942 – 13,008	67			GTT	-2
rrnS	+	13,007– 13,723	717				-17

*Gene borders are defned based on borders with adjacent genes.

[a]Plus strand (+)/minus strand (−).

[b]Negatve values represent overlapping nucleotdes.

Feature	Strand[a]	Position	Length (bp)	Initiation Codon	Stop Codon	Anticodon	Intergenic nucleotide
trn-I	+	13707–13768	62			TCA	4
trn-W	+	13,773 – 13,826	54			TCA	292
trn-G	+	14,119 – 14,176	58			TCC	485
trn-T	+	14,662 – 14,725	64			TGT	879
trn-S1	+	15,605 – 15,677	73			TCT	341
*Gene borders are defned based on borders with adjacent genes.							
[a]Plus strand (+)/minus strand (−).							
[b]Negatve values represent overlapping nucleotdes.							

Table 2
Base composition of whole genome, protein-coding gene, rRNA

Region (strand)	A%	C%	G%	T%	A + T%	G + C%	AT skew	GC skew	+/-strand Isopoda ground pattern
Whole genome	37.13	10.68	14.00	38.19	75.32	24.68	-0.014	0.134	
cob (-)	30.17	15.97	11.47	42.46	72.63	27.37	-0.169	-0.162	-
nad5 (+)	33.80	7.82	15.47	42.91	76.71	23.29	-0.119	0.328	+
nad4 (-)	30.21	15.85	8.78	45.16	75.37	24.63	-0.198	-0.287	-
nad4l (-)	34.41	9.32	9.68	46.59	81.00	19.00	-0.150	**0.019**	-
nad6 (+)	33.74	6.71	10.98	48.58	82.32	17.68	-0.180	0.241	+
rrnL (-)	40.24	9.97	10.24	39.55	79.79	20.21	0.009	**0.013**	-
nad2 (+)	36.42	8.06	13.53	41.99	78.41	21.59	-0.071	0.230	+
cox1 (+)	27.93	14.52	17.19	40.36	68.29	31.71	-0.182	0.084	+
cox2 (+)	33.33	13.10	14.73	38.84	72.17	27.83	-0.076	0.059	+
atp8 (+)	37.18	10.90	8.97	42.95	80.13	19.87	-0.072	**-0.097**	+
atp6 (+)	32.00	11.26	14.81	41.93	73.93	26.07	-0.134	0.136	+
cox3 (+)	27.78	14.86	16.18	41.18	68.96	31.04	-0.194	0.043	+
nad3 (+)	31.37	9.24	14.85	44.54	75.91	24.09	-0.173	0.233	+

Region (strand)	A%	C%	G%	T%	A + T%	G + C%	AT skew	GC skew	+/-strand Isopoda ground pattern
nad1 (-)	30.30	13.1	12.88	43.72	74.03	25.97	-0.181	-0.008	-
rrnS (+)	36.12	13.39	18.13	32.36	68.48	31.52	0.055	0.150	+

The circular mitogenome of *Ligia exotica* is composed of 13 PCGs, 21 tRNA genes, two rRNA genes, one non-coding region, while only lacking the trnG gene (Fig. 2). The average A + T content of the *L. exotica* mt genome is approximately 59.13%. 17 overlapping regions and 13 intergenic regions are found in the genome (Table 3). Nucleotide frequencies of all mt genes of *M. sinensis* are listed (Table 4).

Table 3
Gene content of the *Ligia exotica*. Mitogenome

Feature	Strand[a]	Position	Length (bp)	Initiation Codon	Stop Codon	Anticodon	Intergenic nucleotide
trnE	+	663–724	62			TTC	
trnS1	+	725–787	63			TCT	17
cob	-	805-1,938	1134	ATA	TAA		*
trnT	-	1,939-1,997	59			TGT	-8[b]
nad5	+	1,990-3,717	1728	ATT	TAG		-8
trnF	+	3,710-3,768	59			GAA	-2
trnH	-	3,767-3,828	62			GTG	*
nad4	-	3,829-5,158	1330	ATG	T		-7
nad4l	-	5,152-5,448	297	ATA	TAA		6
trnP	-	5,455-5,516	62			TGG	1
nad6	+	5,518-6,024	507	ATT	TAG		-2
trnS2	+	6,023 – 6,084	62			TGA	*
rrnL	-	6,085 – 7,268	1184				-7
trnV	-	7,262-7,320	59			TAC	2
trnQ	-	7,323-7,377	55			TTG	4

*Gene borders are defned based on borders with adjacent genes.

[a]Plus strand (+)/minus strand (−).

[b]Negatve values represent overlapping nucleotdes.

Feature	Strand[a]	Position	Length (bp)	Initiation Codon	Stop Codon	Anticodon	Intergenic nucleotide
trnM	+	7,382-7,445	64			CAT	-21
nad2	+	7,425-8,441	1017	ATG	TAG		-15
trnC	-	8,427-8,479	53			GCA	-1
trnY	-	8,479-8,540	62			GTA	6
cox1	+	8,547 – 10,079	1533	CGA	TAA		-5
trnL2	+	10,075 – 10,136	62			TAA	*
cox2	+	10,137 – 10,820	684	ATA	TAG		-2
trnK	+	10,819 – 10,880	62			TTT	-2
trnD	+	10,879 – 10,938	60			GTC	9
atp8	+	10,948 – 11,097	150	ATA	TAA		-7

*Gene borders are defned based on borders with adjacent genes.

[a]Plus strand (+)/minus strand (−).

[b]Negatve values represent overlapping nucleotdes.

Feature	Strand[a]	Position	Length (bp)	Initiation Codon	Stop Codon	Anticodon	Intergenic nucleotide
atp6	+	11,091 – 11,762	672	ATG	TAA		-1
cox3	+	11,762 – 12,565	804	ATG	TAA		-17
trnR	+	12,549 – 12,608	60			TCG	9
nad3	+	12,618 – 12,962	345	ATT	TAG		-2
trnA	+	12,961 – 13,021	61			GCA	24
nad1	-	13,046 – 13,957	912	ATC	TTA		18
trnL1	-	13,976 – 14,035	60			TAG	-4
rrns	+	14,096 – 14,794	699				2
trnI	+	14,797 – 14,860	64			GAT	14
trnW	+	14,875 – 14,938	64			TCA	39

*Gene borders are defned based on borders with adjacent genes.

[a]Plus strand (+)/minus strand (−).

[b]Negatve values represent overlapping nucleotdes.

Table 4
Base composition of whole genome, protein-coding gene, rRNA

Region (strand)	A%	C%	G%	T%	A + T%	G + C%	AT skew	GC skew	+/-strand Isopoda ground pattern
Whole genome	28.29	18.04	22.83	30.85	59.13	40.87	-0.043	0.117	
cob (-)	24.34	24.96	17.02	33.69	58.02	41.98	-0.161	-0.189	-
nad5 (+)	25.87	15.39	24.59	34.14	60.01	39.99	-0.138	0.230	+
nad4 (-)	24.59	25.34	17.22	32.86	57.44	42.56	-0.144	-0.191	-
nad4l (-)	21.55	23.23	19.87	35.35	56.90	43.10	-0.243	-0.078	-
nad6 (+)	23.67	14.40	24.06	37.87	61.54	38.46	-0.231	0.251	+
rrnL (-)	32.18	19.51	17.91	30.41	62.58	37.42	0.028	-0.043	-
nad2 (+)	22.91	17.11	25.07	34.91	57.82	42.18	-0.207	0.189	+
cox1 (+)	22.96	19.70	22.37	34.96	57.93	42.07	-0.207	0.064	+
cox2 (+)	24.71	19.44	24.42	31.43	56.14	43.86	-0.120	0.113	+
atp8 (+)	30.00	12.67	22.00	35.33	65.33	34.67	-0.082	0.269	+
atp6 (+)	24.85	19.79	21.43	33.93	58.78	41.22	-0.154	0.040	+
cox3 (+)	22.01	21.14	24.13	21.71	54.73	45.27	-0.195	0.066	+
nad3 (+)	22.32	16.52	26.09	35.07	57.39	42.61	-0.222	0.224	+

Region (strand)	A%	C%	G%	T%	A + T%	G + C%	AT skew	GC skew	+/-strand Isopoda ground pattern
nad1 (-)	22.59	22.59	20.29	34.54	57.13	42.87	-0.209	-0.054	-
rrnS (+)	20.03	20.03	22.03	26.32	57.94	42.06	0.091	0.048	+

Protein-coding genes and codon usage

PCGs of *M. sinensis* are 11,088 bp in size, with its A + T content reaching 74.18%, the highest among all known Oniscidea. The ATG codon is the most commonly found start codon (six PCGs), with ATT next (found in nad3, nad4l, nad5, nad6, and atp8), and ATA only found in nad1. As for the terminal codons, TAA is found in 12 genes, and TAG only found in nad2 (Table 1).

The PCGs of *L. exotica* are 11113 bp in size, with its A + T content being 58.06%, much lower than *M. sinensis*. ATA is the start codon for four PCGs (cob, nad4l, cox2, nad8*)*, with ATG found in four PCGs (nad2, nad4, atp6, cox3), ATT in three (nad3, nad5, nad6), and ATC and CGA in one each (nad1 and cox1 respectively). As for terminal codons, TAA is found in 7 genes, TAG in five (cox2, nad2, nad3, nad5, nad6) and T found only in Nad4 (Table 3). The unfinished T codon is not counted separately, as we presumed that it would be completed (TAA) by posttranscriptional polyadenylation (Ojala et al., 1981; Schuster & Stern, 2009).

The two species exbibit the same protein-coding genes location as 9 genes (cox1-3, atp8, atp6, nad2-3, and nad5-6) are encoded by the plus strand and four genes (cob, nad1, nad4, and nad4L) by the minus strand (Table 2 and Table 4). Codon usage, RSCU, and codon family proportion (corresponding to the amino acid usage) of *M. sinensis* and *L. exotica* are investigated (Suppl. Materials 2). The four most abundant codon families of *M. sinensis* (Phe, Ile, Leu, and Ser) encompass 48.75% of all codon families; The four most abundant codon families of *L. exotica* (Gly, Leu, Ser and Val) encompass 43.17% of all codon families. Among these codon families, A + T-rich codons are favored over synonymous codons with lower A + T content in *M. sinensis* and *L. exotica*. These nonpolar hydrophobic amino acids and polar neutral amino acids occur so frequently that they are associated with most proteins encoded in the mitochondrial genome as transmembrane proteins.

Transfer and ribosomal RNA genes

The two rRNAs, rrnL and rrnS, of *M. sinensis* are 1143 and 717 bp in size, with 79.79% and 68.48% A + T content respectively. All 19 commonly found tRNAs are present in the mitochondrial genome of *M. sinensis*, ranging from 51 bp (trnS2) to 73 bp in size (trnS1), and adding up to 1158 bp in total combined

length. As for *L. exotica*, the two rRNAs, rrnL and rrnS, are 1184 and 699 bp in size, with 62.58% and 57.94% A + T content, respectively. All 21 commonly found tRNAs are present in the mitochondrial genome of *L. exotica*, ranging from 53 bp (trnC) to 64 bp in size (trnI, trnM, trnW), and adding up to 1278 bp in combined length. tRNA genes are distributed throughout the mitogenome and are found on both strands. The putative secondary structures of all identified tRNAs are shown in Fig. 3 and Fig. 4. The majority of tRNAs have a common t-shaped or clover-leaf secondary structure. Exceptions include the trnC of *M. sinensis*, where the DHU-arm is absent, and the trnG, trnK, trnP, trnS2, trnW and trnV, where the TΨC-arm is absent. In *L. exotica*, all of the secondary structures (predicted by MITOS and ARWEN) exhibit the conventional cloverleaf structure, except for trnC and trnV, which lack DHU arms and trnF, trnP, trnM, trnL2, trnK, trnD, trnR which lack the TΨC-arm.

Non-coding regions

In *M. sinensis*. there is one biggest non-coding region of 879 bp length located between trnT and trnS1. And other three non-coding regions are between trn-W and cob the length is 292 bp, 485 bp and 341 bp. We have not detected a hairpin structure in the mt control region of *M. sinensis* (Fig. 1). There are six discontinuous repeats of length 53 bp. As for *L. exotica*, one non-coding regions (NCR), 662 bp in size, are located between trnE and trnW. There are no tandem repeats of sections in this region and a GC-rich region containing the putative hairpin structure (Fig. 5).

Phylogeny and gene order

The phylogenies produced using BI and ML methods show concordant topologies; however, support values differ between approaches. BI analyses produced very high statistical support while the ML topology exhibited a mixture of mostly high but several lower support values (Fig. 7 and Fig. 8). Neither analyses recovered the monophyly of the Oniscidea. Though *M. sinensis* formed a clade with three species of Porcellionidae Brandt□Ratzeburg,1831 (Oniscidea) which was placed within a larger clade including all other Oniscidean species included in this study except *Ligia* species. But *L. exotica* formed a clade with *L. oceanica*, and the clade of two species of family Ligidae was not placed into the main clade consisting subfamily Oniscidea.

The gene order of the mt monomer of *M. sinensis* and *L. exotica* are shown in Fig. 6. Comparison with the putative isopoda ground pattern mt genome [11] revealed four gene rearrangements in *M. sinensis*, three gene rearrangements in *L. exotica*. As for *M. sinensis*, the first rearrangement trnR is between the cox3 and nad3, and it is between nad3 and nad1 of isopoda ground pattern. The second rearrangement, trnV, is between nad3 and nad1, but it is between 16S rRNA and nad3 in the isopoda ground pattern. The third rearrangement of trnT is between the 12S rRNA and cob, and it is between cob and nad5 in isopoda ground pattern. Another rearrangement of trnT was translocated near trnS1 and cob.

As for *L. exotica*, the first rearrangement trnR is between the cox3 and nad3, and it is between nad3 and nad1 of isopoda ground pattern; The second and third rearrangement trnW and trnE are interchanging. In a word, comparing with isopoda ground pattern, two isopods have gene translocation as other known mitochondrial genomes in isopod. Missing tRNA is found in all oniscideas including the present two species, and it is universal phenomenon. So, the order of mitochondrial genes in the Oniscidea is weakly conserved.

Discussion

The object of this work was to reconstruct the evolutionary relationships of the Oniscidea based on complete mitogenomes, with a particular focus on the Ligiidae as *Ligia* have been proposed to represent the intermediate forms in the evolution of terrestriality in isopods [7].

Phylogeny of Oniscidea based on the morphological characters or molecular data are debated [12, 18, 19, 20, 21]. The monophyly of Oniscidea has been well supported by morphological characters such as the complex water-conducting system and reduced first antenna with only three articles [6, 8, 9, 12, 22, 23]. The first antenna with rudimentary three articles has been regarded as a prominent synapomorphy for Oniscidea [6, 9], but some have suggested it to be a possible homoplasy [6]. Hornung (2011) found the water-conducting systems within Oniscidea are changeable[24]. So, the two main morphological characters to support the monophyly of Oniscidea actually be convergent related to the terrestrial environment challenges [12]. Indeed, monophyly of Oniscidea have been denied by various molecular data [15, 21, 25, 26, 27]. But further phylogenetic analyses of suborder Oniscidea and taxonomic rank and origin of Ligia spp. remain debated. Especially, Dimitriou et al., (2019) produced a phylogenetic tree using four nuclear genes and reported that *Ligia* does not form a monophyletic clade with the rest of the Oniscidea. We observe similar findings in this study based on mitochondrial genomes, as we found the suborder Oniscidea to be a polyphyletic group, as the 12 species of the Oniscidea did not converge into a single clade clade. The *L. exotica* and *L. oceanica*, although forming a clade themselves, were not placed with other Oniscidea, instead being shown to be most closely related to members of marine isopod suborders. Lins et al., (2017), Zou et al., (2018) and Hua et al., (2018) reported similar findings about *Ligia oceanica*, but they did not propose a taxonomic change to account for these observations. Similarly, Dimitriou et al., (2019) failed to recover the monophyly of *Ligia* + Oniscidea but did not suggest a taxonomic solution. So, here, we propose that the family Ligiidae should be elevated to its own suborder: Ligiaidea.

Our phylogenetic tree does indicate that the other ten species of oniscidea are indeed closely grouped together. Three species of Armadillidiidae Brandt,1833 and Porcellionidae Brandt & Ratzeburg,1831 are also closely clustered. Monophyly of Agnaridae is supported by two mitochondrial and three nuclear genetic markers [28], and it is closely related to Porcellionidae, rather than Armadillidiidae. *M. sinensis* formed a clade with three species of Porcellionidae (Oniscidea), and three species of Armadillidiidae formed another clade. So, our findings are in agreement with above research, but add further evidence to definite the relationship between families of Oniscidea. *Oniscus asellus*, with its longer horizontal

branches, morphologically with much flatter body and more varied body than other isopods. With regards to *M. sinensis*, they formed a clade with three species of Porcellionidae Brandt & Ratzeburg,1831 and were indeed much closer in morphology and life environment. So, *Trachelipus rathkii* clustered with *Cylisticus convexus*, and two different heterogeneous nodes are found, by changing the anti-codon of tRNA to produce additional double tRNA genes, the results are consistent in both species. This may be the reason why two species have always been one clade. So, monophyly of Crinocheta is strongly supported, and the close relatives of Ligiaidea are Valvifera + Cymothoida + Sphaeromatidea.

The sequence of mitochondrial genomes of *M. sinensis* and *L. exotica* are highly derived compared with other known isopod species pattern. All 24 oniscideans studied to date have missing tRNAs (see Suppl. Materials 1), possibly due to the short length of tRNA. We report the complete absence of the DHU arm for the TrnC of *M. sinensis* and trnC, trnV of *L. exotica*. Complete absence of tRNA DHU arm is also found in other isopods, for instance the trnC and trnV in *L. oceanica*.

Doublet et al., (2012) and Chandler et al., (2015) pointed out mitogenomes of isopods especially oniscideans with unique linear/circular organization [27, 29]. It may be associated with an ancient and conserved constitute heteroplasmic sites, or faciliate the evolution of stable mitochondrial heteroplasmies, and then constrain the multimeric structure of mitogenomes. The heteroplasmic site alters the anticodon in tRNA genes, allow it to code for different tRNA. As for *M. sinensis* and *L. exotica*, some tRNA genes appear to be missing not by annotation software or RNA editing, and it is usual in isopods (Suppl. Materials. 1) [10, 30].

In *Ligia exotica*, *Ligia oceanica* and *Idotea baltica*, there is a positive GC skew in the plus strand coding gene and a negative GC skew in the minus strand. Most of the genes in *M. sinensis* follow the above rules, and only some of the genes are reversed, such as ad4l and trn-l are in the negative chain but their GC is positively skew, and atp8 is in the positive chain but its GC is negatively skew. This is common in most other crustaceans, probably because of the reversal of the mitochondrial gene replication region.

Conclusions

The family Ligiidae with its own suborder Ligiaidea is proposed. Monophyly of Agnaridae is strong supported by the complete mitochondrial, and it is closely related to Porcellionidae rather than Armadillidiidae. Overall, the conquest of the land is one of the major events in the history of the life on Earth, so the key groups from the ocean to land should attract much more attention. In this study we add evidence to the growing consensus behind the polyphyly of Oniscidea. Thus, taxonomic revision of this group should be urgently considered. Mitogenomes and nuclear genes of more species such as genera including *Caucasoligidium*, *Ligidoides*, *Ligidium*, *Tauroligidium*, *Typhloligidium* and other current and previous members of the Ligiidae family should be sequenced. Morphological and ecological characters of isopods should be analyzed with multi-evidences.

Methods

Sample and DNA extraction

As only eight isopods have their complete mitogenomes and other 14 isopods with incomplete mitogenomes reported in public databases (Table 1) [29, 30, 31], we generated the entire mitogenome for an additional two species including *Mongoloniscus sinensis* and *Ligia exotica*. Specimens for the form*er* were collected by An and Zhao on 5 June 2017, from Tao Village in Yantai, Shandong province of China (37°28.8N,121°27.6'E); while *L. exotica* specimens were collected by An and Zhang on 10 June 2017 from the first bathing beach in Qingdao, Shandong province of China (36°18'N,120°06'E). All specimens used for DNA extraction were stored in Ethanol (> 99.7%). Total genomic DNA was isolated from pereopods/pleopods using a gDNA rapid extraction kit (Aidlab Biotechnologies Co., Ltd). Voucher specimens are deposited in Shanxi Normal University under the collection number CIO201706050006 *(Mongoloniscus sinensis)* and CIO201706100009 *(Ligia exotica)*.

Genome determination, gene annotation and sequence analysis

Libraries were constructed using Whole Genome Shotgun (WGS) strategies with libraries sequenced using paired-end (PE) approaches on the Illumina MiSeq sequencing platform. Raw data was quality tested using the Read Quality Inspection Tool FastQC (http://www.bioinformatics.ac.uk/projects/fastqc). The data quality assessment was carried out by inspecting the single base quality, base content distribution, GC content distribution, and sequence base quality. High-quality second-generation sequencing data was assembled by using A5-miseq v20150522 and SPAdesv3.9.0 to construct contig and scaffold sequences. Newly produced sequences were extracted according to the sequenced depth of splicing-sequence using blastn (BLAST v2.2.31+) in nt bank of NCBI to the mitochondrial sequence. The mitochondrial spliced results obtained by the above different software were co-linearly analyzed by mummer v3.1 software, and the positional relationship of contigs was determined, with gaps between filled.

Results were corrected with the pilon v1.18 software to obtain the final mitochondrial sequence. Pilon significantly improves draft genome assemblies by correcting bases, fixing mis-assemblies and filling gaps. Mitogenomes were annotated using the MITOS webserver [32] with tRNAs re-detected using two computer programs: tRNAscan-SE v1.21 [33] and ARWEN v1.2.3.c [34]. tRNAs were confirmed by performing blast searches in MitoZoa 2.0 [35] CGView [36] used to display the circular mitogenome of the two focus species, while nucleotide composition was analyzed with MEGA 5. The complete genome sequence has been submitted to NCBI (GenBank: MG709492 and MK028672).

Phylogenetic analysis of mt gene sequences

The 13 protein-coding genes for twenty-four available isopod species were used for phylogenetic analyses. An additional six decapod species were used as outgroups: *Alvinocaris longirostris* Kikuchi &

Ohta, 1995, *Halocaridina rubra* Holthuis, 1963, *Panulirus japonicus* von Siebold, 1824, *Shinkaia crosnieri* Baba & Williams, 1998, *Geothelphusa dehaani* White, 1847, *Austinograea rodriguezensis* Tsuchida & Hashimoto, 2002. Nucleotide sequence alignments were built for all 13 protein-coding genes individually using MEGA 5.1 [37]; and then concatenated to a single alignment. Modeltest software [38] and PhyML 3.0 [39] webserver were used to determine the optimal nucleotide sequence model for phylogenetic analysis (evaluated according to the Akaike information criterion), with the GTR + G + I model (general time reversible model with a proportion of invariant sites and gamma distributed rate variation across sites) found to be most appropriate. Maximum likelihood phylogenetic reconstructions were performed using the PhyML 3.0 webserver, with support for phylogenetic relationships established by carrying out 1,000 bootstrap replicates. Bayesian inference reconstructions were performed in MrBayes 3.1.2 [40] assuming a fixed nucleotide sequence GTR + G + I model, a randomly chosen starting tree, and four Monte Carlo Markov chains run for 10,000,000 generations with trees sampled every 1,000 generations. The initial 25% of the trees are discarded as 'burn-in'.

Abbreviations

bp: base pair

DNA: Deoxyribonucleic acid

g DNA: genome Deoxyribonucleic acid

L.exotica: *Ligia exotica*

M.sinensis: *Mongoloniscus sinensis*

NCBI: National Center of Biotechnology Information

NCR: non-coding regions

NGS: Next Generation Sequencing

PCGs: protein-coding genes

PE: paired-end

rRNA: ribosomal Ribonucleic Acid

rrnL: 16s ribosomal Ribonucleic Acid

rrns: 12s ribosomal Ribonucleic Acid

RSCU: relative synonymous codon usage

tRNA: Transfer Ribonucleic Acid

WGS: Whole Genome Shotgun

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Genetic data used in the present study are deposited at GenBank and publicly accessible through the provided accession numbers, and other relevant data are presented within paper and Suppl. Materials.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (No. 31471970); Program of Ministry of Science and Technology of the People”s Republic of China (2015FY210300) and Natural Science Foundation of Shanxi Province (No. 201901D111274).

Author’s contributions

JA: conceived, designed and supervised the study, obtained samples, and revised the manuscript;

RZ analyzed and interpreted the data and drafted a manuscript;

RC: performed the experiments and contributed materials/analysis tools;

CAS: revised the manuscript.

All authors read and approved the final manuscript.

Acknowledgements

We are indebted to all collectors of the present specimens.

References

1. Wilson GD. Global diversity of isopod crustaceans (Crustacea; Isopoda) in freshwater. Hydrobiologia. 2008; 595(1): 231-240.
2. Wilson GD. The phylogenetic position of the Isopoda in the Peracarida (Crustacea: Malacostraca). Arthropod Systematics & Phylogeny. 2009; 67: 159-198.
3. Schmalfuss H. World catalog of terrestrial isopods (Isopoda: Oniscidea). Stuttgar Beiträge zur Naturkunde, Serie A. 2003; 654, 1-341.
4. Sfenthourakis S, Taiti S. Patterns of taxonomic diversity among terrestrial isopods. Zookeys. 2015; 515: 13-25.
5. Schmalfuss H. Phylogenetics in Oniscidea. Monit. Zool. Ital. Monogr. 1989; 4: 3-27.
6. Schmidt C. Phylogeny of the terrestrial Isopoda (Oniscidea): a review. Arthropod Systematics & Phylogeny. 2008; 66: 191-226.
7. Carefoot T, Taylor B. Ligia: a prototypal terrestrial isopod. Crustacean. 1995; 47-60.
8. Tabacaru I, Danielopol D. Phylogénie des Isopodes terrestres. Comptes rendus de l'Académie des sciences. Série 3, Sciences de la vie. 1996;319:71-80.
9. Wägele JW. Evolution und phylogenetisches System der Isopoda. Stand der Forschung und neue Erkenntnisse. Zoologica. 1989;140:1-262.
10. Hua CJ, Li WX, Zhang D, Zou H, Li M, Jakovlić I, et al. Basal position of two new complete mitochondrial genomes of parasitic Cymothoida (Crustacea: Isopoda) challenges the monophyly of the suborder and phylogeny of the entire order. Parasites & vectors. 2018;11,628.
11. Kilpert F, Held C, Podsiadlowski L. Multiple rearrangements in mitochondrial genomes of Isopoda and phylogenetic implications. Molecular phylogenetics and evolution. 2012;64:106-117.
12. Lins LS, Ho SY, Lo N. An evolutionary timescale for terrestrial isopods and a lack of molecular support for the monophyly of Oniscidea (Crustacea: Isopoda). Organisms Diversity & Evolution. 2017;17:813-820.
13. 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. Ecology and Evolution. 2017;7:1869-1881.
14. Wetzer R, Perez LM, Bruce NL. Phylogenetic relationships of the family Sphaeromatidae Latreille, 1825 (Crustacea: Peracarida: Isopoda) within Sphaeromatidea based on 18S-rDNA molecular data. Zootaxa. 2013;3599:161-177.
15. Zhang D, Zou H, Hua CJ, Li WX, Mahboob S, Al-Ghanim KA., et al. Mitochondrial architecture rearrangements produce asymmetrical nonadaptive mutational pressures that subvert the phylogenetic reconstruction in Isopoda. Genome Biol. Evol. 2019;11:1797-1812.

16. Ojala D, Montoya J, Attardi G. Trna punctuation model of rna processing in human mitochondria. Nature. 1981;290(5806):470-474.
17. Schuster G, Stern D. RNA polyadenylation and decay in mitochondria and chloroplasts.[J]. Progress in molecular biology and translational science. 2009;85.
18. Schmalfuss H. The terrestrial isopod genus Schizidium (Isopoda: Oniscidea): systematics, distribution, morphology. Stuttgar Beiträge zur Naturkunde. 2008;1:143-151.
19. Lins LSF, Ho SYW, Wilson GDF, Lo N. Evidence for Permo-Triassic colonization of the deep sea by isopods. Biology letters. 2012;8:979-982.
20. Hurtado LA., Mateos M, Wang C., Santamaria CA, Jung J, Khalaji-Pirbalouty V, et al. Out of asia: mitochondrial evolutionary history of the globally introduced supralittoral isopod ligia exotica. Peerj. 2018;6(3):e4337.
21. Dimitriou AC, Taiti S, Sfenthourakis S. Genetic evidence against monophyly of Oniscidea implies a need to revise scenarios for the origin of terrestrial isopods. Scientific Reports. 2019;9:18508.
22. Brusca RC, Wilson GDF. A phylogenetic analysis of the Isopoda with some classificatory recommendations. Mem Queensl Mus. 1991;31:143-204.
23. Erhard F. Das pleonale Skelet-Muskel-System von Titanethes albus (Synocheta) und weiterer Taxa der Oniscidea (Isopoda), mit Schlussfolgerungen zur Phylogenie der Landasseln. Stuttgarter Beiträge zur Naturkunde A. 1997;550:1-77.
24. Hornung E. Evolutionary adaptation of oniscidean isopods to terrestrial life: structure, physiology and behavior. Terrestrial Arthropod Reviews. 2011;4(2):95-130.
25. Michel-Salzat A, Bouchon D. Phylogenetic analysis of mitochondrial lsu rrna in oniscids. Comptes Rendus de l'Academie des Sciences Serie III Sciences de la Vie. 2000;323(9):0-837.
26. Zou H, Jakovlić I, Zhang D, Hua CJ, Chen R, Li WX, et al. Architectural instability, inverted skews and mitochondrial phylogenomics of Isopoda: outgroup choice affects the long-branch attraction artefacts. Royal Society Open Science. 2000;7(2):191887.
27. Doublet V, Raimond R, Grandjean F, Lafitte A, Souty-Grosset C, Marcadé I. Widespread atypical mitochondrial DNA structure in isopods (Crustacea, Peracarida) related to a constitutive heteroplasmy in terrestrial species. 2012;55(3):234-44.
28. Dimitriou A, Taiti S, Schmalfuss H, Schmalfuss H.. A molecular phylogeny of Porcellionidae (Isopoda, Oniscidea) reveals inconsistencies with present taxonomy. 2018;801:163-176.
29. Chandler CH, Badawi MM, Mouziane B, Grève P, Cordaux R. Multiple Conserved Heteroplasmic Sites in tRNA Genes in the Mitochondrial Genomes of Terrestrial Isopods (Oniscidea). 2015;5(7):1317-22.
30. 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.
31. Marcadé I, Cordaux R, Doublet V, Debenest C, Bouchon D, Raimond R. Structure and evolution of the atypical mitochondrial genome of Armadillidium vulgare (Isopoda, Crustacea). 2007;65(6):651-9.

32. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. Molecular phylogenetics and evolution. 2013;69:313-319.
33. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic acids research. 1997;25:955-964.
34. Laslett D, Canbäck B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics. 2007;24:172-175.
35. D'Onorio de Meo P, D'Antonio M, Griggio F, Lupi R, Borsani M, Pavesi G, et al. MitoZoa 2.0: a database resource and search tools for comparative and evolutionary analyses of mitochondrial genomes in Metazoa. Nucleic acids research. 2011;40:1168-1172.
36. Stothard P, Wishart DS. Circular genome visualization and exploration using CGView.Bioinformatics. 205; 21: 537-539.
37. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences. 2004; 101: 11030-11035.
38. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics (Oxford, England). 1998; 14: 817-818.
39. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology. 2003; 52: 696-704.
40. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003; 19: 1572-1574.

Figures

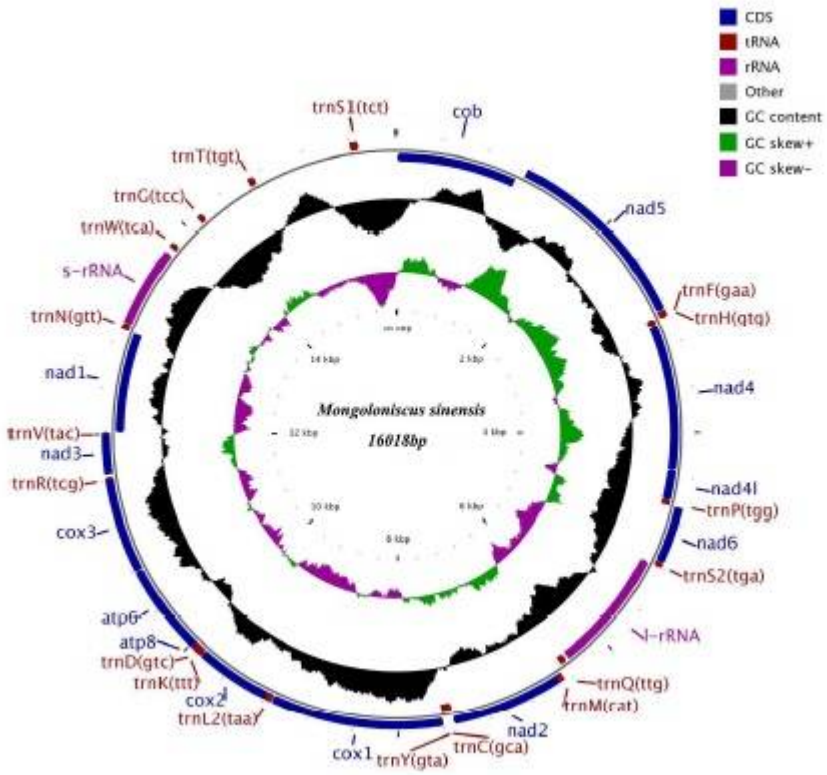

Figure 1

Map of the Mongoloniscus sinensis. mitogenome. From inside to outside, the first circle represents the scale; the second GC skew (-); the third GC content; the fourth the protein-coding gene, tRNA, rRNA and its arrangement on the genome.

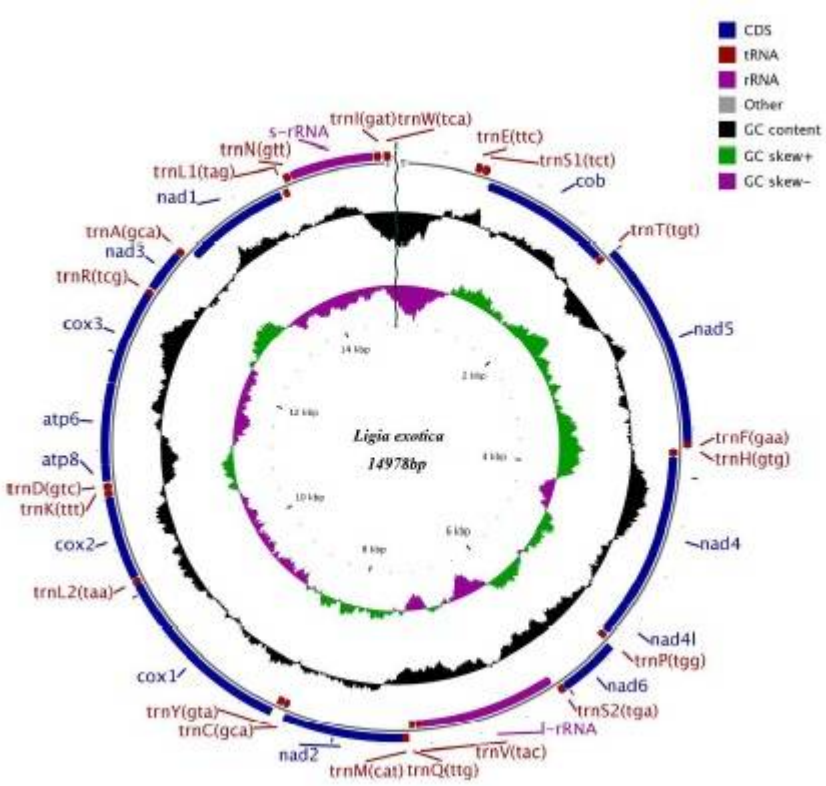

Figure 2

Map of the Ligia exotica. mitogenome. From inside to outside, the first circle represents the scale; the second GC skew (-); the third GC content; the fourth the protein-coding gene, tRNA, rRNA and its arrangement on the genome.

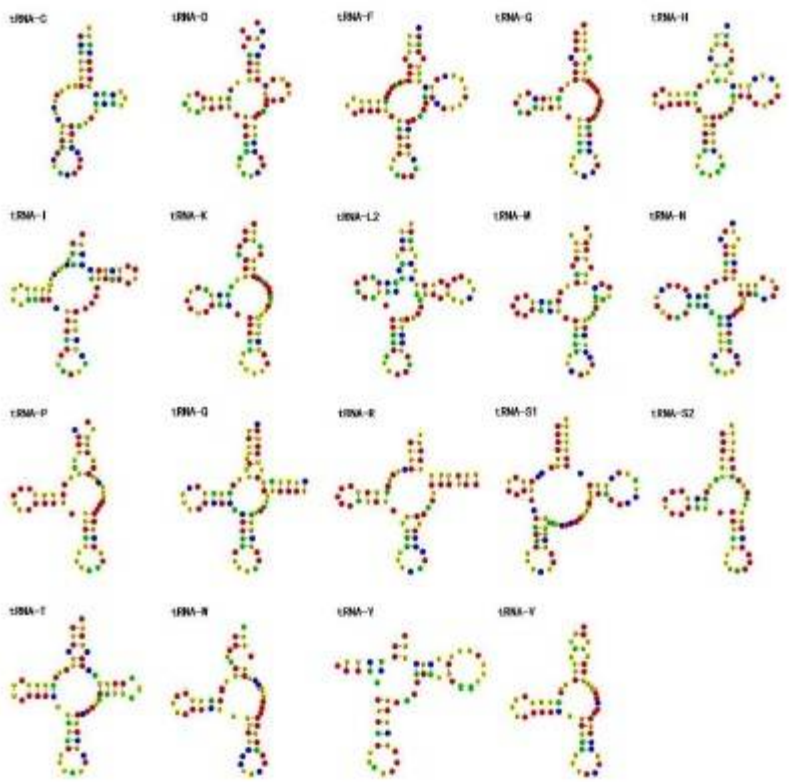

Figure 3

Putative secondary structures for the 19 transfer RNAs of the Mongoloniscus sinensis. mitogenome. All 19 structures were generated using the MITOS webserver and verified using tRNAscan-SE and ARWEN. Most tRNAs feature a standard clover-leaf structure. Exceptions: The TΨC arm is absent from trnG, trnK, trnP, trnS2, trnW, trnV and the DHU arm is absent from trnC

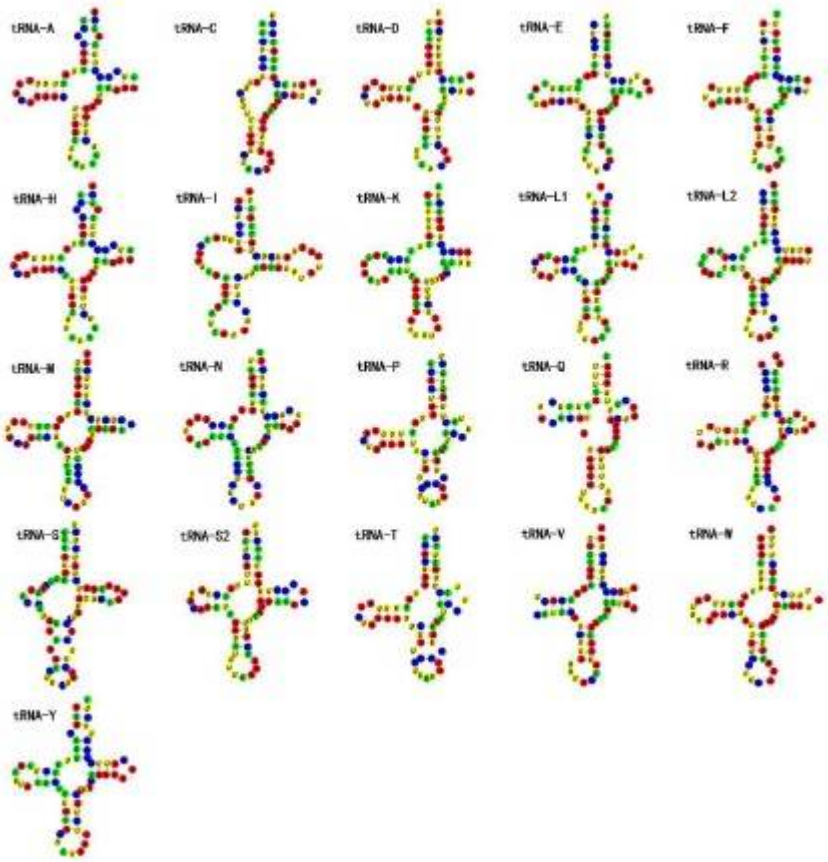

Figure 4

Putative secondary structures for the 21 transfer RNAs of the Ligia exotica. mitogenome. All 21structures were generated using the MITOS webserver and verified using tRNAscan-SE and ARWEN. Most tRNAs feature a standard clover-leaf structure. Exceptions: The TΨC arm is absent from trnF, trnP, trnM, trnL2, trnK, trnD,trnR and the DHU arm is absent from trnC,trnV

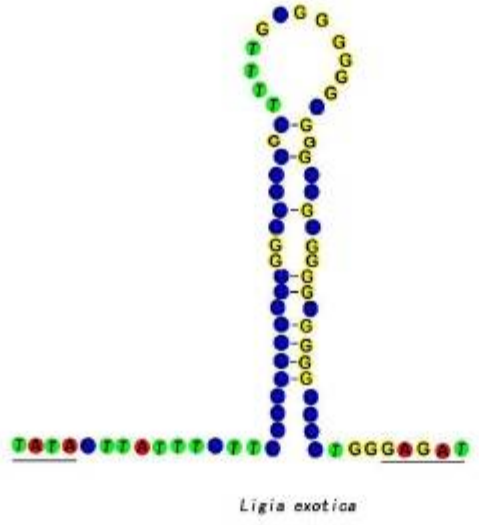

Figure 5

Hairpin structures in mitochondrial control regions of Ligia exotica (Isoppda).

Figure 6

Gene order in isopod mitogenomes.

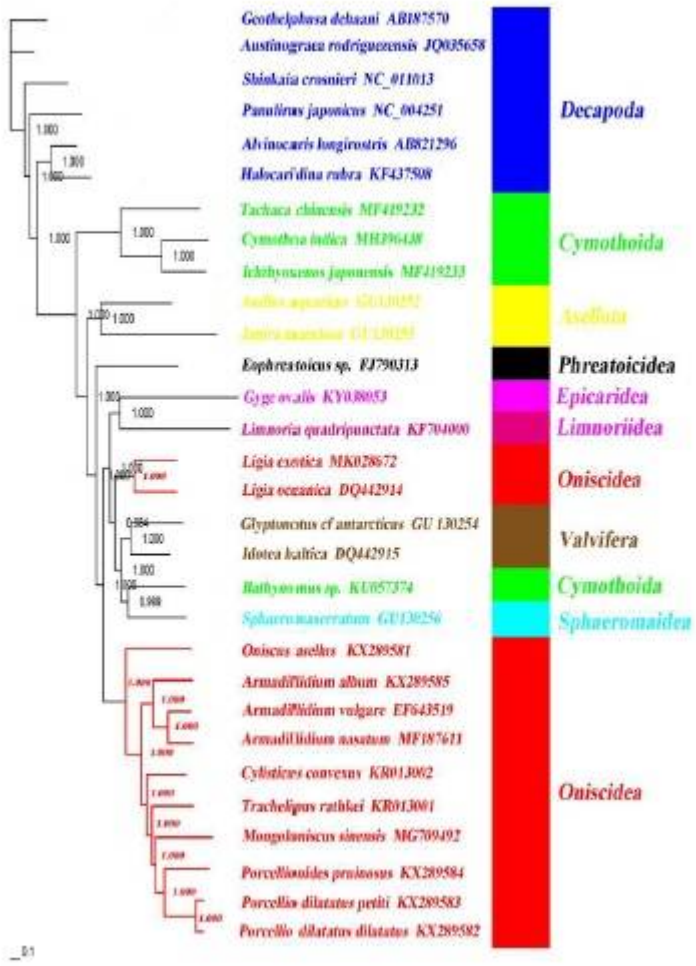

Figure 7

Mitochondrial phylogenomics of Isopoda: Bayesian inference analysis. The analysis is conducted using nucleotide sequences of all PCGs. 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.

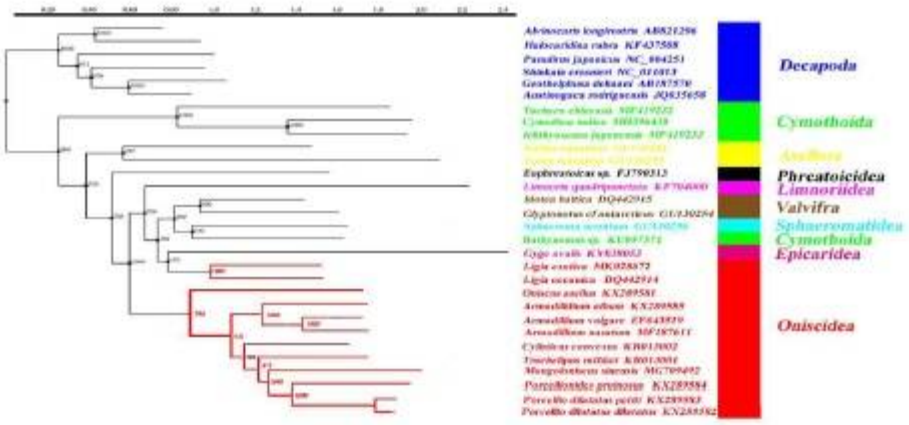

Figure 8

Mitochondrial phylogenomics of Isopoda: Maximum likelihood analysis. The analysis is conducted using nucleotide sequences of all PCGs. Scale bar corresponds to the estimated number of substitutions per site. Bootstrap support values (lower than 1000) are shown next to corresponding nodes. GenBank accession numbers are shown next to species names. Taxonomic rank (suborder/superfamily) is shown to the right.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Materials2.pdf
- Materials1.pdf