Biologia
https://doi.org/10.1007/s11756-021-00729-6

ORIGINAL ARTICLE

First complete mitogenomes of three mayflies in the genus *Afronurus* (Ephemeroptera: Heptageniidae) and their implications for phylogenetic reconstruction

Ran Li[1] · Wei Zhang[1] · Zhenxing Ma[1] · Changfa Zhou[1]

Received: 16 December 2020 / Accepted: 1 March 2021

Abstract
Heptageniidae is one of the most species-rich families within Ephemeroptera, whose evolution pattern of mitogenomes and internal relationships remain unclear. In this study, the complete mitochondrial genomes (mitogenomes) of *Afronurus rubromaculata*, *A. yixingensis* and *A. obliquistrita* were firstly sequenced and analyzed. Three mitogenomes share similar gene arrangement with an extra *trnM* gene compared with the ancestral order for insects. Comprehensive analysis of tRNA secondary structures showed that the extra *trnM* had a high level of nucleotide variation. The conservative intergenic spacer (IGS) region in the mitogenomes was found between the *trnA* and *trnR* genes with a size of 39 nucleotides and all formed a loop-stem structure. Furthermore, tandem repeats were identified in the control region of all three mitogenomes. The phylogenetic analyses using both Bayesian inference (BI) and maximum likelihood (ML) methods based on three datasets all supported that Heptageniidae was a monophyletic family. Our results supported the relationships ((*A. rubromaculata* + *A. obliquistrita*) + *A. yixingensis*) within *Afronurus*. The analysis will improve our understanding of the mitogenome features of heptageniid mayflies and lay the foundation for further phylogenetic study of Ephemeroptera.

Keywords Ephemeroptera · Heptageniidae · *Afronurus* · Mitochondrial genome · Phylogenetic analysis

Introduction

An insect mitochondrial genome (mitogenome) is generally a double-strand circular molecule, 14–19 kilobases (kb) in length (Boore 1999). It generally contains 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), two ribosomal RNA genes (rRNAs), and a large non-coding region (called control region or A + T-rich region) associated with the initiation of transcription and replication (Wolstenholme 1992). Because of the unique features of maternal inheritance, absence of introns, low rate of recombination and high evolutionary rate, the sequences of insect mitogenome have been extensively used in studies of molecular phylogenetics and evolution, population genetics, and phylogeography (Saccone et al. 1999; Lin and Danforth 2004; Simon et al. 2006; Avise 2009; Krzywinski et al. 2011; Cameron 2014; Qin et al. 2015). Moreover, complete mitogenome has been shown to provide more credible results than single or several mitochondrial genes in many studies (Yuan et al. 2016; Liu et al. 2019).

Heptageniidae is one of the most species-rich families within the order Ephemeroptera (mayfly), comprising 36 genera and more than 600 described species (Sartori 1992; Barber-James et al. 2008; Webb and McCafferty 2008). Members of the family are widely distributed in the Palaearctic Region (Wang and McCafferty 2004; Yanai et al. 2017). The heptageniid nymphs usually inhabit lotic habitats, and mainly feed on fine detritus and diatoms. As a consequence, they are important components of the mayfly community and most ecosystems. Over the past decade, studies of Heptageniidae have mainly focused on taxonomy (Ma et al. 2018; Sutthacharoenthad et al. 2019), behaviour (McCafferty and Huff 2017), and ecology (Freimark et al. 2017). Owing to limited molecular informative traits, phylogenetic relationships among Heptageniidae were poorly studied. Meanwhile, most phylogenetic analyses were still based

✉ Changfa Zhou
mayflyzcfnnu@163.com

[1] The Key Laboratory of Jiangsu Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Wenyuan Road No. 1, Nanjing 210023, People's Republic of China

Published online: 15 March 2021

on a single gene or several genes so far (Vuataz et al. 2011; Yanai et al. 2017; Kaneko et al. 2018). Up to now, only five complete mitogenomes representing three genera (three of *Epeorus*, one of *Paegniodes*, one of *Parafronurus*) have been submitted to the GenBank (https://www.ncbi.nlm.nih.gov/; last visited on August 2020) (Gao et al. 2018). Therefore, mitogenomic studies of more mayflies are required for a comprehensive phylogenetic analysis of the family Heptageniidae, and still in demand before a full picture of mitogenomes for Ephemeroptera is depicted.

The genus *Afronurus* Kimmins, 1937 is widely distributed in Asia, and generally considered as a junior synonym of the more widespread Asian and African *Afronurus*. However, there is no mitogenome of *Afronurus* was reported, the gene organization have remained unclear, and its taxonomic status among Heptageniidae is poorly understood. Herein we sequenced and annotated the complete mitogenomes of three widely distributed species in China (*Afronurus rubromaculata* You, Wu, Gui & Hsu, 1981; *A. yixingensis* Wu & You, 1986 and *A. obliquistrita* You, Tian, Hong & Hsu, 1981). Furthermore, we analyzed the main features of the newly generated mitogenomes and that of the related species (*Parafronurus youi* Zhou & Braasch, 2003), including genomic structure, gene order, nucleotide composition, secondary structure of tRNAs and control region. In addition, we reconstructed the phylogenetic relationships within the order of Ephemeroptera using current mitogenome information, and analyzed the taxonomic status of *Afronurus* and its phylogenetic relationships with other heptageniid species at the mitogenome level. Our results will improve our understanding of the mitogenome features of heptageniid mayflies and lay the foundation for further phylogenetic study of Ephemeroptera.

Materials and methods

Sample collection and DNA extraction

All specimens for three *Afronurus* species were collected in 2019, *A. rubromaculata* from the city of Yixing in the province of Jiangsu, China, *A. yixingensis* and *A. obliquistrita* from Yichang in Hubei, China. All specimens were immediately preserved in 100% ethanol and then stored at −20 °C in the College of Life Sciences of Nanjing Normal University, where we conducted this research without permit requirement. The specimens were morphologically identified by Chang-Fa Zhou using available taxonomic keys. Whole genomic DNA for subsequent analyses was extracted separately from the individual specimen using a DNeasy® tissue kit (Qiagen, Hilden, Germany). NanoDrop 2000 spectrophotometer (Thermo, Wilmington, USA) was then used to test the DNA concentration.

Mitogenome sequencing and assembly

The qualified samples were sequenced on the Illumina HiSeq 2500 (Illumina, California, USA) platform following a PE150 strategy (2 × 150 base paired-end reads). Paired-end libraries were constructed using the TruSeq DNA Library Preparation kit (https://support.illumina.com/downloads/truseq) with standard protocols. Three libraries were pooled, and then sequenced together with other projects. Each mitogenome yielded approximately 20 million raw reads. The raw reads were quality-trimmed using NGS-Toolkit by removing adapter contamination and low-quality reads (Patel and Jain 2012). The clean sequence reads were assembled in Geneious 11.1.5 using the mitochondrial genome of *P. youi* (GenBank accession number: EU349015) as mapping reference, with the following parameter settings: Minimum Overlap = 30–50, Minimum Overlap Identity: 80–100, Maximum Mismatches Per Read = 10% (Kearse et al. 2012).

Mitogenome annotation and comparative analysis

Initial annotation of the mitogenomes was performed using the MITOS web server under the code for invertebrate mitochondria (Bernt et al. 2013). The start and stop codons of all PCGs were further adjusted and corrected manually using reference mitogenome of *P. youi*. rRNA genes were identified by aligning homologous genes of previously sequenced mitogenomes from the family Heptageniidae. The non-coding regions (intergenic spacers and overlapping regions) between genes were estimated manually. Tandem Repeat Finder 4.07 (http://tandem.bu.edu/trf/trf.html) was used to find tandem repetitive sequences. The circular map of the mitogenomes was drawn using CGView (Grant and Stothard 2008).

Base composition, codon distribution and relative synonymous codon usage (RSCU) of PCGs were analysed with MEGA X (Kumar et al. 2018). Composition skew: AT-skew = (A − T) / (A + T) and GC-skew = (G − C) / (G + C), was used to measure nucleotide compositional differences (Perna and Kocher 1995).

Phylogenetic analysis

A total of 27 mitogenomes were used for phylogenetic analyses (Online resource 1: Table S1), including the three newly determined sequences, 22 available sequences of Ephemeroptera, and two of Archaeognatha (*Nesomachilis australica* Tillyard, 1924 and *Pedetontus silvestrii* Mendes, 1993 were used as outgroups). Sequences for each PCG were aligned individually with codon-based multiple alignments using MAFFT 7.0 (Katoh et al. 2005) within TranslatorX online platform (with the L-INS-i strategy and default setting) (Abascal et al. 2010). Sequences from each of the two rRNAs were aligned separately with the MAFFT algorithm using the G-INS-i strategy. Additionally, Gblocks

Table 1 Mitochondrial genome organization of *Afronurus rubromaculata* (Aru), *Afronurus yixingensis* (Ayi) and *Afronurus obliquistrita* (Aob)

Gene	Strand	Size (bp)			Intergenic Nucleotides			Anticodon
		Aru	Ayi	Aob	Aru	Ayi	Aob	
trnI	J	65	66	66	0	0	0	GAT
trnM	J	66	66	66	0	0	0	CAT
trnQ	N	69	69	69	0	0	0	TTG
trnM	J	66	65	64	-1	-1	-1	CAT
nad2	J	1035	1035	1035	0	0	0	–
trnW	J	68	68	68	−2	−2	−2	TCA
trnC	N	62	62	62	−8	−8	−8	GCA
trnY	N	64	66	65	0	0	0	GTA
cox1	J	1534	1534	1534	−2	−2	−2	–
trnL2	J	66	67	66	0	0	0	TAA
cox2	J	688	688	688	4	4	4	–
trnK	J	69	68	68	0	0	0	CTT
trnD	J	67	67	67	5	0	0	GTC
atp8	J	159	159	159	0	0	0	–
atp6	J	675	675	675	−4	−4	−4	–
cox3	J	789	789	789	−1	−1	−1	–
trnG	J	64	64	64	−1	−1	−1	TCC
nad3	J	357	357	357	−3	−3	−3	–
trnA	J	64	64	64	−2	−2	−2	TGC
trnR	J	64	64	64	39	39	39	TCG
trnN	J	64	65	65	0	7	0	GTT
trnS1	J	66	66	66	−3	−3	−3	GCT
trnE	J	64	64	64	2	2	3	TTC
trnF	N	65	65	65	−2	−2	−2	GAA
nad5	N	1735	1735	1735	0	0	0	–
trnH	N	64	64	64	0	0	0	GTG
nad4	N	1346	1346	1346	−1	−1	−1	–
nad4L	N	297	297	297	−7	−7	−7	–
trnT	J	65	65	65	2	2	2	TGT
trnP	N	66	65	65	0	0	0	TGG
nad6	J	507	507	507	14	14	14	–
cytb	J	1137	1137	1137	−1	−1	−1	–
trnS2	J	70	70	70	−2	−2	−2	TGA
nad1	N	951	951	951	16	16	16	–
trnL1	N	66	66	66	1	1	1	TAG
rrnL	N	1283	1285	1286	0	0	0	–
trnV	N	71	71	71	0	0	0	TAC
rrnS	N	786	780	779	0	0	0	–
A+T-rich	J	574	846	704	0	0	0	–

0.91b was employed with default settings to eliminate poorly aligned positions and divergent regions (Talavera and Castresana 2007). The DAMBE was used to test the nucleotide saturation (Xia and Xie 2001). Geneious 11.1.5 was selected to concatenated the individual alignment fragments. Phylogenetic analyses were conducted with three datasets: 1) P123R: all codon positions of 13 PCGs and two rRNAs; 2) P12R: first and second codon positions of PCGs and two rRNAs. Partitioned analyses for multi-gene alignments were defined by both gene types (each of 13 PCGs and two rRNAs) and codon positions (first, second, and third codon positions for each PCG). 3) AAR: all PCGs' aa sequences plus two rRNA genes. The best-fit model for each partition was determined by PartitionFinder implemented in PhyloSuite program (Lanfear et al. 2012; Zhang et al. 2020).

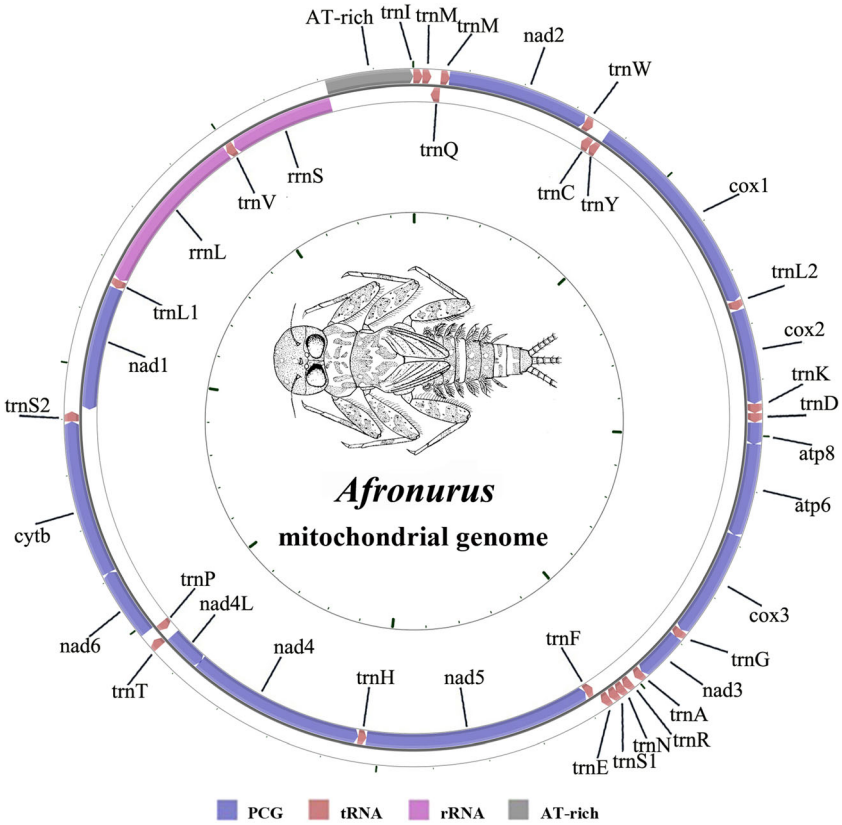

Fig. 1 Mitochondrial map of *Afronurus* genes transcribed clockwisely outside and anti-clockwisely inside

Bayesian Inference (BI) and Maximum Likelihood (ML) methods were employed for each dataset to assess whether the datasets were sensitive to the inference methods. BI analyses were performed in MrBayes 3.2.6 (Ronquist et al. 2012) through the online CIPRES Science gateway (Miller et al. 2011). Two simultaneous runs with Four independent Markov chains (three heated and one cold) were performed for 10 million generations with sampling every 1000 trees. The initial 25% trees of each run were discarded as burn-in, and the consensus tree was computed from the remaining trees. For ML analyses, RAxML 8.2.0 was employed with node support values inferring from a rapid bootstrap method applied with 1000 replications (Stamatakis 2014). The phylogenetic trees were then visualized in FigTree 1.4.2 (http://tree.bio.ed.aA.uk/software/figtree/).

Table 2 Base composition of *Afronurus* mitochondrial genomes

Gene	Species	Total (bp)	A+T (%)	AT-skew	GC-skew
Mito	Aru	15,413	62.7	0.004	−0.205
	Ayi	15,683	64.1	0.002	−0.218
	Aob	15,532	63.9	0.004	−0.205
PCGs	Aru	11,210	62.7	−0.193	−0.040
	Ayi	11,210	64.6	−0.197	−0.019
	Aob	11,210	64.0	−0.198	−0.024
rRNAs	Aru	2069	64.6	0.009	0.310
	Ayi	2065	65.6	−0.025	0.333
	Aob	2065	64.6	−0.008	0.318
tRNAs	Aru	1515	64.1	−0.013	0.107
	Ayi	1517	63.4	−0.010	0.106
	Aob	1514	64.3	−0.021	0.122
CR	Aru	574	52.3	0.060	0.204
	Ayi	846	55.0	0.131	−0.181
	Aob	704	58.5	0.117	0.123

Results and discussion

Mitogenome structure, organization

We successfully obtained the complete mitogenomes of *A. rubromaculata*, *A. yixingensis* and *A. obliquistrita*, whose size was 15,413 bp, 15,683 bp and 15,532 bp, respectively (Table 1; Online resource 1: Tables S2-S4). The size of the newly generated mitogenomes was well within the range of the completely sequenced mitogenomes of other ephemeropteran species, with sizes ranging from 14,589 bp in *Alainites yixiani* (Gui & Lu, 1999) to 16,616 bp in *Siphlonurus chinensis* Ulmer, 1920 (Online resource 1: Table S1).

The three new mitogenomes encoded a complete set of 38 genes (including 13 protein-coding genes, 23 tRNA genes and two rRNA genes) and a longest non-coding region (control region) (Fig. 1; Table 1). In addition to the 37 genes

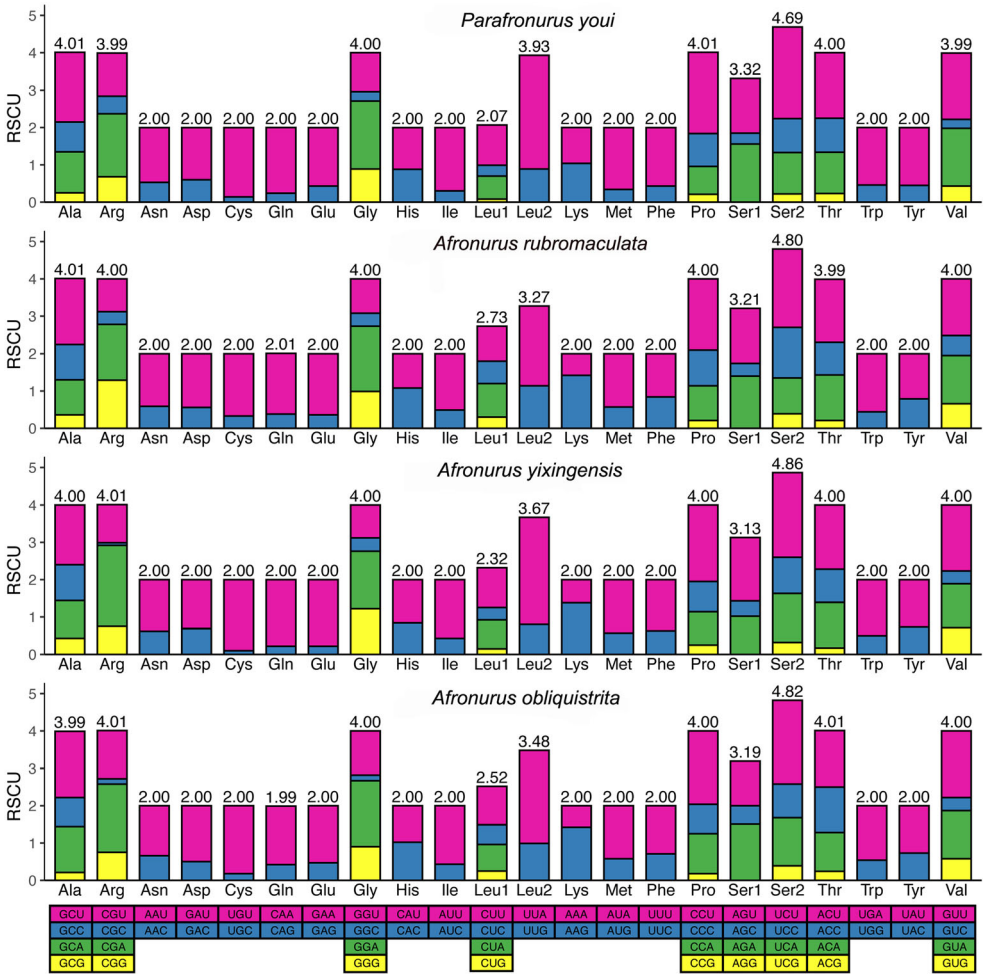

Fig. 2 The relative synonymous codon usage (RSCU) in the mitogenomes of *Parafronurus youi*, *Afronurus rubromaculata*, *A. yixingensis* and *A. obliquistrita*

commonly found in mitogenomes of insects (Wolstenholme 1992), an additional *trnM* gene was identified in the mitogenomes. Four PCGs, eight tRNAs and two rRNAs were encoded on the minority strand (N-strand), while the other 24 genes (nine PCGs and 15 tRNAs) were transcribed from the majority strand (J-strand). The gene rearrangement with tRNA cluster (*trnI-trnM-trnQ-trnM*) was found in the newly generated mitogenomes (Fig. 1), which is similar to four known sequences of heptageniid species (*P. youi*, *Epeorus herklotsi* (Hsu, 1936), *Epeorus* sp1. and *Epeorus* sp2.) (Gao et al. 2018). The conservative gene order is different to the other heptageniid sequence (*Paegniodes cupulatus* Eaton, 1871) with the typical tRNA cluster (*trnI-trnQ-trnM*) (Zhou et al. 2016). It is apparent that more mitogenomes from diverse groups of Heptageniidae are in demand to well understand the mechanism of this gene rearrangement in the following studies.

Nucleotide composition

The nucleotide composition of mitogenome was investigated using the general parameters including A + T content, AT-skew and GC-skew (Yang et al. 2018). All complete mitogenomes showed highly similar nucleotide composition biases towards A and T nucleotides, ranging from 62.7% (*A. rubromaculata*) to 64.1% (*A. yixingensis*) on the J-strand (Table 2). Additionally, skew metrics of the mitogenomes showed positive AT-skew and negative GC-skew, indicating that As and Cs were more abundant than Ts and Gs. Comprehensive analysis for all components of the *Afronurus* mitogenomes showed that they were slight deviation in nucleotide composition (Table 2), except for the AT-skew of rRNA and GC-skew of control region. The comparative analysis also revealed that the A + T content of PCGs, rRNAs and tRNAs was very similar (62.7% – 65.6%, average 64.2%), whereas the lowest in control region (52.3% – 58.5%, average 55.3%).

Protein-coding genes (PCGs) and codon usage

The total length of the 13 PCGs of our sequenced mitogenomes were all 11,210 bp (Table 2). In the three mayflies, nine PCGs (*nad2*, *nad3*, *nad6*, *cox1*, *cox2*, *cox3*, *atp6*, *atp8* and *cytb*) were encoded on the J-strand, and the remaining four PCGs (*nad1*, *nad4*, *nad4L* and *nad5*) on the N-strand, as reported in the other ephemeropterans (Gao et al. 2018).

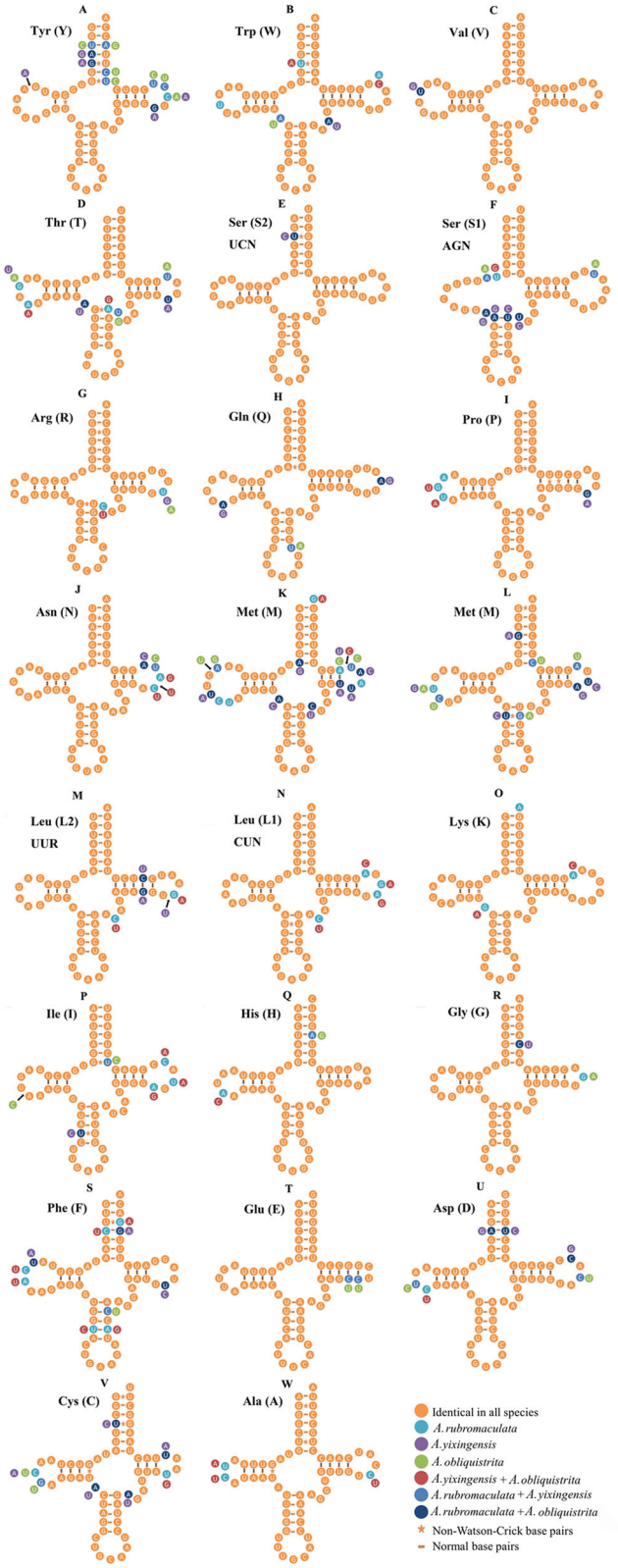

Fig. 3 The putative tRNA second structure for the three *Afronurus* mitogenomes

Predicted initiation and termination codons of all PCGs were consistent across the three mitogenomes, except for *atp8* starting with ATG for *A. rubromaculata* while with GTG for *A. yixingensis* and *A. obliquistrita* (Online resource 1: Tables S2-S4). Additionally, nine PCGs started with typical ATN codon (one with ATT, two with ATA, six with ATG), whereas *nad2* and *nad5* appeared to start with GTG, *cox1* with CCG. Nine PCGs terminated with TAN codon (two with TAG and seven with TAA), and the remaining four terminated with an incomplete stop codon TA (*nad1*) and T (*cox1*, *cox2* and *nad5*). This is a common phenomenon in the mitogenomes of metazoan, which can produce functional terminal codons via polycistronic transcription cleavage and polyadenylation processes (Ojala et al. 1981; Li et al. 2019).

The relative synonymous codon usage (RSCU) values of the three *Afronurus* and *P. youi* mitogenomes were calculated and summarized in Fig. 2 and Online resource 1: Tables S5-S7. In addition to the termination codons, the total number of codons of three *Afronurus* mitogenomes was identical (3726), while *P. youi* was different (3728). The codon of AGG was not found in all analysed sequences. Comparative analysis showed that the codon usage patterns and the major customarily utilized codons of the four mitogenomes were conservative. Meanwhile, the most frequently used amino acid were Leucine and Phenylalanine. Furthermore, RSCU analysis also indicated that codons were biased to use more A / T at the third codon (Fig. 2). Similarly, the biased usage of A + T nucleotides was reflected in the codon frequencies.

RNA genes

Two rRNA genes (*rrnL* and *rrnS*) were observed in the mitogenomes of all three mayflies. The *rrnL* gene was located between *trnL1* and *trnV*, while *rrnS* gene was between *trnV* and the control region (Fig. 1). The size of *rrnL* was 1283 bp of (*A. rubromaculata*), 1285 bp (*A. yixingensis*), and 1286 bp (*A. obliquistrita*) (Online resource 1: Tables S2-S4), with the A + T content accounting for 66.5%, 66.8% and 67.3%, respectively. The *rrnS* of *A. rubromaculata* was 786 bp long with an A + T content of 61.5%, 780 bp (63.5%) of *A. yixingensis* and 779 bp (60.2%) of *A. obliquistrita*. Two rRNAs are similar in size and positions among ephemeropteran species (Gao et al. 2018).

Unlike most ephemeropteran mitogenomes (*trnI-trnM-trnQ*), there was an extra *trnM* recognized in three new sequenced mitogenomes besides 22 typical tRNA genes, forming a conservative gene block (*trnI-trnM-trnQ-trnM*). Comparative analysis found that the duplicated *trnM* gene was found in all available heptageniid mitogenomes except for that of *P. cupulatus*. Due to limited mitogenomes, more heptageniid mitogenomes are need to understand the evolution pattern and mechanism. All of the 23 tRNA genes were dispersed among the protein-coding and the rRNA genes; 15

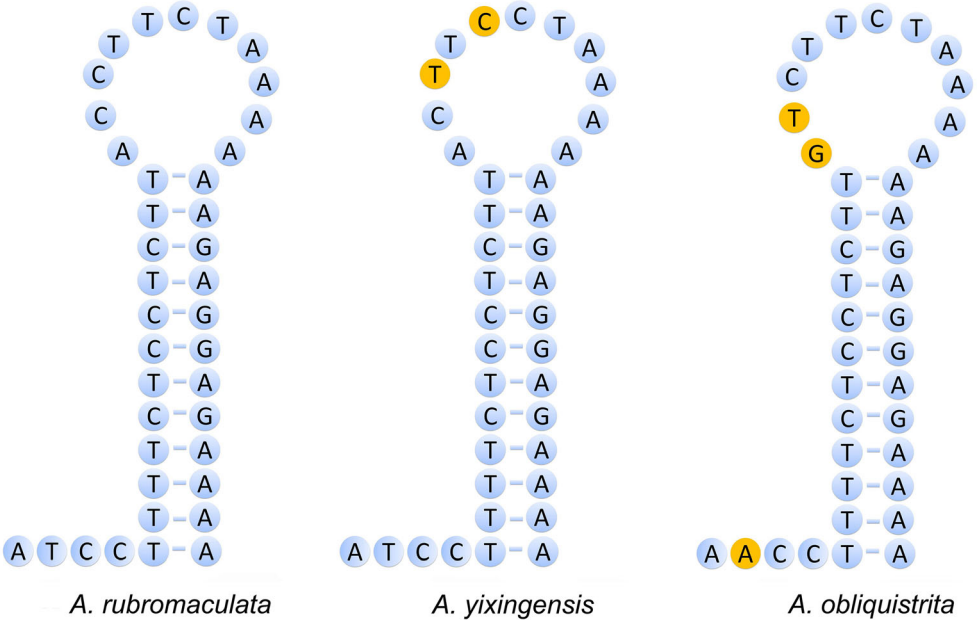

Fig. 4 The putative stem-loop structures in the longest intergenic spacer (IGS) region in *Afronurus rubromaculata*, *A. yixingensis* and *A. obliquistrita*. Yellow: different nucleotide sites

of them lay on J-strand and remaining eight on N-strand of the mitogenomes (Table 1). The length of the tRNAs ranged between 62 bp to 71 bp, with the shortest being that of *trnC* and the longest *trnV*. All tRNAs had standard anticodons and displayed the typical clover-leaf secondary structures, except for *trnS1* lacking the dihydrouridine (DHU) arm (Fig. 3F). The comparison of tRNA secondary structures showed that the nucleotide conservation of tRNAs, including *trnV*, *trnS2* and *trnE*, with the highest percentage of identical nucleotides. The extra tRNA (*trnM*) had a high level of nucleotide variation (Fig. 3K and L). In addition, apart from normal base pairs, mismatched pairs also exist in stems of tRNAs including multiple nonWatson-Crick base pairs (G-U, A-G, A-A, U-U and C-A). These mismatches are common in metazoan animals, which can be corrected through editing processes, and should not affect the transport function (Varani and McClain 2000).

Non-coding regions

In accordance with other insect mitogenomes, the different size of mitogenomes among the mayflies is mainly due to the size variation of the non-coding regions, especially the variable number and size of repeats in the control regions

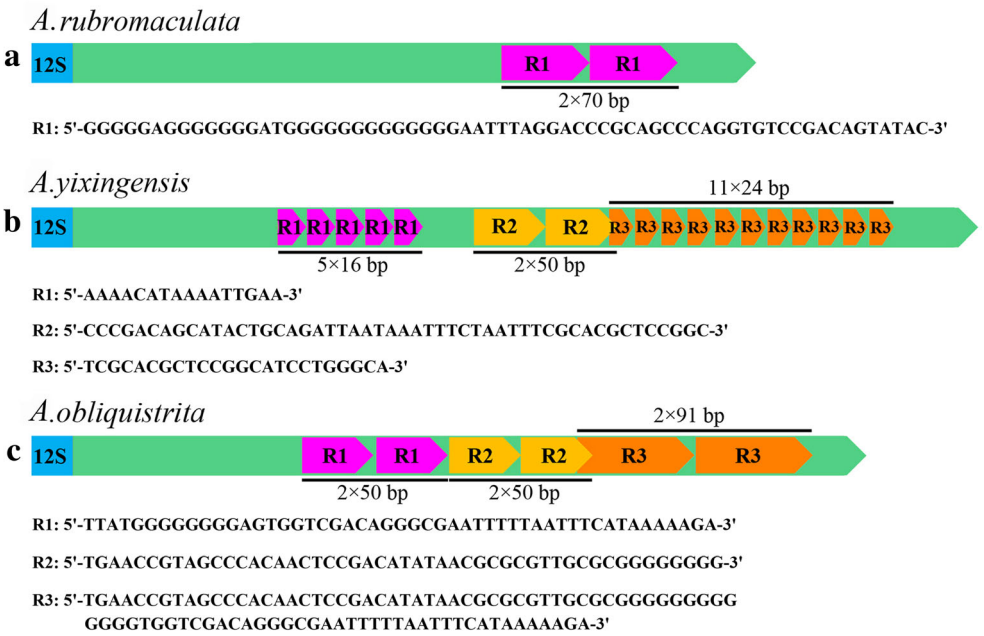

Fig. 5 Organization of the control region structure in the mitogenomes of **a** *Afronurus rubromaculata*, **b** *A. yixingensis* and **c** *A. obliquistrita*. R: repeat unit

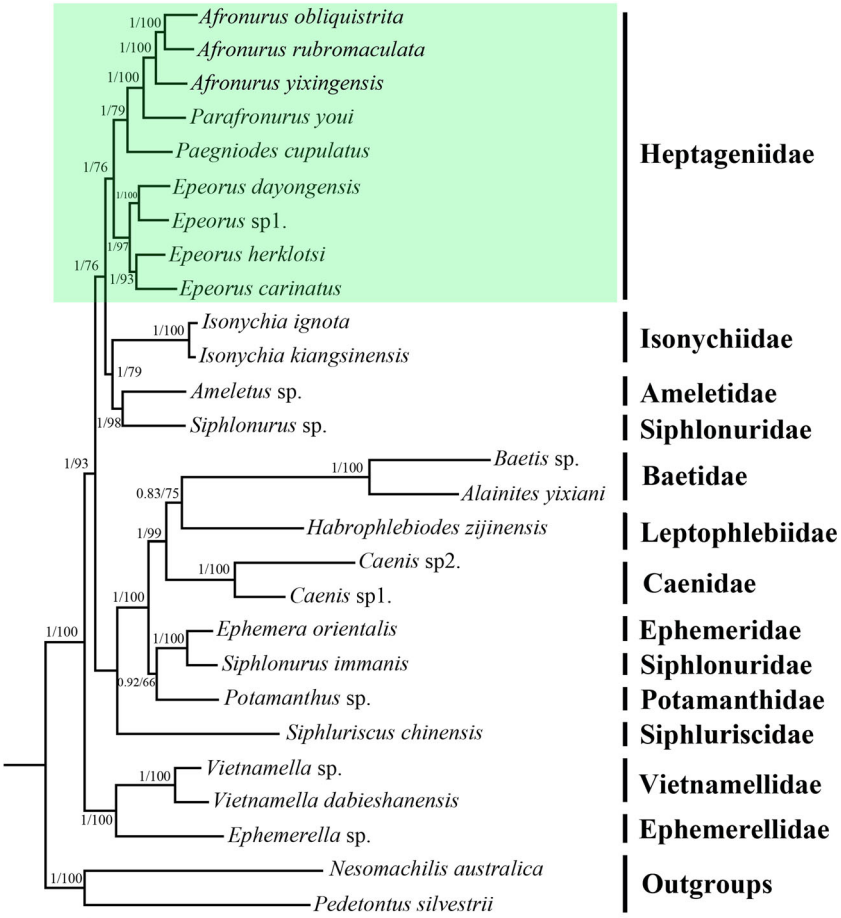

Fig. 6 A phylogenetic tree obtained from ML and BI analysis based on the P123R dataset. Numbers separated by a slash on node are posterior probability (PP) and bootstrap value (BV)

(Lv et al. 2018). While *A. obliquistrita* had eight non-coding regions with total length of 783 bp, *A. rubromaculata* and *A. yixingensis* had nine non-coding regions with total length of 657 bp and 931 bp, respectively (Table 1). In addition to the control region, a total of 79, 83 and 85 intergenic nucleotides were present in the mitogenomes of the three species, respectively. The longest intergenic spacer (IGS) region in the mitogenomes was found between the *trnA* and *trnR* genes with a size of 39 nucleotides (Online resource 1: Tables S2-S4). As shown in Fig. 4, the IGSs was rather conserved and all formed a loop-stem structure. Notably, 15 instances of gene overlap involving a total of 40 bp were found in each mitogenome of three species, which was extremely conservative in *Afronurus*.

All newly sequenced mitogenomes contained only one long non-coding region, i.e., the control region, located between the *rrnS* and the *trnI* genes (Fig. 1). The length was found to be 574 bp (*A. rubromaculata*), 846 bp (*A. yixingensis*) and 704 bp (*A. obliquistrita*) with the A + T content of 52.3%, 55.0% and 58.5%, respectively (Table 2). Because of high A + T content level, this non-coding element was generally defined as the A + T-rich region. However, the A + T content of *Afronurus* sequences was very low, which is not consistent with other Ephemeroptera mitogenomes (except for *P. youi*, 57.0%). The control region of *A. obliquistrita* possessed two copies of tandem repeats with a fragment of 50 bp, 50 bp and 91 bp, respectively (Fig. 5). The region of *A. yixingensis* possessed three tandem repetitive sequences: R1 (5 × 16 bp), R2 (2 × 50 bp) and R3 (11 × 24 bp). But the sequence of *A. rubromaculata* only contained one tandem repeat of a 70 bp fragment (2 times). The existence of tandem repeats in the mitochondrial control region has been observed in many insects, which influence the size of mitogenomes.

Phylogenetic analysis

Phylogenetic analyses were conducted with three datasets, each representing different types of data with specific sites. The dataset of P123R contained 12,346 sites including all codon positions of 13 PCGs plus two rRNAs. The P12R dataset contained 8784 sites with the first and second codon positions of 13 PCGs plus two rRNAs. The matrix of AAR contained 5221 sites with all PCGs' aa sequences and two rRNAs. The DAMBE analyses confirmed the suitability of the nucleotide's unsaturation of all datasets for phylogenetic analyses (ISS < ISS.c). Six phylogenetic trees using two inference methods (BI and ML) with best partition schemes were constructed (Figs. 6, 7, 8). Independent

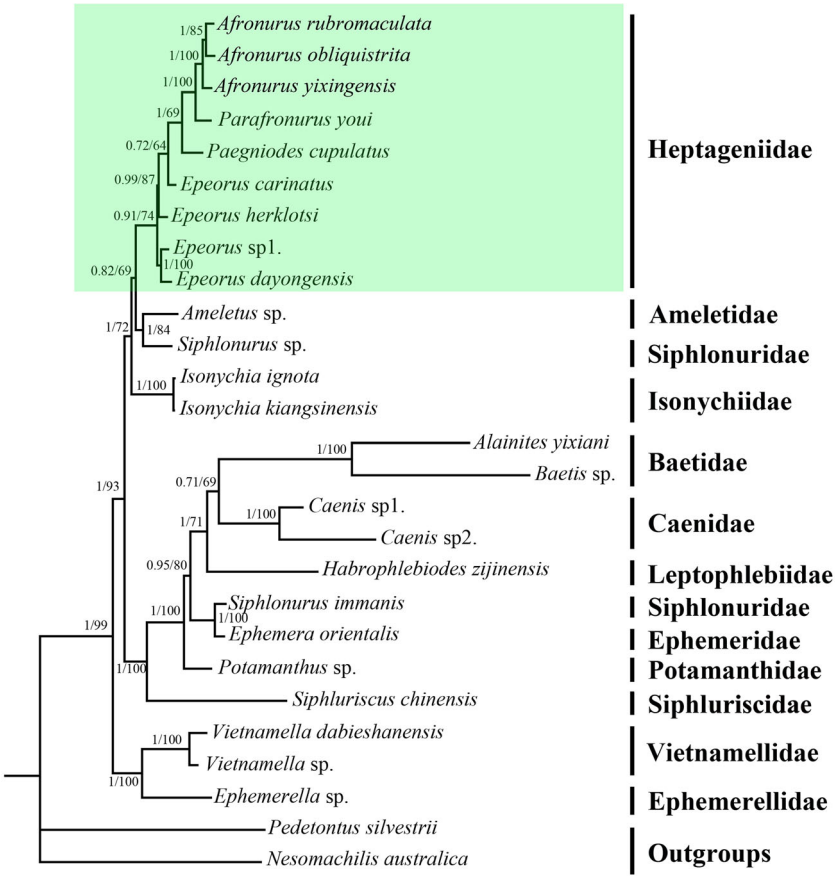

Fig. 7 A phylogenetic tree obtained from ML and BI analysis based on the P12R dataset. Numbers separated by a slash on node are posterior probability (PP) and bootstrap value (BV)

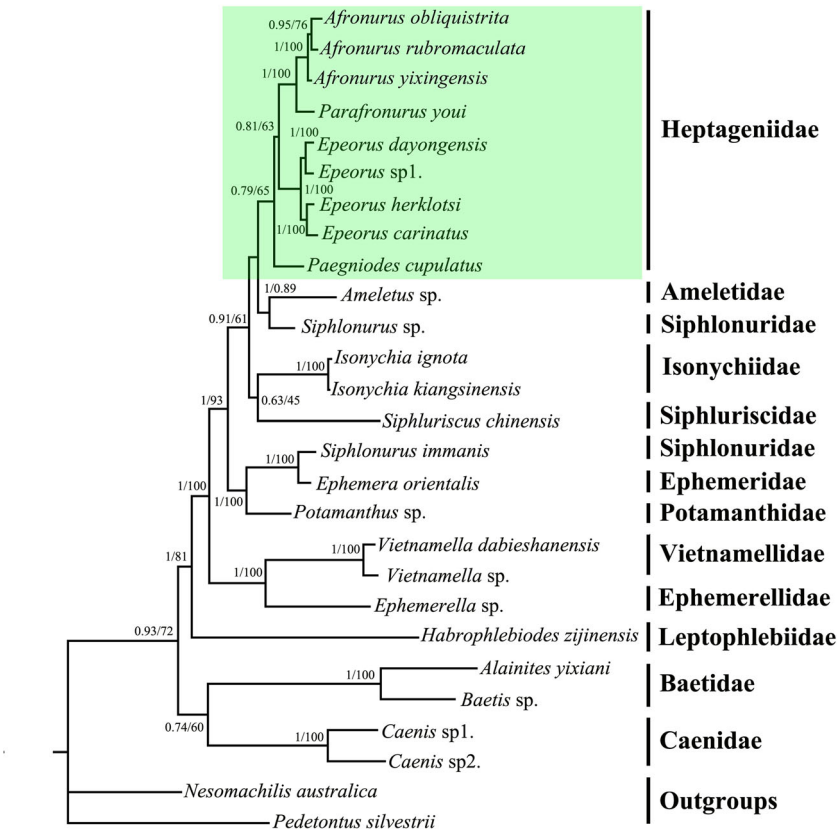

Fig. 8 A phylogenetic tree obtained from ML and BI analysis based on the AAR dataset. Numbers separated by a slash on node are posterior probability (PP) and bootstrap value (BV)

phylogenetic analysis was performed to determine the influence of different datasets and methods on tree topology and nodal support. For the same dataset, the identical topologies were produced using two different analytical methods.

As shown in Figs. 6-8, the phylogenetic analyses based on different dataset yielded obviously variable topologies. Due to the limited number of available mitogenomes for mayflies, most families contained only one species, which resulted in the unstable relationships in different trees. In all of our trees, two species within Baetidae, Caenidae, Isonychiidae and Vietnamellidae were clustered together with high nodal support values (posterior probability: 1 and bootstrap value: 100), respectively. However, two species within Siphlonuridae formed two different clades, which support the non-monophyletic Siphlonuridae in previous research (Wu and Yu 2018). To further discuss the monophyly of this family at mitogenome level, more accurate sequence and increased taxa sampling is necessary. Four phylogenetic analyses based on nucleotide sequences yielded a similar topology except for the position of Isonychiidae, Leptophlebiidae and the genus *Epeorus*. For the position of two species within Isonychiidae, which was the sister group with (*Ameletus* sp. + *Siphlomurus* sp.) based on P12R matrix, but supported as sister group of (Heptageniidae + (Ameletidae + Siphlonuridae)) based on P123R matrix. For *Habrophlebiodes zijinensis* Gui, Zhang & Wu, 1996 of Leptophlebiidae, the result placed the species as a sister group of (Baetidae + Caenidae) using P123R matrix, but Leptophlebiidae and Baetidae were consolidated into one group in the P12R analyses. In addition, the sister group relationship between Baetidae and Caenidae was supported based on P12R and AAR matrixes, but nodal support values were not significant, i.e., posterior probability <0.75 and bootstrap values <70.

In all phylogenetic analyses, the species from the family Heptageniidae formed a monophyletic group. This result was also supported by the combined analysis of morphological and molecular data (Ogden et al. 2009). The genus *Epeorus* was recovered non-monophyletic in P12R analyses, however, monophyletic in P123R and AAR analyses. Meanwhile, the same topology of four species within *Epeorus* was highly supported in the latter two analyses, which was believed more credible and as the final topology. The most controversial point in our results was the taxonomic status of the genus *Afronurus* in Heptageniidae and the phylogenetic relationships among three analysed species. Based on the morphological feature of the unique scattered setae on the ventral surface of the maxillae, *Afronurus* and *Parafronurus* belong to the subfamily Ecdyonurinae. Our results suggest that *Afronurus* is a monophyletic group, as a sister to *P. youi*, which is in accordance with the previous hypotheses in morphological taxonomy (Webb and McCafferty 2008). Among the three species of *Afronurus*, the following relationships were reconstructed with the support of high nodal values in all analyses: ((*A. rubromaculata* + *A. obliquistrita*) + *A. yixingensis*), which should lay the foundation for further population genetics and phylogeography study of *Afronurus*.

Conclusion

This study presented the complete mitogenome sequences of *A. rubromaculata*, *A. yixingensis* and *A. obliquistrita*. The three mayfly mitogenomes shared similar gene organization with an extra *trnM* gene compared with the ancestral order for insects. Nucleotide composition of the mitogenomes were relatively conservative, biased towards A and T nucleotides. Comprehensive analysis of tRNA secondary structures showed that the extra *trnM* had a high level of nucleotide variation. The longest intergenic spacer (IGS) region in the mitogenomes was found between the *trnA* and *trnR* genes with a size of 39 nucleotides and all formed a conservative loop-stem structure. Control region of *Afronurus* mitogenomes possessed different level of A + T content with the insects previously reported. The phylogenetic analyses for Ephemeroptera inferred from mitogenomes provided support for a monophyletic Heptageniidae. Meanwhile, our results elaborated the taxonomic status of the genus *Afronurus* in Ephemeroptera and the phylogenetic relationships among three analysed species at the mitogenome level. In addition, more studies as well as more mitogenomes are needed to better understand the phylogenetic relationships within Ephemeroptera in the future.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11756-021-00729-6.

Acknowledgments We thank Dr. Xiaohan Shu for helping with graphics. This work was supported by the National Natural Science Foundation of China (Grant No. 31750002) to Changfa Zhou, and the China Postdoctoral Science Foundation (Grant No. 2019M661877) to Ran Li.

Declarations

Ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors Ran Li, Wei Zhang, Zhenxing Ma, Changfa Zhou declare that they have no conflict of interest.

References

Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res 38:7–13. https://doi.org/10.1093/nar/gkq291

Avise JC (2009) Phylogeography: retrospect and prospect. J Biogeogr 36: 3–15. https://doi.org/10.1111/j.1365-2699.2008.02032.x

Barber-James HM, Gattolliat JL, Sartori M, Hubbard MD (2008) Global diversity of mayflies (Ephemeroptera, Insecta) in freshwater.

Hydrobiologia 595:339–350. https://doi.org/10.1007/s10750-007-9028-y

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69:313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27: 1767–1780. https://doi.org/10.1093/nar/27.8.1767

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

Freimark C, Jameson E, Jubera K, Schneider B (2017) Effect of abiotic factors on the diversity and abundance of aquatic macroinvertebrates in the east and west branches of the Maple River, Michigan, USA. Biological Station, Unversity of Michigan, Michigan

Gao XY, Zhang SS, Zhang LP, Yu DN, Zhang JY, Cheng HY (2018) The complete mitochondrial genome of *Epeorus herklotsi* (Ephemeroptera: Heptageniidae) and its phylogeny. Mitochondrial DNA B 3:303–304. https://doi.org/10.1080/23802359.2018.1445482

Grant JR, Stothard P (2008) The CGView server: a comparative genomics tool for circular genomes. Nucleic Acids Res 36:181–184. https://doi.org/10.1093/nar/gkn179

Kaneko H, Ishiwata SI, Liao YC, Takamura-Enya T (2018) The genetic diversities and phylogenetic relationships of two genera, *Ecdyonurus* and *Afronurus*, of Heptageniidae (Ephemeroptera) in the Yaeyama Islands and Taiwan. Biogeography 20:6–20

Katoh K, Kuma K, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33:511–518. https://doi.org/10.1093/nar/gki198

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Krzywinski J, Li C, Morris M, Conn JE, Lima JB, Povoa MM, Wilkerson RC (2011) Analysis of the evolutionary forces shaping mitochondrial genomes of a Neotropical malaria vector complex. Mol Phylogenet Evol 58:469–477. https://doi.org/10.1016/j.ympev.2011.01.003

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096

Lanfear R, Calcott B, Ho SY, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol 29:1695–1701. https://doi.org/10.1093/molbev/mss020

Li R, Shu XH, Li XD, Meng L, Li BP (2019) Comparative mitogenome analysis of three species and monophyletic inference of Catantopinae (Orthoptera: Acridoidea). Genomics 111:1728–1735. https://doi.org/10.1016/j.ygeno.2018.11.027

Lin CP, Danforth BN (2004) How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Mol Phylogenet Evol 30:686–702. https://doi.org/10.1016/S1055-7903(03)00241-0

Liu JB, Yu JJ, Zhou M, Yang JD (2019) Complete mitochondrial genome of *Japalura flaviceps*: deep insights into the phylogeny and gene rearrangements of Agamidae species. Int J Biol Macromol 125:423–431. https://doi.org/10.1016/j.ijbiomac.2018.12.068

Lv C, Li Q, Kong L (2018) Comparative analyses of the complete mitochondrial genomes of *Dosinia clams* and their phylogenetic position within Veneridae. PLoS One 13:e0196466. https://doi.org/10.1371/journal.pone.0196466

Ma ZX, Han N, Zhang W, Zhou CF (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. https://doi.org/10.1080/01650424.2018.1475675

McCafferty WP, Huff BL Jr (2017) The life cycle of the mayfly *Stenacron interpunctatum* (Ephemeroptera: Heptageniidae). The Great Lakes Entomol 11:1

Miller MA, Wayne P, Terri S (2011) The CIPRES science gateway: a community resource for phylogenetic analyses. In: Proceedings of the 2011 TeraGrid conference: extreme digital discovery. ACM digital library, Utah, USA, pp 1–8. https://doi.org/10.1145/2016741.2016785

Ogden TH, Osborne JT, Jacobus LM, Whiting MF (2009) Combined molecular and morphological phylogeny of Ephemerellinae (Ephemerellidae: Ephemeroptera), with remarks about classification. Zootaxa 1991:28–42. https://doi.org/10.1177/0192512115584008

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

Patel RK, Jain M (2012) NGS QC toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7:e30619. https://doi.org/10.1371/journal.pone.0030619

Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol 41:353–358. https://doi.org/10.1007/BF00186547

Qin J, Zhang YZ, Zhou X, Kong XB, Wei SJ, Ward RD, Zhang AB (2015) Mitochondrial phylogenomics and genetic relationships of closely related pine moth (Lasiocampidae: *Dendrolimus*) species in China, using whole mitochondrial genomes. BMC Genomics 16: 428. https://doi.org/10.1186/s12864-015-1566-5

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029

Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A (1999) Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene 238:195–209. https://doi.org/10.1016/S0378-1119(99)00270-X

Sartori M (1992) Mayflies from Israel (Insecta: Ephemeroptera) I. Heptageniidae, Ephemerellidae, Leptophlebiidae, Palingeniidae. Revue Suisse Zool 99:835–858. https://doi.org/10.5962/bhl.part.79856

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annu Rev Ecol Evol Syst 37:545–579. https://doi.org/10.1146/annurev.ecolsys.37.091305.110018

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

Sutthacharoenthad W, Sartori M, Boonsoong B (2019) Integrative taxonomy of *Thalerosphyrus* Eaton, 1881 (Ephemeroptera, Heptageniidae) in Thailand. J Nat Hist 53:1491–1514. https://doi.org/10.1080/00222933.2019.1657513

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

Varani G, McClain WH (2000) The G-U wobble base pair: a fundamental building block of RNA structure crucial to RNA function in diverse biological systems. EMBO Rep 1:18–23. https://doi.org/10.1093/embo-reports/kvd001

Vuataz L, Sartori M, Wagner A, Monaghan MT (2011) Toward a DNA taxonomy of Alpine *Rhithrogena* (Ephemeroptera: Heptageniidae) using a mixed yule-coalescent analysis of mitochondrial and nuclear DNA. PLoS One 6(5):e19728. https://doi.org/10.1371/journal.pone.0019728

Wang TQ, McCafferty WP (2004) Heptageniidae (Ephemeroptera) of the world. I. Phylogenetic higher classification. Trans Am Entomol Soc 130:11–45

Webb JM, McCafferty WP (2008) Heptageniidae of the world. Part II. Key to the genera. Canadian J Arthrop Ident 7:10.37551

Wolstenholme DR (1992) Genetic novelties in mitochondrial genomes of multicellular animals. Curr Opin Genet Dev 2:918–925. https://doi.org/10.1016/S0959-437X(05)80116-9

Wu MJ, Yu LL (2018) The complete mitochondrial genome of *Epeorus herklotsi* (Ephemeroptera: Heptageniidae) from Longquan, Zhejiang, China and its phylogeny. Mitochondrial DNA B 3:1254–1255. https://doi.org/10.1080/23802359.2018.1532351

Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. J Heredity 92:371–373. https://doi.org/10.1093/molbev/mst064

Yanai Z, Zohar S, Dor M, 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. https://doi.org/10.1111/syen.12203

Yang QQ, Liu SW, Song F, Liu GF, Yu XP (2018) Comparative mitogenome analysis on species of four apple snails (Ampullariidae: Pomacea). Int J Biol Macromol 118:525–533. https://doi.org/10.1016/j.ijbiomac.2018.06.092

Yuan S, Xia Y, Zheng Y, Zeng X (2016) Next-generation sequencing of mixed genomic DNA allows efficient assembly of rearranged mitochondrial genomes in *Amolops chunganensis* and *Quasipaa boulengeri*. PeerJ 4:e2786. https://doi.org/10.7717/peerj.2786

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020) PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Res 20:348–355. https://doi.org/10.1111/1755-0998.13096

Zhou D, Wang YY, Sun JZ, Han YK, Zhou CF (2016) The complete mitochondrial genome of *Paegniodes cupulatus* (Ephemeroptera: Heptageniidae). Mitochondrial DNA A 27:925–926. https://doi.org/10.3109/19401736.2014.926488