Gene 777 (2021) 145467

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Complete mitochondrial genomes of *Epeorus carinatus* and *E. dayongensis* (Ephemeroptera: Heptageniidae): Genomic comparison and phylogenetic inference

Wei Zhang, Ran Li, Changfa Zhou *

The Key Laboratory of Jiangsu Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing 210023, China

ARTICLE INFO

Keywords:
Mitochondrial genome
Heptageniid mayfly
Epeorus
Comparative analysis
Phylogeny

ABSTRACT

The current research on Ephemeroptera is mainly based on its morphology, since only small numbers of mitogenomes have been reported. In this study, the mitogenomes of *Epeorus carinatus* (15,338 bp) and *E. dayongensis* (15,609 bp) were sequenced, annotated and compared to genome data from congeners. Both mitogenomes had 23 tRNA genes including standard 22 and one extra $tRNA^{Met}$. The duplicated $tRNA^{Met}$ gene had been found in other heptageniid species except *Paegniodes cupulatus*, suggesting it could be used as a molecular synapomorphy for partial Heptageniidae. The phylogenetic analyses based on Bayesian Inference (BI) and Maximum Likelihood (ML) showed that Heptageniidae was monophyletic and the relationships among known *Epeorus* species were ((*E. carinatus* + *E. herklotsi*) + (*E. dayongensis* + *E.* sp. 1)), which implied the focal species *E. carinatus* and *E. dayongensis* should be grouped into different subgenera.

1. Introduction

The mitochondrial genome (mitogenome) of insects is typically a circular molecule 14–20 kb in length. It contains 37 genes, including 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), two ribosomal RNA genes (rRNAs), and a large non-coding region (AT-rich region) which contains initiation sites for transcription and replication of the genome (Wolstenholme, 1992; Boore, 1999; Cameron, 2014). Due to the advantages of its small size, almost complete maternal inheritance, rapid evolution rate and lack of introns, mitogenomes are widely used in comparative and evolutionary genomics, molecular evolution, population genetics and species identifications (Boore, 1999; Cameron, 2014).

Heptageniidae is the third largest family of Ephemeroptera, and includes 37 genera with more than 600 species. Previous taxonomic and phylogenetic analyses were based almost exclusively on morphological characters (Barber-James et al., 2008; Webb and McCafferty, 2008; Sartori, 2014a, 2014b, 2014c, 2014d; Yanai et al., 2017). Although some mayfly taxonomists have utilized molecular data to investigate this family in recent years, only a few molecular markers such as cytochrome *c* oxidase I (COI), ribosomal RNA genes (12S rRNA gene, 16S rRNA gene, 28S rRNA gene) and Histone (H3) have been adopted individually or together (Yanai et al., 2017; Ogden et al., 2005, 2009; Vuataz et al., 2013). Currently, only five complete mitogenomes of this family have been reported (22 mitogenomes for Ephemeroptera as a whole) (Zhou and Braasch, 2003; Eaton, 1871; Hsu, 1936; Zhang et al., 2008; Zhou et al., 2016; Tang et al., 2014; Gao et al., 2018).

The heptageniid genus *Epeorus* includes 93 species distributed around the world and is the second largest genus in the family Heptageniidae (Satori et al., 2015). Historically, this genus was morphologically divided into seven subgenera (Wang and McCafferty, 2004; Kluge, 2004; Braasch, 2006; Chen et al., 2010; Kluge and Tiunova, 1989; Boonsoong and Braasch, 2013; Hrivniak et al., 2017, 2019). However, the validity and monophyly of those subgenera were very controversial. Some questions, like whether the *Iron* was a subgenus or a valid genus, or whether some other subgenera (i.e. *Belovius*, *Ironopsis*, *Caucasiron*) existed or not have been debated for a long time (Chen et al., 2010; Kluge and Tiunova, 1989; Boonsoong and Braasch, 2013; Hrivniak et al., 2019). More molecular data, especially mitogenomes, will be helpful to test those theories.

In this study, the mitogenomes of *Epeorus carinatus* (a new record from China) and *E. dayongensis* were further analyzed and compared to that of *E. herklotsi* in GenBank, in an attempt to identify the evolutionary character of the *Epeorus* mitogenomes. In addition, phylogenetic

* Corresponding author.
E-mail address: zhouchangfa@njnu.edu.cn (C. Zhou).

https://doi.org/10.1016/j.gene.2021.145467
Received 31 July 2020; Received in revised form 12 January 2021; Accepted 22 January 2021
Available online 29 January 2021

analyses based on 13 PCGs were performed to test previous topologies of Heptageniidae. Hopefully, the new mitogenome information in this study will contribute to and shed more light on further studies on Ephemeroptera.

2. Materials and methods

2.1. Sample collection, morphological identification and DNA extraction

Specimens of *E. carinatus* were collected in Fan-jing Mountain, Guizhou Province, 2019-VIII-9 and *E. dayongensis* specimens were collected in Lei-shan county, Guizhou Province, 2019-VIII-6. Nymphs were collected in running water by hand nets, and imagoes were attracted by lights. All materials were stored in ethanol (more than 95%) and inspected under Nikon SMZ 645 or SMZ 1500 stereomicroscopes. After carefully identification (Braasch and Soldán, 1984; Gui and Zhang, 1992), these specimens were stored in anhydrous ethanol immediately and stored at –20℃ until DNA extraction. Total DNA was isolated using the Trelief™ Animal Genomic DNA Kit (TSINGKE Biotech, China) following the manufacturer's instructions. The voucher specimen numbers of *E. carinatus* (voucher number: NNU–EP1235) and *E. dayongensis* (voucher number: NNU–EP1283) are 12 and 18 individuals. All specimens used in this study are deposited in the Mayfly collection, College of Life Sciences, Nanjing Normal University.

2.2. Mitogenome sequencing and assembly

For library preparation, Illumina TruSeq® DNA Sample Prep Kit (Illumina, USA) was used, with an average insert size of 350 bp. Whole genomic sequencings were performed using an Illumina HiSeq X Ten platform, with 150 bp pair-ended reads. Samples were sequenced together with other projects and two lanes were used for each mitogenome. More than 2 GB raw reads (deposited in NCBI SRA, SAMN16427085 and SAMN16427086) of each mitogenome were trimmed of adapter contamination using NGC QC Toolkit (Patel et al., 2012) and low quality and short reads were removed by Prineseq (Schmieder and Edwards, 2011). The final data set was *de novo* assembled using Velvet 1.2.10 imbedded in Geneious R11, with the following parameter settings: Minimum Overlap = 30–50 bp, Minimum Overlap Identity: 80–100 bp, Maximum Gap Size = 2 bp, Maximum mismatches Per Read = 2% (Zerbino and Birney, 2008).

2.3. Mitogenome annotation and analysis

The coding regions of 13 PCGs were identified using the NCBI ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/) using the invertebrate mitochondrial genetic code, then translated into amino acid sequences using MEGA 7 (Kumar et al., 2016). tRNA genes were identified by tRNAscan-SE program and MITOS Web Server (Lowe and Eddy, 1997; Bernt et al., 2013). rRNA genes were determined by MITOS Web Server (Bernt et al., 2013). Intergenic spacers and overlapping regions between genes were estimated manually.

Base composition, relative synonymous codon usage (RSCU) and nucleotide substitution statistics were analyzed with MEGA 7. The bias of nucleotide composition was measured as AT-skew = [A – T] / [A + T] and GC-skew = [G – C] / [G + C] (Perna and Kocher, 1995).

2.4. Phylogenetic analysis

We selected 22 species of Ephemeroptera as the ingroup, including the two newly sequenced *Epeorus* species, and two species of Archaeognatha as outgroups, which is evolutionarily older than Ephemeroptera (Table S1). All 13 protein sequences were employed to perform phylogenetic analyses. PCGs were aligned using MAFFT 7 online server with the G-INS-i strategy (Katoh and Standley, 2013). Nucleotide saturation was tested in DAMBE 5 (Xia and Xie, 2001). The individual alignment fragments were then concatenated using the software Geneious 10.1.3. The PCG data was created to test the influence of the gene sequences. The best partitioning scheme and corresponding nucleotide substitution model for the PCG data were selected by PartitionFinder (Lanfear et al., 2012) with Bayesian Information Criterion (BIC). Both Maximum likelihood (ML) and Bayesian inference (BI) were employed for phylogenetic analyses. The ML analysis was performed with RAxMLHPC2 on XSEDE 8.0.0 (Stamatakis, 2014) through the CIPRES Science Gateway (Miller et al., 2010) and the nucleotide substitution model used for ML was GTRGAMMAI. Support for clades was assessed with bootstrap 1000 replicates. The BI analysis was performed using MrBayes 3.2.6 (Ronquist et al., 2012) also on CIPRES and the best model was listed in Table S2, two simultaneous runs with four chains (one cold chain and three hot chains) for 10 million generations, sampling every 1000 trees. The first 25% samples were discarded as burn-in.

3. Results and discussion

3.1. General characters of mitogenomes

The complete mitogenome sequences of the two species were 15,338 bp (*E. carinatus*) and 15,609 bp (*E. dayongensis*) in size (GenBank accession numbers: MT112896 and MT112895). Both sequences contained 13 PCGs (COI–III, ND1–6, ND4L, Cytb, ATP6 and ATP8), 23 tRNA genes (one for each amino acid, two for Leucine and Serine, and an extra Methionine), two rRNA genes (12S and 16S rRNA) and an AT-rich region (Figs. 1–2). Among them, 4 PCGs, 8 tRNA genes and two rRNA genes were encoded on the minority strand (N-strand) while another 24 genes were encoded on the majority strand (J-strand) (Figs. 1–2, Table 1). The arrangement and orientation of the two mitogenomes was identical to the ancestral gene order (Cameron, 2014), except that each had an extra tRNA ($tRNA^{Met}$), which had already been recorded in *E. herklotsi* (Fig. 2).

In the mitogenome of *E. carinatus*, there were 12 non-coding regions with a total length of 568 bp and each non-coding region ranged from 1 to 490 bp, 14 overlap regions were present and with 37 bp in length totally and each overlap region ranged from 1 to 8 bp. Comparatively, *E. dayongensis* had 13 non-coding regions with totally 853 bp in length and each non-coding region ranged from 1 to 765 bp, 13 overlap regions with totally 34 bp in length and ranging from 1 to 8 bp (Table 1). In addition, the complete mitogenome of *E. herklotsi* was 15,502 bp, with 12 non-coding regions (725 bp) and ranging from 1 to 634 bp, 13 overlap regions (34 bp) and ranging from 1 to 8 bp. Remarkably, overlaps (4 bp) were detected in the junction between ATP8 and ATP6 in the mitogenomes of three species, although the length of this region was not 7 bp. Based on the hypothesis of Lavrov, the overlaps between genes may be a product of the selective pressure to reduce genomes size noted in mitochondrial (Lavrov and Brown, 2001).

3.2. Nucleotide composition of the mitogenome

The nucleotide composition of the three mitogenomes was compared in Table 2. The overall A + T content of them was 64.65% (*E. carinatus*), 67.2% (*E. dayongensis*) and 65.67% (*E. herklotsi*), and each species showed similar A and T nucleotides biases. The A + T content composition in regions of *E. carinatus* was 64.07% (PCGs), 64.49% (tRNAs), 66.23% (rRNAs) and 72.25% (AT-rich region), in *E. dayongensis* was 66.26% (PCGs), 67.4% (tRNAs), 67.8% (rRNAs) and 77.38% (AT-rich region), and in *E. herklotsi* was 65% (PCGs), 65.75% (tRNAs), 65.78% (rRNAs) and 74.45% (AT-rich region). In addition, skew metrics of the mitogenomes showed negative AT-skew (-0.005 to −0.002) and GC-skew (−0.246 to −0.208), indicating that Ts and Cs were more abundant than As and Gs. Just as mentioned before (Perna and Kocher, 1995; Carapelli et al., 2007; Xu et al., 2020), the majority strand showed negative AT-skews in the all known mitogenomes of Ephemeroptera, differing from the majority of hexapod species, which showed positive AT-skews. This suggested that a special strand asymmetry reverse

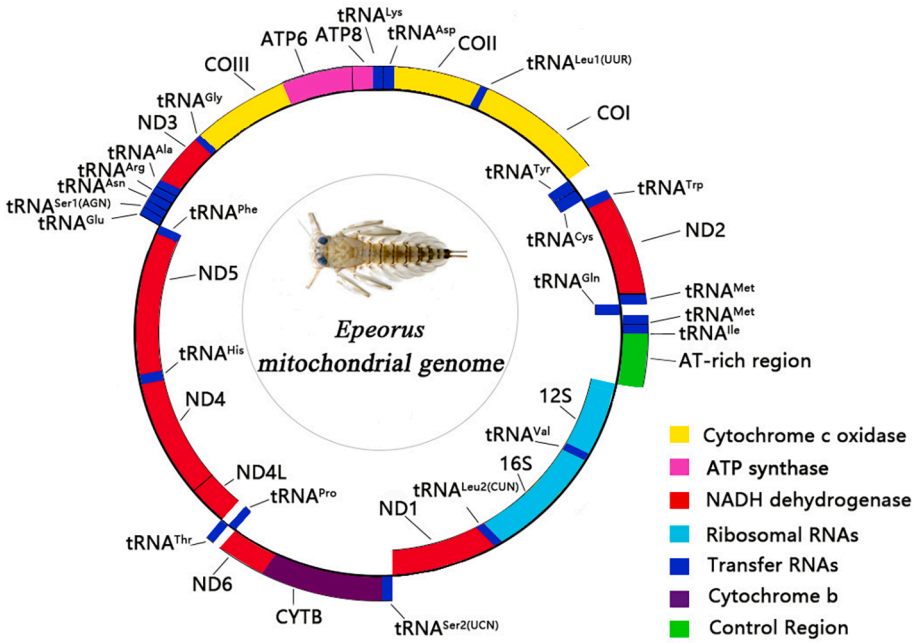

Fig. 1. Mitochondrial map of *Epeorus* genes transcribed clockwisely inside and anti-clockwisely outside.

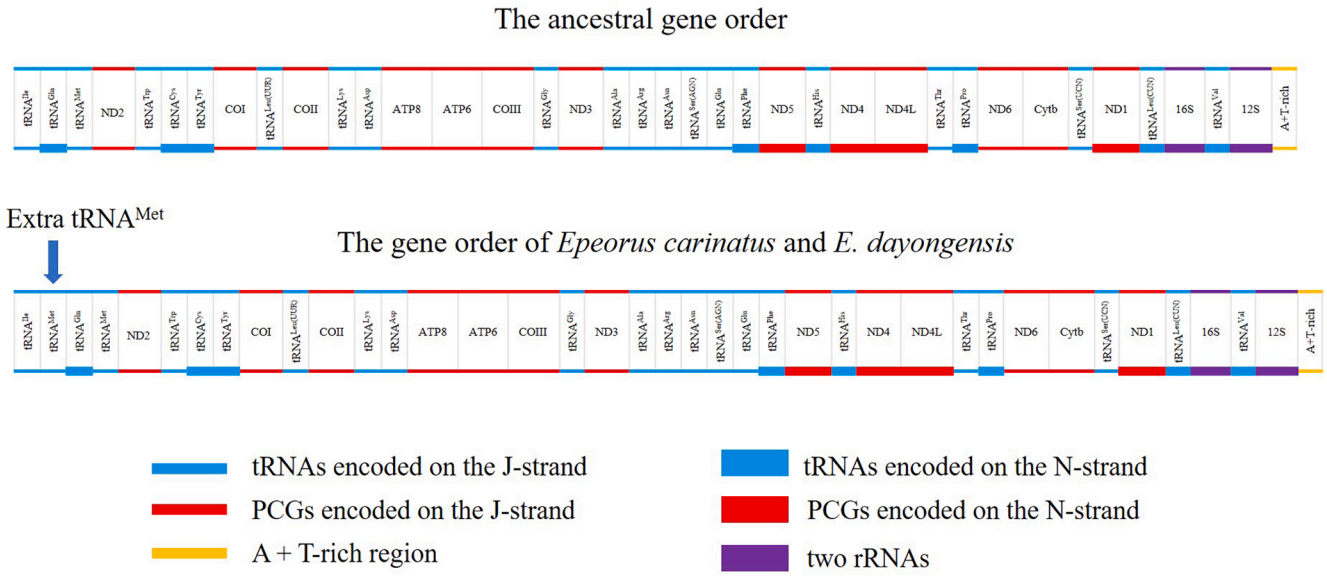

Fig. 2. Landscape of genetic rearrangement of *Epeorus carinatus* and *E. dayongensis.*

happened in mayflies (Wei et al., 2010; Li et al., 2014; Xu et al., 2020).

A comprehensive analysis of the genus *Epeorus* exhibited that the lowest A + T content was found in PCGs (64.07%–66.26%) and the highest A + T content in AT-rich region (72.25%–77.38%).

3.3. Protein-coding genes

The mitogenomes of *E. carinatus* and *E. dayongensis* had 9 PCGs encoded on the J-strand and 4 PCGs on the N-strand, as in most other insects. The initiation codons of all PCGs in the genus *Epeorus* were typical ATN (ATA, ATT, ATC and ATG) pattern except for APT8, COI and ND5. The ATP8 gene of both sequenced species started with GTG while that of *E. herklotsi* with ATG. The COI and ND5 genes of all three species used the start codon CCG and GTG, ND3 started with the codon ATA in *E. carinatus*, ATC in *E. dayongensis* and ATG in *E. herklotsi*. In terms of termination codons, COI, COII, ND4 and ND5 of three species were terminated with incomplete TA or T and the remaining genes with TAA or TAG (Table 3), which was consistent with other invertebrate insects. The possible interpretation was the incomplete TA or T could be transcribed to be the entire codon (UAA) via post-transcriptional polyadenylation (Li et al., 2014; Ojala et al., 1981).

The relative synonymous codon usage (RSCU) values of the mitogenomes of the three species was summarized in Fig. 3. The total number of codons in PCGs was 3726 in *E. carinatus*, 3724 in *E. dayongensis* and 3729 in *E. herklotsi* (Tables S3, S4 and S5). The result indicated that the codons UAA, UAG and AGG were not detected in these three species, and the three most frequently used codons were the same, i.e. UAA, UUU and AUU, corresponding to amino acid Leucine 2 (Leu2), Phenylalanine (Phe) and Isoleucine (Ile).

Table 1
Annotation of the mitogenomes of *E. carinatus* (Ecar) and *E. dayongensis* (Eday).

Feature	Strand	Nucleotide Number		Intergenic Nucleotides		Anticodon	Start/Stop Codons	
		Ecar	Eday	Ecar	Eday		Ecar	Eday
$tRNA^{Ile}$	J	1–65	1–65	0	0	GAT	–	–
$tRNA^{Met}$	J	66–131	67–132	0	1	CAT	–	–
$tRNA^{Gln}$	N	133–201	134–202	1	1	TTG	–	–
$tRNA^{Met}$	J	203–268	206–269	1	3	CAT	–	–
ND2	J	269–1303	270–1304	0	0	–	ATG/TAA	ATG/TAA
$tRNA^{Trp}$	J	1302–1369	1303–1370	–2	–2	TCA	–	–
$tRNA^{Cys}$	N	1362–1425	1363–1425	–8	–8	GCA	–	–
$tRNA^{Tyr}$	N	1427–1492	1427–1493	1	1	GTA	–	–
COI	J	1491–3024	1492–3025	–2	–2	–	CCG/T	CCG/T
$tRNA^{Leu(UUR)}$	J	3025–3090	3026–3091	0	0	TAA	–	–
COII	J	3095–3782	3096–3783	4	4	–	ATG/T	ATG/T
$tRNA^{Lys}$	J	3783–3851	3784–3852	0	0	CTT	–	–
$tRNA^{Asp}$	J	3852–3917	3853–3918	0	0	GTC	–	–
ATP8	J	3918–4076	3919–4077	0	0	–	GTG/TAA	GTG/TAA
ATP6	J	4073–4747	4074–4748	–4	–4	–	ATA/TAA	ATA/TAA
COIII	J	4747–5535	4748–5536	–1	–1	–	ATG/TAA	ATG/TAA
$tRNA^{Gly}$	J	5538–5602	5539–5602	2	2	TCC	–	–
ND3	J	5600–5956	5603–5956	–3	0	–	ATA/TAG	ATC/TAG
$tRNA^{Ala}$	J	5955–6018	5955–6018	–2	–2	TGC	–	–
$tRNA^{Arg}$	J	6052–6115	6060–6123	33	41	TCG	–	–
$tRNA^{Asn}$	J	6115–6178	6123–6187	–1	–1	GTT	–	–
$tRNA^{Ser(AGN)}$	J	6178–6243	6187–6252	–1	–1	GCT	–	–
$tRNA^{Glu}$	J	6247–6310	6255–6318	3	2	TTC	–	–
$tRNA^{Phe}$	N	6309–6372	6317–6380	–2	–2	GAA	–	–
ND5	N	6373–8107	6381–8115	0	0	–	GTG/T	GTG/T
$tRNA^{His}$	N	8108–8171	8116–8179	0	0	GTG	–	–
ND4	N	8171–9516	8179–9524	–1	–1	–	ATG/TA	ATG/TA
ND4L	N	9510–9806	9518–9814	–7	–7	–	ATG/TAA	ATG/TAA
$tRNA^{Thr}$	J	9809–9872	9817–9880	2	2	TGT	–	–
$tRNA^{Pro}$	N	9873–9937	9881–9945	0	0	TGG	–	–
ND6	J	9952–10,458	9960–10,466	14	14	–	ATT/TAA	ATT/TAA
Cytb	J	10,458–11,594	10,466–11,602	–1	–1	–	ATG/TAG	ATGTAG
$tRNA^{Ser(UCN)}$	J	11,593–11,661	11,601–11,669	–2	–2	TGA	–	–
ND1	N	11,678–12,628	11,686–12,636	16	16	–	ATG/TAA	ATG/TAA
$tRNA^{Leu(CUN)}$	N	12,630–12,695	12,638–12,702	1	1	TAG	–	–
16S	N	12,696–13,977	12,703–13,975	0	0	–	–	–
$tRNA^{Val}$	N	13,978–14,048	13,976–14,046	0	0	TAC	–	–
12S	N	14,049–14,848	14,047–14,844	0	0	–	–	–
A + T-rich	J	14,849–15,338	14,845–15,609	0	0	–	–	–

Notes: J refers to the majority strand and N refers to the minority strand. Position numbers refer to positions on the majority strand.

Table 2
Nucleotide composition in regions of mitogenomes of the genus *Epeorus*.

Species	Mitochondrial genome				PCGs		tRNAs		rRNAs		AT-rich region	
	Length (bp)	AT%	AT-skew	GC-skew	Length (bp)	AT%	Length (bp)	AT%	Length (bp)	AT%	Length (bp)	AT%
E. carinatus	15,338	64.65	−0.002	−0.208	11,210	64.07	1515	64.49	2082	66.23	490	72.25
E. dayongensis	15,609	67.2	−0.005	−0.244	11,175	66.26	1512	67.4	2071	67.8	765	77.38
E. herklotsi	15,502	65.67	−0.002	−0.246	11,180	65	1512	65.75	2086	65.78	634	74.45

Table 3
Start and stop codons of protein-coding genes in mitogenomes of *E. carinatus* (Ecar), *E. dayongensis* (Eday) and *E. herklotsi* (Eher).

Species	Gene												
	ATP6	ATP8	COI	COII	COIII	Cytb	ND1	ND2	ND3	ND4	ND4L	ND5	ND6
Ecar	ATA/TAA	GTG/TAA	CCG/T	ATG/T	ATG/TAA	ATG/TAG	ATG/TAA	ATG/TAA	ATA/TAG	ATG/TA	ATG/TAA	GTG/T	ATT/TAA
Eday	ATA/TAA	GTG/TAA	CCG/T	ATG/T	ATG/TAA	ATG/TAG	ATG/TAA	ATG/TAA	ATC/TAG	ATG/TA	ATG/TAA	GTG/T	ATT/TAA
Eher	ATA/TAA	ATG/TAA	CCG/T	ATG/T	ATG/TAA	ATG/TAG	ATG/TAA	ATG/TAA	ATG/TAG	ATG/TA	ATG/TAA	GTG/T	ATT/TAA

Note: The different codons are shown in red.

3.4. tRNAs, rRNAs and AT-rich regions

Gene structures and their arrangements are useful information to molecular evolution and phylogenetic reconstruction (Boore, 1999; Boore and Brown, 1998; Serb and Lydeard, 2003; Boore et al., 2004). For vertebrates, the tandem duplication random loss (TDRL) of gene duplicates was considered as the main mechanism, whereas in invertebrates, the mechanism still remained unclear due to lacking of available mitogenomes from important lineages (Eberhard and Wright, 2016; San Mauro et al., 2006; Grande et al., 2008; Guerra et al., 2018; Uribe et al., 2016; Xie et al., 2019). In this study, we found the mitogenomes of *Epeorus* species contained one extra $tRNA^{Met}$ gene (Fig. 2)

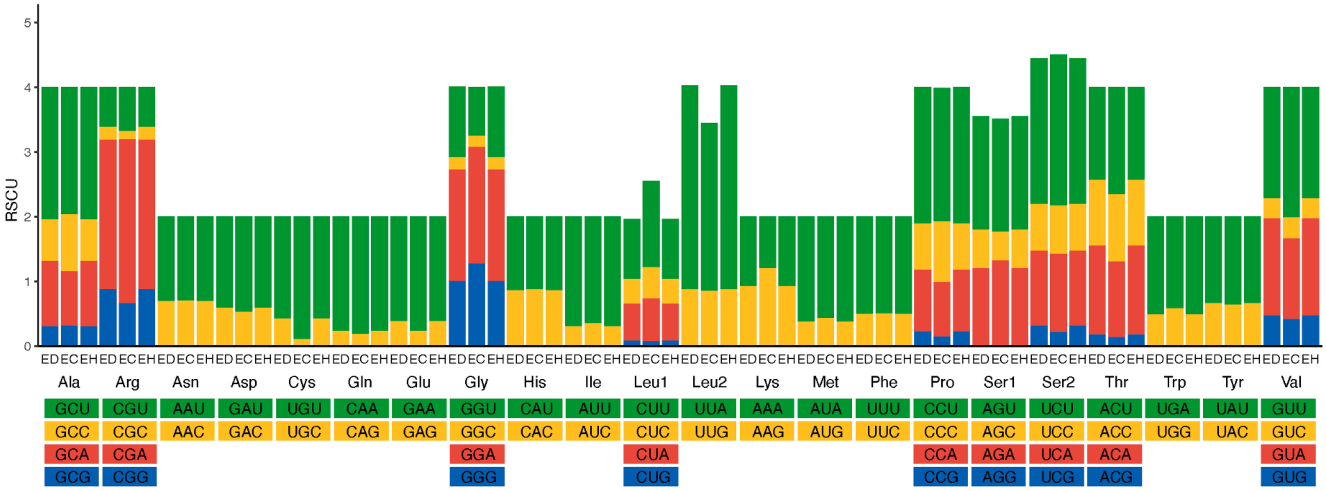

Fig. 3. The relative synonymous codon usage (RSCU) in the mitogenomes of *Epeorus dayongensis* (ED), *E. carinatus* (EC) and *E. herklotsi* (EH).

and the duplicate tRNA^{Met} of *E. dayongensis* varied considerably from the original tRNA^{Met} and between different species, especially in amino acid accepter (AA) arm and anticodon (AC) arm (Fig. S1). Moreover, the duplicate tRNA^{Met} gene was found in all known mitogenomes of heptageniid species except *Paegniodes cupulatus* (Zhang et al., 2008; Zhou et al., 2016; Tang et al., 2014; Gao et al., 2018; Song et al., 2019; Wu and Yu, 2018). It may indicate the duplicate tRNA^{Met} gene could be used as the molecular synapomorphy for a portion of the family. In this research, all tRNAs of the three *Epeorus* species had standard anticodons and could fold into cloverleaf secondary structures except tRNA$^{Ser(AGN)}$ (Figs. 4 and S1). The tRNA$^{Ser(AGN)}$ could not form the complete cloverleaf secondary structure due to lack of dihydrouridine (DHU) loop. In addition to normal base pairs, the stems of the secondary structures also contained non-Waston-Crick base pairs. In *E. carinatus*, there were 40 non-canonical G-U (or U-G) pairs and mismatched pairs U-U for twice and C-U for once (Fig. S1). In *E. dayongensis*, 40 non-canonical G-U (or U-G) base pairs and C-U, U-U, A-G and C-A mismatches (each for once) were recognized (Fig. S1). Additionally, this phenomenon also performed in *E. herklotsi*, there were 40 non-canonical G-U (or U-G) pairs and mismatched base pairs U-U for twice and A-C for once (Fig. S1). These mismatches could be corrected through editing process, and should not affect the transport function (Wang et al., 2018).

The large ribosomal RNA subunit (16S rRNA) was located between tRNA$^{Leu(CUN)}$ and tRNA^{Val}, and the small ribosomal RNA subunit (12S rRNA) was located between tRNA^{Val} and AT-rich region. The length of 16S rRNA was 1282 bp in *E. carinatus*, 1273 bp in *E. dayongensis* and 1283 bp in *E. herklotsi*. The length of 12S rRNA was 800 bp in *E. carinatus*, 798 bp in *E. dayongensis* and 803 bp in *E. herklotsi*. The A + T content accounting for 66.23% in *E. carinatus*, 67.8% in *E. dayongensis* and 65.78% in *E. herklotsi*, slightly higher than PCGs and tRNAs.

The AT-rich region was the only major non-coding segment in the mitogenome of insects, and contributed to the size variation of mitogenomes. For example, *Siphluriscus chinensis* had the longest mitogenome (16,616 bp) in mayflies. At the same time, its mitogenome contained the longest AT-rich region (1,829 bp). In contrast, the mitogenome of *Baetis rutilocylindratus* (14,883 bp) was relatively shorter while it had the shortest AT-rich region (340 bp) (Li et al., 2014; Xu et al., 2020). In this research, the AT-rich region of the three *Epeorus* species was between 12S rRNA and tRNA^{Ile}, with the length 490 bp in *E. carinatus*, 765 bp in *E. dayongensis* and 634 bp in *E. herklotsi*. These data showed the size of AT-rich region could change dramatically in different species, even they were in same family or genus.

3.5. Phylogenetic analysis

In the present study, 13 protein-coding genes from 22 mayfly species were used as dataset to construct phylogenetic trees by BI and ML methods, the species *Pedetontus silvestrii* (Archaeognatha, Machilidae) and *Nesomanchilis australica* (Archaeognatha, Meinertellidae) were selected as outgroups. The GenBank accession numbers of those 24 species used in this study were shown in Table S1 (Xu et al., 2020a, 2020b; Zhang et al., 2008a, 2008b; Zhou et al., 2016; Tang et al., 2014; Gao et al., 2018; Wang et al., 2018; Cameron et al., 2004; Ye et al., 2018). The Phylogenetic topologies constructed by two methods were similar except the positions of Baetidae, Leptophlebiidae and Caenidae (Fig. 5, shown in red block).

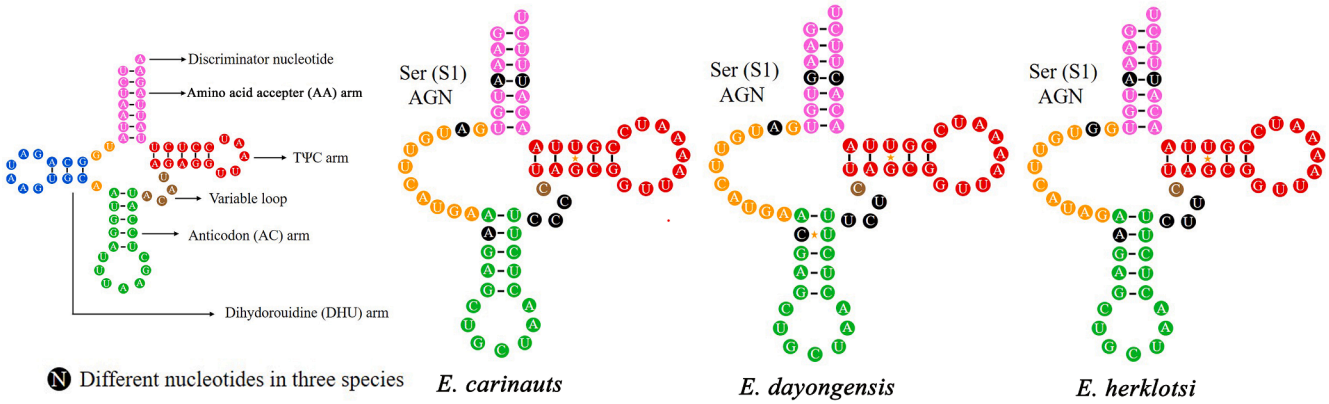

Fig. 4. Putative tRNA second structure and the secondary structure of tRNA$^{Ser(AGN)}$ for the three *Epeorus* species.

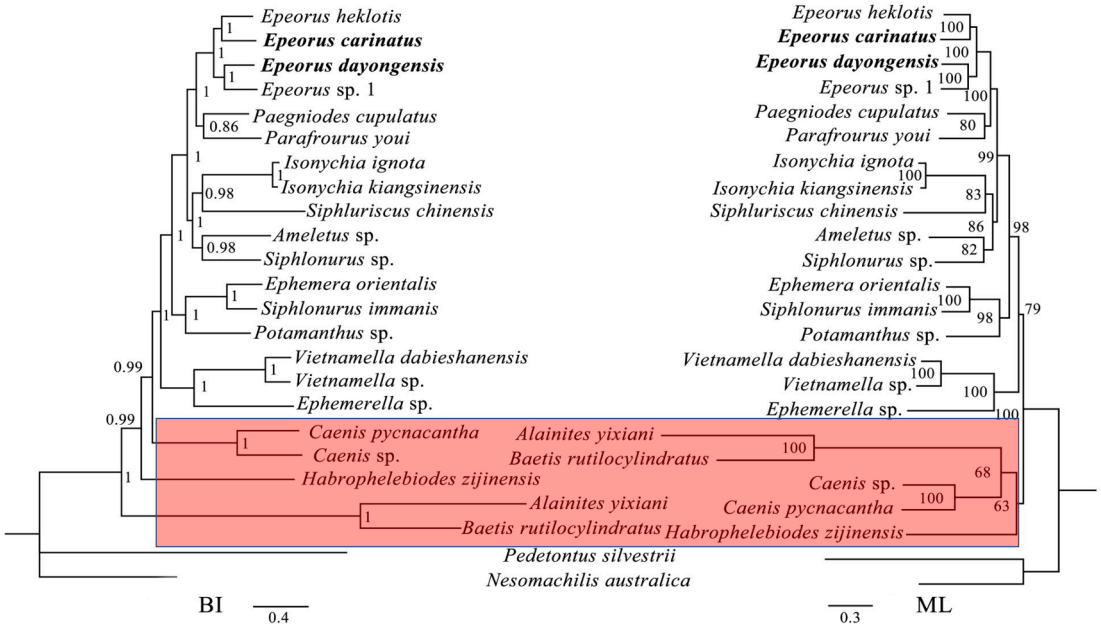

Fig. 5. Phylogenetic trees constructed by Bayesian and Maximum-likelihood analyses based on the 13 protein-coding genes (the differences are shown in red block). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The BI tree indicated that ((*Baetis rutilocylindratus* + *Alainites yixiani*) + (*Habrophelebiodes zijinensis* + ((*Caenis* sp. + *Caenis pycnacantha*))) with high supporting values. By contrast, ML tree provided their relationships as (*Habrophelebiodes zijinensis* + ((*Caenis* sp. + *Caenis pycnacantha*) + (*Baetis rutilocylindratus* + *Alainites yixiani*))). These competing results were found in related previous studies (Zhang et al., 2008; Gao et al., 2018; Xu et al., 2020a, 2020b; Song et al., 2019; Cai et al., 2018; Ye et al., 2018), the main reason was believed due to the limited number of mitogenomes available. Despite that, our trees showed some valuable information. First, the families Ephemerellidae and Vietnamellidae were clustered together, which was consistent with the recent studies (Gao et al., 2018; Xu et al., 2020a, 2020b; Wu and Yu, 2018; Cai et al., 2018; Ye et al., 2018; Miller et al., 2018). Second, the monophyly of burrowing mayflies (presenting by the families Potamanthidae and Ephemeridae in this study) was supported, and this finding had been inferred stably by other mayfly taxonomists (Miller et al., 2018; Ogden et al., 2019). Third, our results also supported the following relationships: ((Siphlonuridae + Ameletidae) + (Siphluriscidae + Isonychiidae)), in which the two siphlonurid species demonstrated the family Siphlonuridae was a polyphyletic group.

The family Heptageniidae was strongly corroborated as monophyletic in our topologies although the relationships within *Epeorus* remained unstable. In our trees, all four *Epeorus* species (*E. carinatus*, *E. herklotsi*, *E.* sp. 1 and *E. dayongensis*), which supported this genus was a monophyletic group. Furthermore, our topologies also showed the focal species *E. carinatus* and *E. dayongensis* were in different subgenera, which was consistent with the morphological data (Kluge, 2004; Braasch, 2006).

In our trees, the species *Paegniodes cupulatus* (Rhithrogeninae) was the sister group of *Parafronurus youi* (Ecdyonurinae) instead of other Rhithrogeninae members (*Epeorus* species), which was totally contrast to previous morphological results (Webb and McCafferty, 2008; Zhou and Braasch, 2003; Wang and McCafferty, 2004; Ma et al., 2018). This may due to its plesiomorphic characters (Webb and McCafferty, 2008; Wang and McCafferty, 2004; Ma et al., 2018) and the duplicate $tRNA^{Met}$ genes.

Due to limited molecular data and available ingroups, the phylogeny discussed in this research is very preliminary. With the growing number of available mitogenomes of Ephemeroptera, the results presented here will be tested by further studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Xiaohan Shu for helping with graphics. This work was supported by the National Natural Science Foundation of China (Grant 31750002), funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and supported by key projects of science-technology basic condition platform from The Ministry of Science and Technology of the People's Republic of China (Grant No. 2005DKA21402).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2021.145467.

References

Wolstenholme, D.R., 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. 141, 173–216.
Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.
Cameron, S.L., 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. 59, 95–117.
Barber-James, H.M., Gattolliat, J.L., Sartori, M., Hubbard, M.D., 2008. Global diversity of mayflies (Ephemeroptera, Insecta) in freshwater. Hydrobiologia 595, 339–350.
Webb, J.M., McCafferty, W.P., 2008. Heptageniidae of the World. Part II: key to the genera. Can. J. Arthropod Ident. 7, 1–55.
Sartori, M., 2014a. About the status of the enigmatic Oriental genus Rhithrogeniella Ulmer, s1939 (Ephemeroptera, Heptageniidae), ZooKeys 429, 47–61.
Sartori, M., 2014b. The concept of Compsoneuria Eaton, 1881 revisited in light of historical and new material from the Sudan Islands (Ephemeroptera, Heptageniidae, Ecdyonurinae), Zootaxa 3835, 1–32.
Sartori, M., 2014c. The species of Thalerosphyrus Eaton, 1881 (Insecta, Ephemeroptera, Heptageniidae, Ecdyonurinae) in Java and Sumatra, with some comments on the diversity of the genus in the Oriental Realm. ZooKeys 420, 19–39.

Sartori, M., 2014d. What is *Ecdyonurus sumatranus* Ulmer, 1939? A contribution to the knowledge of the genus *Rhithrogena* in the Oriental Region (Ephemeroptera, Heptageniidae). Zootaxa 3802, 193–208.
Yanai, Z., Sartori, M., Dor, R., Dorchin, N., 2017. Molecular phylogeny and morphological analysis resolve a long-standing controversy over generic concepts in Ecdyonurinae mayflies (Ephemeroptera: Heptageniidae). Syst. Entomol. 42, 182–193.
Ogden, T.H., Whiting, M.F., Wheeler, W.C., 2005. Poor taxon sampling, poor character sampling, and non-repeatable analyses of a contrived dataset do not provide a more credible estimate of insect phylogeny: a reply to Kjer. Cladistics 21, 295–302.
Ogden, T.H., Gattolliat, J.L., Sartori, M., Staniczek, A.H., Soldán, T., Whiting, M.F., 2009. Towards a new paradigm in mayfly phylogeny (Ephemeroptera): combined analysis of morphological and molecular data. Syst. Entomol. 34, 616–634.
Vuataz, L., Sartori, M., Gattolliat, J.L., Monaghan, M.T., 2013. Endemism and diversification in freshwater insects of Madagascar revealed by coalescent and phylogenetic analysis of museum and field collections. Mol. Phylogenet. Evol. 66, 979–991.
Zhou, C.F., Braasch, D., 2003. Eine neue Gattung und Art der Heptageniidae aus dem ostlichen China (Ephemeroptera). Entomologische Nachrichten und Berichte 47, 147–151.
Eaton, A.E., 1871. A monograph on the Ephemeridae. Trans. Entomol. Soc. London 1, 1–164.
Hsu, Y.C., 1936. Mayflies of Hong Kong with description of two new species (Ephemeroptera). Hong Kong Naturalist 7, 233–238.
Zhang, J.Y., Zhou, C.F., Gai, Y.H., Song, D.X., Zhou, K.Y., 2008. The complete mitochondrial genome of *Parafronurus youi* (Insecta: Ephemeroptera) and phylogenetic position of the Ephemeroptera. Gene 424, 18–24.
Zhou, D., Wang, Y.Y., Sun, J.Z., Han, Y.K., Zhou, C.F., 2016. The complete mitochondrial genome of *Paegniodes cupulatus* (Ephemeroptera: Heptageniidae), Mitochondrial DNA A DNA. Mapp. Seq. Anal. 27, 925–926.
Tang, M., Tan, M., Meng, G., Yang, S., Su, X., 2014. Multiplex sequencing of pooled mitochondrial genomes a crucial step toward biodiversity analysis using mito-metagenomics. Nucleic Acids Res. 42, e166.
Gao, X.Y., Zhang, S.S., Zhang, L.P., Yu, D.N., Zhang, J.Y., Cheng, H.Y., 2018. The complete mitochondrial genome of *Epeorus herklotsi* (Ephemeroptera: Heptageniidae) and its phylogeny. Mitochondrial DNA B 3, 303–304.
Satori, M., Brittain, J.E., 2015. Order Ephemeroptera. In: Thorp, J., Rogers, D.C. (Eds.), Ecology and General Biology: Thorp and Covich's Freshwater Invertebrates. Academic Press, pp. 873–891.
Wang, T.Q., McCafferty, W.P., 2004. Heptageniidae (Ephemeroptera) of the world, Part I: Phylogenetic higher classification. Trans. Amer. Entomol. Soc. 130, 11–45.
Kluge, N.J., 2004. The Phylogenetic System of Ephemeroptera. Kluwer Academic Publishers, Dordrecht, p. 442.
Braasch, D., 2006. Neue Eintagsfliegen der Gattungen *Epeorus* und *Iron* aus dem Himalaja (Ephemeroptera, Heptageniidae). Entomologische Nachrichten und Berichte 50, 79–88.
Chen, P., Wang, Y.Y., Zhou, C.F., 2010. A new mayfly species of *Epeorus* (*Caucasiron*) from southwestern China (Ephemeroptera: Heptageniidae. Zootaxa 2527, 61–68.
Kluge, N.J., Tiunova, T.M., 1989. Palearctic mayflies of the group longimanus of the subgenus *Iron* of *Epeorus* (Ephemeroptera, Heptageniidae). Vestnik zoologii 4 (1989), 7–14.
Boonsoong, B., Braasch, D., 2013. Heptageniidae (Insecta, Ephemeroptera) of Thailand. ZooKeys 272, 61–93.
Hrivniak, L. Sroka, P., Godunko, R.J., Žurovcovà, M., 2017. Mayflies of the genus Epeorus Eaton, 1881 s.l. (Ephemeroptera: Heptageniidae) from the Caucasus Mountains: a new species of Caucasiron Kluge, 1997 from Georgia and Turkey. Zootaxa 4341, 353–374.
Hrivniak, L., Sroka, P., Türkmen, C., Godunko, R.J., Kazanci, N., 2019. A new *Epeorus (Caucasiron*) (Ephemeroptera: Heptageniidae) species from Turkey based on molecular and morphological evidence. Zootaxa 4550, 58–70.
Braasch, D., Soldán, T., 1984. Eintagsfliegen (Gattungen Epeorus und Iron) aus Vietnam (Ephemeroptera, Heptageniidae), 1984, pp. 109–114, in: Landa, V., Soldán, T., Tonner, M. (Eds.), Proceedings of the Fourth International Conference on Ephemeroptera, Institute of Entomology, Czechoslovak Academy of Sciences, Ceské Budejovice.
Gui, H., Zhang, J., 1992. A new species of genus *Epeorus* Eaton (Ephemeroptera: Heptageniidae). Acta Zootaxonomica Sinica 17, 61–64.
Patel, R.K., Jain, M., 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLOS ONE 7, e30619.
Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863–864.
Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. 18, 821–829.
Kumar, S., Stecher, G., Tamura, K., MEGA,, 2016. 7: molecular evolutionary genetics analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874.
Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.
Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: improved *de novo* metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69, 313–319.
Perna, N.T., Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41, 353–358.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Xia, X., Xie, Z., 2001. DAMBE: software package for data analysis in molecular biology and evolution. J. Hered. 92, 371–373.
Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post–analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE), IEEE, 2010, pp. 1–8.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., et al., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Lavrov, D.V., Brown, W.M., 2001. Trichinella spiralis mtDNA: a nematode mitochondrial genome that encodes a putative ATP8 and normally structured tRNAs and has a gene arrangement relatable to those of coelomate Metazoans. Genetics 157, 621–637.
Carapelli, A., Liò, P., Nardi, F., E.van der Wath, F. Frati,, 2007. Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. BMC Evol. Biol. 7, S8.
Wei, S.J., Shi, M., Chen, X.X., Sharkey, M.J., van Achterberg, C., Ye, G.Y., He, J.H., 2010. New views on strand asymmetry in insect mitochondrial genomes. PloS One 5, e12708.
Li, D., Qin, J.C., Zhou, C.F., 2014. The phylogeny of Ephemeroptera in Pterygota revealed by the mitochondrial genome of *Siphluriscus chinensis* (Hexapoda: Insecta). Gene 545, 132–140.
Xu, X.D., Jia, Y.Y., Cao, S.S., Zhang, Z.Y., Storey, K.B., Yu, D.N., Zhang, J.Y., 2020. Six complete mitochondrial genomes of mayflies from three genera of Ephemerellidae (Insecta: Ephemeroptera) with inversion and translocation of trnI rearrangement and their phylogenetic relationships. Peer J. 8, e9740.
Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature 290, 470–474.
Boore, J.L., Brown, W.M., 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Curr. Opin. Genet. Dev. 8, 668–674.
Serb, J.M., Lydeard, C., 2003. Complete mtDNA sequence of the North American freshwater mussel, *Lampsilis ornata* (Unionidae): an examination of the evolution and phylogenetic utility of mitochondrial genome Organization in Bivalvia (Mollusca). Mol. Biol. Evol. 20, 1854–1866.
Boore, J.L., Medina, M., Rosenberg, L.A., 2004. Complete sequences of the highly rearranged molluscan mitochondrial genomes of the Scaphopod *Graptacme eborea* and the Bivalve *Mytilus edulis*. Mol. Biol. Evol. 21, 1492–1503.
Eberhard, J.R., Wright, T.F., 2016. Rearrangement and evolution of mitochondrial genomes in parrots. Mol. Phylogenet. Evol. 94, 34–46.
San Mauro, D., Gower, D.J., Zardoya, R., Wilkinson, M., 2006. A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. Mol. Biol. Evol. 23, 227–234.
Grande, C., Templado, J., Zardoya, R., 2008. Evolution of gastropod mitohondrial gene arrangements. BMC Evol. Biol. 8, 61.
Guerra, D., Bouvet, K., Breton, S., 2018. Mitochondrial gene order evolution in Mollusca: Inference of the ancestral state from the mtDNA of *Chaetopleura apiculate* (Polyplacophora, Chaetopleuridae). Mol. Phylogenet. Evol. 120 (2018), 233–239.
Uribe, J.E., Kano, Y., Templado, J., Zardoya, R., 2016. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zool. Scr. 45, 145–159.
Xie, G.L., Köhler, F., Huang, X.C., Wu, R.W., Zhou, C.H., Ouyang, S., Wu, X.P., 2019. A novel gene arrangement among the Stylommatophora by the complete mitochondrial genome of the terrestrial slug *Meghimatium bilineatum* (Gastropoda, Arionoidea). Mol. Phylogenet. Evol. 135, 177–184.
Song, N., Li, X., Yin, X., Li, X., Yin, J., Pan, P., 2019. The mitochondrial genomes of palaeopteran insects and insights into the early insect relationships. Sci. Rep. 9, 17765.
Wu, M.J., Yu, L.L., 2018. The complete mitochondrial genome of *Epeorus herklotsi* (Ephemeroptera: Heptageniidae) from Longquan, Zhejiang, China and its phylogeny. Mitochondrial DNA B 3, 1254–1255.
Wang, Z.F., Wang, Z.Q., Shi, X.J., Wu, Q., Tao, Y.T., Guo, H.Y., et al., 2018. Complete mitochondrial genome of *Parasesarma affine* (Brachyura: Sesarmidae): gene rearrangements in Sesarmidae and phylogenetic analysis of the Brachyura. Int. J. Biol. 118, 31–40.
Cameron, S.L., Miller, K.B., D'Haese, C.A., Whiting, M.F., Barker, S.C., 2004. Mitochondrial genome data alone are not enough to unambiguously resolve the relationships of Entognatha, Insecta and Crustacea *sensu lato* (Arthropoda). Cladistics 20, 534–557.
Zhang, J.Y., Song, D.X., Zhou, K.Y., 2008. The complete mitochondrial genome of the bristletail *pedetontus silvestrii* (Archaeognatha: Machilidae) and an examination of mitochondrial gene variability within four bristletails. Ann. Entomol. Soc. Am. 101, 1131–1136.
Cai, Y.Y., Gao, Y.J., Zhang, L.P., Yu, D.N., Storey, K.B., Zhang, J.Y., 2018. The mitochondrial genome of *Caenis* sp. (Ephemeroptera: Caenidae) and the phylogeny of Ephemeroptera in Pterygota. Mitochondrial DNA B 3, 577–579.
Ye, Q.M., Zhang, S.S., Cai, Y.Y., Storey, K.B., Yu, D.N., Zhang, J.Y., 2018. The complete mitochondrial genome of *Isonychia kiangsinensis* (Ephemeroptera: Isonychiidae). Mitochondrial DNA B 3, 541–542.
Xu, X.D., Jia, Y.Y., Dai, X.Y., Ma, J.L., Storey, K.B., Zhang, J.Y., Yu, D.N., 2020. The mitochondrial genome of *Caenis* sp. (Ephemeroptera: Caenidae) from Fujian and the phylogeny of Caenidae within Ephemeroptera. Mitochondrial DNA B 5, 152–153.

Miller, D.B., Bartlett, S., Sartori, M., Breinholt, J.W., Ogden, T.H., 2018. Anchored phylogenomics of burrowing mayflies (Ephemeroptera) and the evolution of tusks. Syst. Entomol. 43, 692–701.
Ogden, T.H., Breinholt, J.W., Bybee, S.M., Miller, D.B., Sartori, M., Shiozawa, D., Whiting, M.F., 2019. Mayfly phylogenomics: initial evaluation of anchored hybrid enrichment data for the order Ephemeroptera. Zoosymposia 16, 167–181.
Ma, Z.X., Han, N., Zhang, W., Zhou, C.F., 2018. Position and definition of the genus Paegniodes (Eaton, 1881) based on redescription on the type species *Paegniodes cupulatus* (Eaton, 1871) (Ephemeroptera: Heptageniidae). Aquat. Insect. 39, 362–374.