The Complete Mitochondrial Genome of *Corizus tetraspilus* (Hemiptera: Rhopalidae) and Phylogenetic Analysis of Pentatomomorpha

Ming-Long Yuan*[☯], Qi-Lin Zhang[☯], Zhong-Long Guo, Juan Wang, Yu-Ying Shen

State Key Laboratory of Grassland Agro-Ecosystems, College of Pastoral Agricultural Science and Technology, Lanzhou University, Lanzhou, Gansu, People's Republic of China

☯ These authors contributed equally to this work.

* yuanml@lzu.edu.cn

Abstract

Insect mitochondrial genome (mitogenome) are the most extensively used genetic information for molecular evolution, phylogenetics and population genetics. Pentatomomorpha (>14,000 species) is the second largest infraorder of Heteroptera and of great economic importance. To better understand the diversity and phylogeny within Pentatomomorpha, we sequenced and annotated the complete mitogenome of *Corizus tetraspilus* (Hemiptera: Rhopalidae), an important pest of alfalfa in China. We analyzed the main features of the *C. tetraspilus* mitogenome, and provided a comparative analysis with four other Coreoidea species. Our results reveal that gene content, gene arrangement, nucleotide composition, codon usage, rRNA structures and sequences of mitochondrial transcription termination factor are conserved in Coreoidea. Comparative analysis shows that different protein-coding genes have been subject to different evolutionary rates correlated with the G+C content. All the transfer RNA genes found in Coreoidea have the typical clover leaf secondary structure, except for *trnS1* (AGN) which lacks the dihydrouridine (DHU) arm and possesses a unusual anticodon stem (9 bp vs. the normal 5 bp). The control regions (CRs) among Coreoidea are highly variable in size, of which the CR of *C. tetraspilus* is the smallest (440 bp), making the *C. tetraspilus* mitogenome the smallest (14,989 bp) within all completely sequenced Coreoidea mitogenomes. No conserved motifs are found in the CRs of Coreoidea. In addition, the A+T content (60.68%) of the CR of *C. tetraspilus* is much lower than that of the entire mitogenome (74.88%), and is lowest among Coreoidea. Phylogenetic analyses based on mitogenomic data support the monophyly of each superfamily within Pentatomomorpha, and recognize a phylogenetic relationship of (Aradoidea + (Pentatomoidea + (Lygaeoidea + (Pyrrhocoroidea + Coreoidea)))).

Introduction

The insect mitochondrial genome (mitogenome) is a circular double-strand molecule of 15–18 kb in size and usually codes for 37 genes: 13 protein-coding genes (PCGs), two ribosomal RNA genes (rRNAs), and 22 transfer RNA genes (tRNAs) [1, 2]. In addition, mitogenome usually contains a large non-coding region, known as control region (also called A+T-rich region in insects due to high A+T content). This region contains essential regulatory elements for transcription and replication [1, 3], and has been identified as the source of size variation in the whole mitogenome [4]. Compared to single mitochondrial gene, mitogenome contains even more genetic information and provides genome-level features (e.g. gene rearrangements and RNA secondary structures) [5–8]. Due to its maternal inheritance, relatively rapid evolutionary rate, and lack of genetic recombination, mitogenome sequences have been extensively used in the study of molecular evolution, phylogenetics, phylogeography and population genetics [2, 9–11].

Pentatomomorpha, which consists of 40 families representing more than 14,000 species, is the second largest among the seven infraorders of Heteroptera [12]. Most members of this group are phytophagous, and economically important in agriculture and forestry. Currently, the classification system of Pentatomomorpha includes five superfamilies (Aradoidea, Pentatomoidea, Coreoidea, Lygaeoidea and Pyrrhocoroidea), and the superfamilies except the Aradoidea are grouped as Trichophora [12, 13]. Based on morphological and molecular evidence, the hypothesis of (Aradoidea + (Pentatomoidea + the remainder of Trichophora)) has been accepted by most researchers [13–17]. However, the phylogenetic relationships among the superfamilies within the Trichophora are still controversial.

To date, complete or nearly complete mitogenomes have been determined for 25 species from Pentatomomorpha (GenBank, September 10, 2014), of which four are from Coreoidea. In this study, we sequenced and annotated the complete mitogenome sequence of *Corizus tetraspilus* (Hemiptera: Rhopalidae), an important pest of alfalfa in China. We analyzed the main features of *C. tetraspilus* mitogenome, including nucleotide composition, codon usage, rRNA structures and evolutionary pattern of PCGs, and provided a comparative analysis with four other Coreoidea species. To investigate the phylogenetic relationships among the superfamilies of Pentatomomorpha, we also performed phylogenetic analyses with Bayesian inference (BI) and maximum likelihood (ML) methods using the concatenated nucleotide sequences of 13 mitochondrial PCGs and 24 RNA genes.

Materials and Methods

Ethics Statement

No specific ethics permits were required for the described studies. The insect specimens were collected from alfalfa field by net sweeping, and no specific permissions were required for these locations/activities. The species in our study is an agricultural pest and are not included in the "List of Protected Animals in China".

Sample and DNA Extraction

Adult specimens of *C. tetraspilus* were collected from alfalfa field in Shishe Town, Xifeng District, Qingyang City, Gansu Province, China, in July 2013. Samples and voucher specimens have been deposited in the State Key Laboratory of Grassland Agro-Ecosystems, College of Pastoral Agricultural Science and Technology, Lanzhou University, Lanzhou, China. All specimens were initially preserved in 100% ethanol in the field, and transferred to -20°C until used for DNA extraction. The total genomic DNA was extracted from thorax muscle of a single

specimen using the OMEGA Insect DNA Kit (OMEGA, USA) according to the manufacturer's protocols.

PCR Amplification and Sequencing

The whole mitogenomes of *C. tetraspilus* were amplified with ten overlapping fragments by using universal insect mitogenome primers [9] and species-specific primer pairs designed from sequenced fragments (S1 Table). PCRs were performed with TaKaRa LA Taq under the following conditions: 2 min initial denaturation at 92°C, followed by 35 cycles of 10 s at 92°C, 1 min at 48–55°C, and 1–3 min at 68°C, and a final elongation for 10 min at 68°C. All PCR products were electrophoresed on a 1.2% agarose gel, purified, and then directly sequenced or cloned into the pEASY-T1 vector (TransGen Biotech, Beijing, China). All fragments were sequenced in both directions on an ABI3730 automated sequencer.

Annotation and Sequence Analysis

Sequence files were proof read and assembled into contigs with BioEdit 7.0.9.0 [18]. PCGs were identified by ORF Finder implemented at the NCBI website with the invertebrate mitochondrial genetic codes. To ensure the accurate boundaries of different genes, PCGs and rRNAs were aligned with the sequenced mitochondrial sequences of other true bugs using Muscle as implemented in MEGA 6.06 [19]. The tRNAs were predicted by their cloverleaf secondary structure using tRNAscan-SE 1.21 [20]. Some tRNAs not detected by tRNAscan-SE were determined in the unannotated regions by sequence similarity to tRNAs of other true bugs.

Nucleotide composition and codon usage were analyzed with MEGA 6.06 [19]. The number of synonymous substitutions per synonymous site (Ks), the number of nonsynonymous substitutions per nonsynonymous site (Ka), the effective number of codons (ENC) and codon bias index (CBI) for each PCG were determined with DnaSP 5.0 [21]. Strand asymmetry was calculated using the formulas: AT-skew = [A-T]/[A+T] and GC-skew = [G-C]/[G+C] [22]. The tandem repeats of the control region were identified by tandem repeats finder online server (http://tandem.bu.edu/trf/trf.html) [23].

Construction of rRNA Secondary Structure

The secondary structure of the large and small subunits of rRNAs (*rrnL* and *rrnS*) were inferred following the models proposed for other insects, *Drosophila melanogaster* (Diptera: Drosophilidae) [24], *Apis mellifera* (Hymenoptera: Apidae) [25], and *Manduca sexta* (Lepidoptera: Sphingidae) [26]. Helix numbering follows the convention established at the CRW site [24]. Regions lacking significant homology and other non-coding regions were folded using the Mfold Web Server [27].

Phylogenetic Analysis

Twenty-five Pentatomomorpha species with complete or nearly complete mitogenomes were used in phylogenetic analyses, representing five superfamilies and seventeen families. Two species, *Adelphocoris fasciaticollis* and *Apolygus lucorum* from Cimicomorpha, were used as outgroups. Details of the species used in this study were listed in Table 1.

Sequences of 13 PCGs (excluding stop codons), two rRNAs and 22 tRNAs were used for phylogenetic analyses. Each PCG was aligned individually with codon-based multiple alignments using MAFFT as implemented in the TranslatorX online server [28]. Gaps and ambiguous sites were removed from the protein alignment before back-translate to nucleotides using GBlocks within the TranslatorX with default settings. The rRNA genes were aligned with

Table 1. List of the species included in the present study.

Infraorder	Superfamily/family	Species	Size (bp)	Accession number	Reference
Cimicomorpha	Miroidea				
	Miridae	*Adelphocoris fasciaticollis*	15,434	NC_023796	[65]
		Apolygus lucorum	14,768	NC_023083	[46]
Pentatomomorpha	Aradoidea				
	Aradidae	*Aradacanthia heissi*	15,528	HQ441233	[66]
		Brachyrhynchus hsiaoi	15,250	NC_022670	[67]
		Neuroctenus parus	15,354	NC_012459	[13]
	Coreoidea				
	Alydidae	*Riptortus pedestris*	17,191	NC_012462	[13]
	Coreidae	*Hydaropsis longirostris*	16,521	NC_012456	[13]
	Rhopalidae	*Corizus tetraspilus*	14,989	KM983397	This study
		Aeschyntelus notatus	14,532	NC_012446*	[13]
		Stictopleurus subviridis	15,319	NC_012888	[56]
	Lygaeoidea				
	Berytidae	*Yemmalysus parallelus*	15,747	NC_012464	[13]
	Colobathristidae	*Phaenacantha marcida*	14,540	NC_012460*	[13]
	Geocoridae	*Geocoris pallidipennis*	14,592	NC_012424*	[13]
	Lygaeidae	*Kleidocerys resedae resedae*	14,688	KJ584365	[68]
	Malcidae	*Chauliops fallax*	15,739	NC_020772	[37]
		Malcus inconspicuus	15,575	NC_012458	[13]
	Pentatomoidea				
	Cydnidae	*Macroscytus gibbulus*	14,620	EU427338*	[13]
	Dinidoridae	*Coridius chinensis*	14,648	JQ739179*	[69]
	Pentatomidae	*Dolycoris baccarum*	16,549	NC_020373	[47]
	Pentatomidae	*Halyomorpha halys*	16,518	NC_013272	[70]
	Pentatomidae	*Nezara viridula*	16,889	NC_011755	[13]
	Plataspidae	*Coptosoma bifaria*	16,179	NC_012449	[13]
	Plataspidae	*Megacopta cribraria*	15,647	NC_015342	Direct Submission
	Tessaratomidae	*Eusthenes cupreus*	16,229	NC_022449	[71]
	Urostylididae	*Urochela quadrinotata*	16,587	NC_020144	[72]
	Pyrrhocoroidea				
	Largidae	*Physopelta gutta*	14,935	NC_012432	[13]
	Pyrrhocoridae	*Dysdercus cingulatus*	16,249	NC_012421	[13]

*Incomplete mitochondrial genome.

doi:10.1371/journal.pone.0129003.t001

MAFFT (http://mafft.cbrc.jp/alignment/server/) using the Q-INS-i method [29], and poorly aligned positions and divergent regions were removed using GBlocks Server (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) with allowing gap positions within the final blocks. Each tRNA was aligned using ClustalW implemented in MEGA 6.06 [19], and the resulting alignments of tRNA were carefully adjusted by eye according to the secondary structures. Alignments of individual genes were then concatenated as two datasets: 1) sequences of 13 PCGs (PCG) with 10,422 residues, and 2) sequences of 13 PCGs, 2 rRNA and 22 tRNAs (PCGRNA) with 13,406 residues. To determine if sequence saturation exists in our alignments we performed a test of substitution saturation using DAMBE 5.3.74 [30]. Saturation plots indicated that no substitution saturation was found for each data partition, even in the third

position of 13 PCGs (S1 Fig). Therefore, all sites for 13 PCGs, 2 rRNAs and 22 tRNAs were used in phylogenetic analyses.

The best partitioning schemes and corresponding nucleotide substitution models for each dataset were selected by PartitionFinder 1.1.1 [31]. We created data blocks based on genes and/ or codon positions, i.e. 39 partitions for the PCG dataset or 42 for the PCGRNA dataset. We used the Bayesian information criterion (BIC) and the "greedy" algorithm with branch lengths estimated as "unlinked" to search for the best-fit scheme (S2 Table). The best-fit partitioning schemes determined by PartitionFinder were implemented in the following phylogenetic analyses.

Phylogenetic analyses were performed with ML and BI methods available on the CIPRES Science Gateway 3.3 [32]. ML analysis was conducted with RAxML-HPC2 on XSEDE 8.0.24 [33] using GTRGAMMAI model, and 1000 bootstraps (BS) were used to estimate the node reliability. Bayesian analysis was carried out using MrBayes 3.2.2 [34] on XSEDE. wo independent runs with four chains (three heated and one cold) were conducted simultaneously for 10,000,000 generations, with sampling every 100 generations. Stationarity is considered to be reached when ESS (estimated sample size) value is above 100 and PSRF (potential scale reduction factor) approach 1.0 as MrBayes 3.2.2 suggested [34]. After discarding the first 25% samples as burn-in, posterior probabilities (PP) were calculated in a consensus tree.

Results and Discussion

Genome Organization

The mitogenome of *C. tetraspilus* is a typical circular DNA molecule of 14,989 bp in size (Gen-Bank accession no. KM983397; Fig 1, Table 2). This mitogenome is the smallest among the five sequenced Coreoidea mitogenomes (Table 1), primarily due to the significant size reduction of the putative control region. The mitogenome of *C. tetraspilus* contains a typical set of 37 mitochondrial genes (13 PCGs, 22 tRNA genes, 2 rRNA genes) and a large non-coding region (putative control region) (Fig 1, Table 2). The order and orientation of the mitochondrial genes is identical to that of the putative ancestral insect mitogenome [1].

The *C. tetraspilus* mitogenome is highly compact in genome size as that in other animals [1]. Gene overlaps have been observed at six gene junctions and involved a total of 27 nucleotides, ranging from 1 to 8 nucleotides (Fig 1, Table 1). The longest overlap (8 bp) exists between *trnW* and *trnC*, which are also present in other Coreoidea species and highly conserved with the same size (S2 Fig). Two PCG pairs *atp8*/*atp6* and *nad4L*/*nad4* overlap seven identical nucleotides in all the five Coreoidea mitogenomes (S2 Fig).

Nucleotide Composition and Codon Usage

The nucleotide composition of the *C. tetraspilus* mitogenome is significantly biased toward A and T. The total A+T content of the J-strand is 74.88%, which is slightly lower than those of other completely sequenced Coreoidea species (S3 Fig). Among 13 PCGs, the lowest A+T content is 69.02% in *cox1*, while the highest is 86.27% in *atp8*. The analysis of the nucleotide composition at each codon position of the concatenated 13 PCGs of *C. tetraspilus* demonstrates that the third codon position (86.19%) has an A+T content higher than that of the first (70.34%) and second (67.46%) positions. The similar nucleotide composition patterns are also observed in other Coreoidea species (S3 Fig).

The *C. tetraspilus* mitogenome has more As and Cs, indicating a positive AT-skew (0.14) and a negative GC-skew (-0.19). The PCGs and rRNAs have a negative AT-skew and a positive GC-skew in the five Coreoidea species. For the most species of Coreoidea, values for both AT-skew and GC-skew of the second and third codon positions of PCGs are negative, whereas
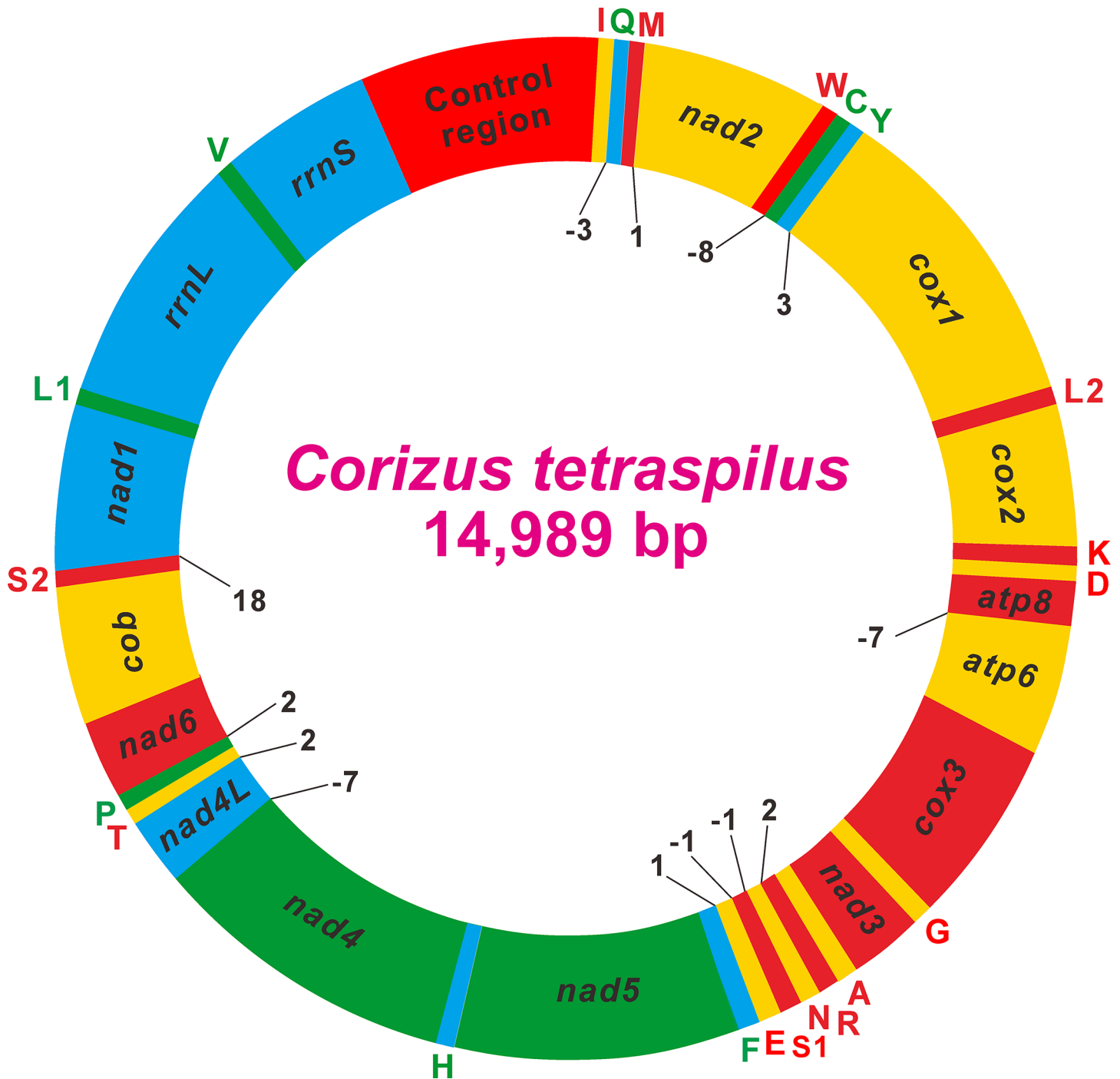
Fig 1. Circular map of the mitochondrial genome of *Corizus tetraspilus*. Protein coding and ribosomal genes are shown with standard abbreviations. Genes for tRNAs are abbreviated by a single letter, with S1 = AGN, S2 = UCN, L1 = CUN, and L2 = UUR. Genes coded in the J-strand (clockwise orientation) are red or orange colored. Genes coded in the N-strand (counterclockwise orientation) are green or cyan colored. Numbers at gene junctions indicate the length of small non-coding regions where negative numbers indicate overlap between genes.

doi:10.1371/journal.pone.0129003.g001

AT- and GC-skews of the first position are positive (S3 Fig). However, negative AT-skew of the first position and positive AT-skew of the third are found in *Hydaropsis longirostris* and *C. tetraspilus*, respectively, which are different from those of the other species.

Table 2. Annotation and organization of the complete mitogenome of *Corizus tetraspilus*.

Gene	Strand	Position	Length (bp)	Intergenic nucleotides	Start codon	Stop codon	Anticodon
trnI	J	1–65	65	0			GAU
trnQ	N	63–131	69	-3			UUG
trnM	J	132–200	69	0			CAU
nad2	J	202–1198	997	1	ATG	T	
trnW	J	1199–1267	69	0			UCA
trnC	N	1260–1322	63	-8			GCA
trnY	N	1323–1387	65	0			GUA
cox1	J	1391–2924	1534	3	TTG	T	
trnL2(UUR)	J	2925–2989	65	0			UAA
cox2	J	2990–3668	679	0	ATT	T	
trnK	J	3669–3742	74	0			CUU
trnD	J	3743–3807	65	0			GUC
atp8	J	3808–3963	156	0	ATA	TAA	
atp6	J	3957–4624	668	-7	ATG	TA	
cox3	J	4625–5411	787	0	ATG	T	
trnG	J	5412–5474	63	0			UCC
nad3	J	5475–5826	352	0	ATA	T	
trnA	J	5827–5889	63	0			UGC
trnR	J	5891–5950	60	1			UCG
trnN	J	5954–6020	67	3			GUU
trnS1(AGN)	J	6020–6088	69	-1			GCU
trnE	J	6088–6154	67	-1			UUC
trnF	N	6156–6222	67	1			GAA
nad5	N	6223–7934	1712	0	ATC	TA	
trnH	N	7935–7998	64	0			GUG
nad4	N	7999–9314	1316	0	ATG	TA	
nad4L	N	9308–9598	291	-7	ATT	TAA	
trnT	J	9601–9663	63	2			UGU
trnP	N	9664–9726	63	0			UGG
nad6	J	9729–10216	488	2	ATA	TA	
cob	J	10217–11348	1132	0	ATG	T	
trnS2(UCN)	J	11349–11417	69	0			UGA
nad1	N	11436–12362	927	18	ATT	TAA	
trnL1(CUN)	N	12363–12429	67	0			UAG
rrnL	N	12430–13695	1266	0			
trnV	N	13696–13763	68	0			UAC
rrnS	N	13764–14549	786	0			
control region	J	14550–14989	440	0			

doi:10.1371/journal.pone.0129003.t002

Excluding termination codons, the 13 PCGs in the mitogenome are composed of 3,672 codons in total (Fig 2). Approximately equivalent codon numbers are found in other four Coreoidea species, ranging from 3,671 in *Aeschyntelus notatus* to 3,679 in *Stictopleurus subviridis* (Fig 2). The codon families exhibit the same pattern of codon usage as elsewhere in the five Coreoidea species (Figs 2 and 3). The four most predominant codon families are Leu2 (UUR), Ile, Phe, and Met, each of which has at least 80 codons (CDs) per thousand CDs. For the RSCU in the mitogenomes of five Coreoidea species, the six most frequently used codons, TTT (F), TTA (L), ATT (I), ATA (M), TAT (Y) and AAT (N), are all completely composed of A and/or T,
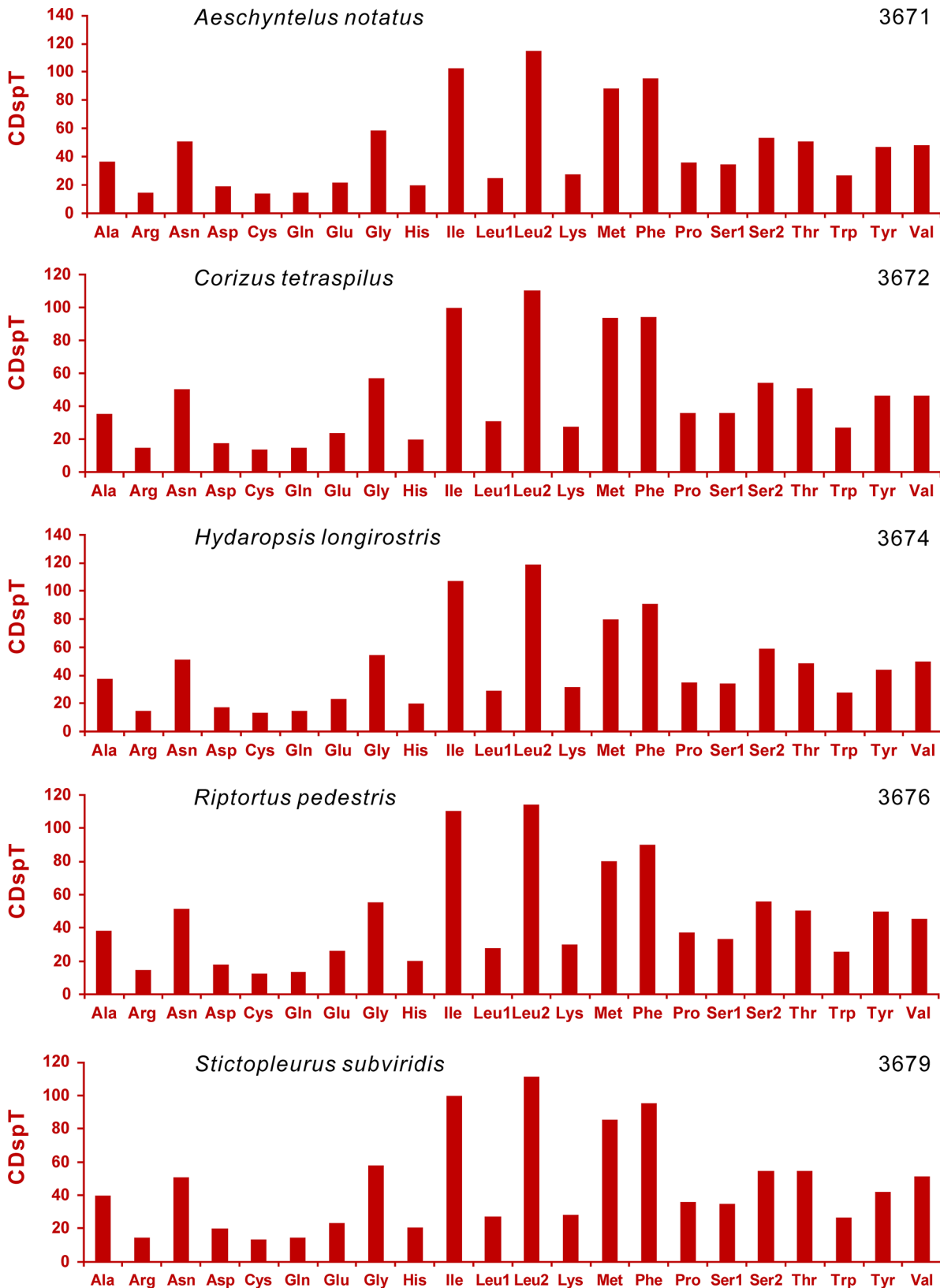
Fig 2. Codon distributions in the mitochondrial genomes of five Coreoidea species. Numbers to the left refer to the total number of codon. CDspT, codons per thousands codons.
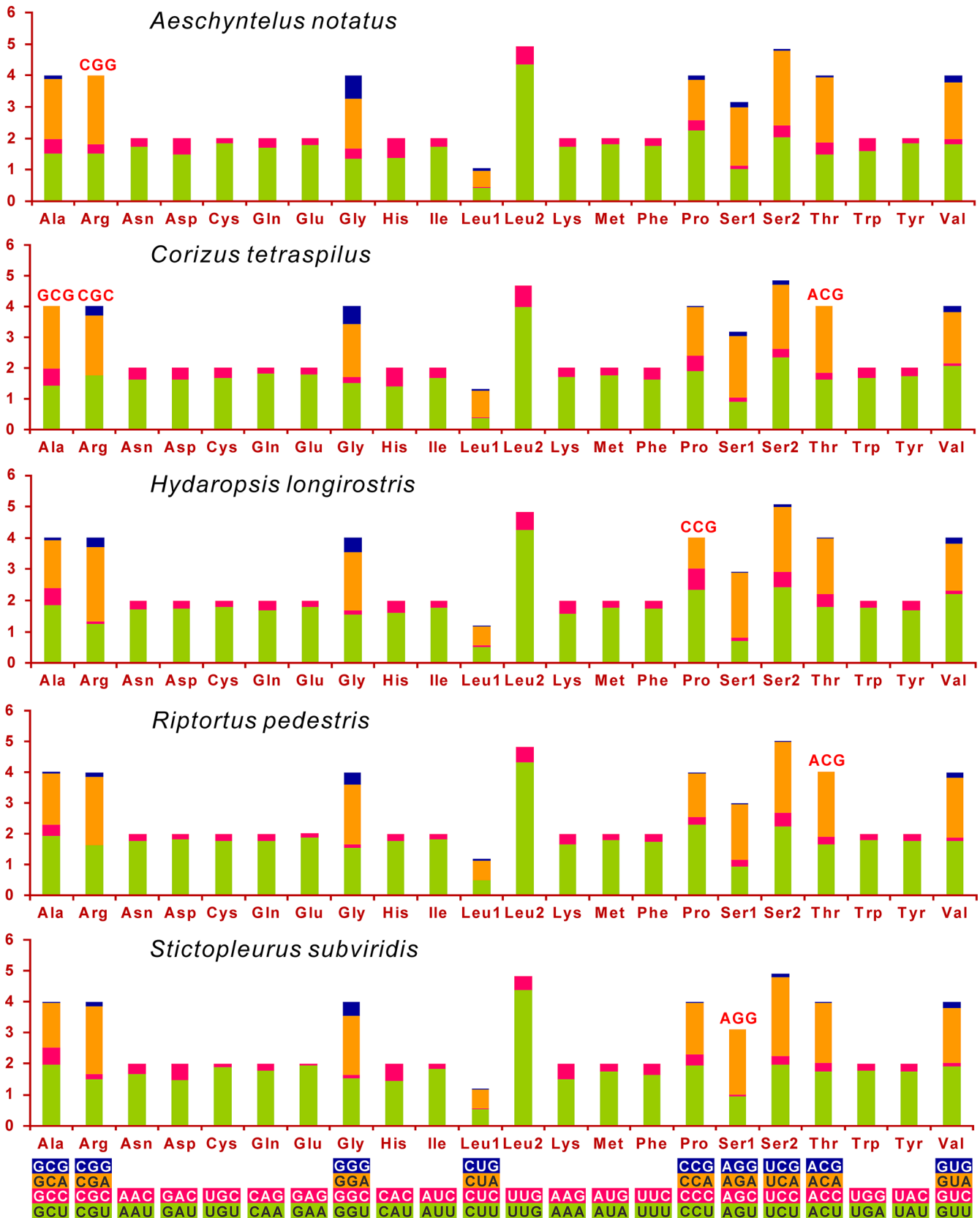
doi:10.1371/journal.pone.0129003.g002

Fig 3. Relative Synonymous Codon Usage (RSCU) in the mitochondrial genomes of five Coreoidea species. Codons that are not present in the genome are indicated in red. Codon Families are provided on the x axis.

which reflects a strong compositional bias toward A+T. The four- and two-fold degenerate codon usages are biased to use more As and Ts than Gs and Cs in the third codon positions (Fig 3). Furthermore, three GC-rich codon families, i.e. GCG (A), CGC (R) and ACG (T), are not utilized in *C. tetraspilus*, whereas only one GC-rich codon is not used in each of other Coreoidea species (Fig 3).

The correlations between ENC, CBI, the G+C content of all codons, and the G+C content of the 3rd codon positions in all sequenced Coreoidea mitogenomes are analyzed (S4 Fig). A positive correlation is observed between ENC and G+C content of all codons ($R^2 = 0.82$) (S4A Fig) and the 3rd codon positions ($R^2 = 0.94$) (S4B Fig). Furthermore, a negative correlation is found between CBI and G+C content of all codons ($R^2 = 0.93$) (S4C Fig), G+C content of the 3rd codon positions ($R^2 = 0.88$) (S4D Fig), and ENC ($R^2 = 0.86$) (S4E Fig). These results are consistent with prevailing neutral mutational theories positing that genomic G+C content is the most significant factor in determining codon bias among organisms [35, 36].

Protein Coding Genes

Twelve of the 13 PCGs start with a typical ATN codon: one (*nad5*) with ATC, three (*cox2*, *nad4L* and *nad1*) with ATT, three (*atp8*, *nad3*, and *nad6*) with ATA, and five (*nad2*, *atp6*, *cox3*, *nad4* and *cob*) with ATG. The only exception is *cox1*, which uses TTG as a start codon. This unconventional codon has also been commonly found in the other Coreoidea mitogenomes (S3 Table) and many other true bugs [37–39]. Four PCGs (*atp6*, *atp8*, *nad1* and *nad4L*) have complete stop codon TAA, while the remaining nine terminate with either TA (*nad4*, *nad5* and *nad6*) or T (*cox1*, *cox2*, *cox3*, *cob*, *nad2* and *nad3*). Incomplete stop codons were also observed in the other Coreoidea species (S3 Table), and it has been proposed that the complete stop codon TAA could be generated via post-transcriptional polyadenylation [40, 41].

The evolutionary patterns among the mitochondrial PCGs in Coreoidea are different (Fig 4). The Ks of *cob* is the highest, but its value of Ka is much lower, while the values of Ka and ω
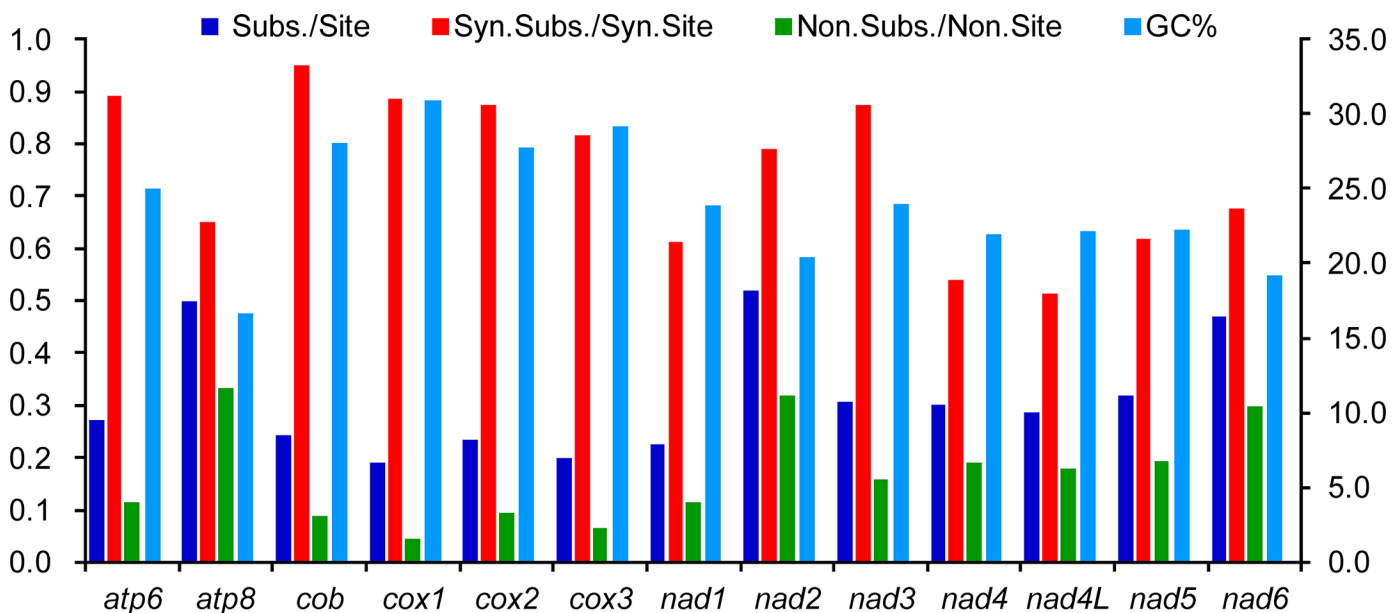

Fig 4. Evolutionary rates of 13 protein-coding genes in the mitochondrial genomes of five Coreoidea species.

for *atp8* are the highest. The *cox1* gene has been widely used as a DNA barcode in true bugs [42–44], but this gene shows the lowest evolutionary rates, compared to other genes. Similarly, *cob*, *cox2* and *cox3* also show relatively slow revolutionary rates. By contrast, the nucleotide substation rate per site and Ka values of *nad2* and *nad6* are only lower than that of *atp8*, indicating that these two genes may be potential barcoding markers in Coreoidea. The *ω* values for all PCGs are far lower than one (< 0.52), indicating that these genes are evolving under the purifying selection. Therefore, all mitochondrial PCGs can be employed to investigate phylogenetic relationships within Coreoidea. Furthermore, a negative correlation has been found between the *ω* and the G+C content of each PCG ($R^2 = 0.93$), indicating that the variation of G+C content probably causes the different evolutionary patterns among genes.

Transfer and Ribosomal RNAs

All the typical 22 tRNAs are found in the *C. tetraspilus* mitogenome, with size ranging from 60 bp to 74 bp (Table 2). The secondary structures of *C. tetraspilus* tRNAs are consistent with other Coreoidea species (S5 Fig). All the tRNAs could be folded into classic cloverleaf secondary structures (Fig 5), with the exception of *trnS1* (AGN) that lack the dihydrouridine (DHU) arm. The loss of the DHU arm in *trnS1* is common in insect mitogenomes [2], and has been considered a typical feature of metazoan mitogenomes. In addition, *trnS1* possesses a unusual anticodon stem (9 bp vs. the normal 5 bp) and a bulged nucleotide in the middle of the anticodon stem. Although this structure of *trnS1* is abnormal, it is highly conserved within all sequenced Coreoidea mitogenomes, especially for the anticodon arm (S5 Fig). This phenomenon found in *trnS1* has been widely reported for many other hemipterans [37, 45–48]. Furthermore, six mismatched pairs (3 U-U, 3 C-U) and 19 G-U wobble pairs are present in 7 aminoacyl acceptor stems, 10 DHU stems, 5 anticodon stems, and 3 TψC stems of the tRNA secondary structures in *C. tetraspilus* (Fig 5). Mismatched and wobble pairs are also detected in other Coreoidea species (S5 Fig). These mismatches are common phenomenon for invertebrate tRNAs and could be corrected by posttranscriptional RNA editing processes [49, 50].

Like other insect mitogenomes, the two genes encoding the large and small rRNA subunits (*rrnL* and *rrnS*) in *C. tetraspilus* are located at the conserved positions between *trnL1* (CUN) and *trnV*, and between *trnV* and the control region, respectively (Fig 1, Table 2). The ends of rRNA genes are difficult to be precisely determined by DNA sequencing alone, so they are assumed to extend to the boundaries of flanking genes [51, 52]. The *rrnL* is 1,266 bp long with an A+T content of 78.20%, and the *rrnS* is 786 bp long with an A+T content of 76.84%. The lengths and nucleotide compositions of two rRNA genes in *C. tetraspilus* are similar to that of other sequenced Coreoidea species (S3 Fig).

The secondary structures of the two rRNA genes inferred for *C. tetraspilus* have similar stem-loop structures as those proposed for *Drosophila melanogaster* [24], *Apis mellifera* [25], *Manduca sexta* [26] and other hemipterans (e.g. *Chauliops fallax* [37], *Stenopirates* sp. [39] and *Cavariella salicicola* [48]). The secondary structure of *rrnL* consists of six structural domains (domain III is absent in arthropods) and 45 helices (Fig 6), whereas the secondary structure of *rrnS* contains three domains and 26 helices (Fig 7). In *rrnL*, domains IV and V are more conserved than domains I, II, and VI among sequenced Coreoidea species. Four helices (H563, H1775, H2064, H2507) of *rrnL* are most conserved with completely identical nucleotides among Coreoidea. Some helices (H183, H687, H736, H837, H991, H2077 and H2520) are greatly variable in both sequence and secondary structure among Coreoidea, as frequently observed in other insects [37, 39, 48], and their structures are inferred by the Mfold Web Server [27]. Compared to the 5'-end, the 3'-end of *rrnS* structure is more conserved among Coreoidea, especially for the helices H921-960, H1047 and H1399. The helix H47 are highly variable
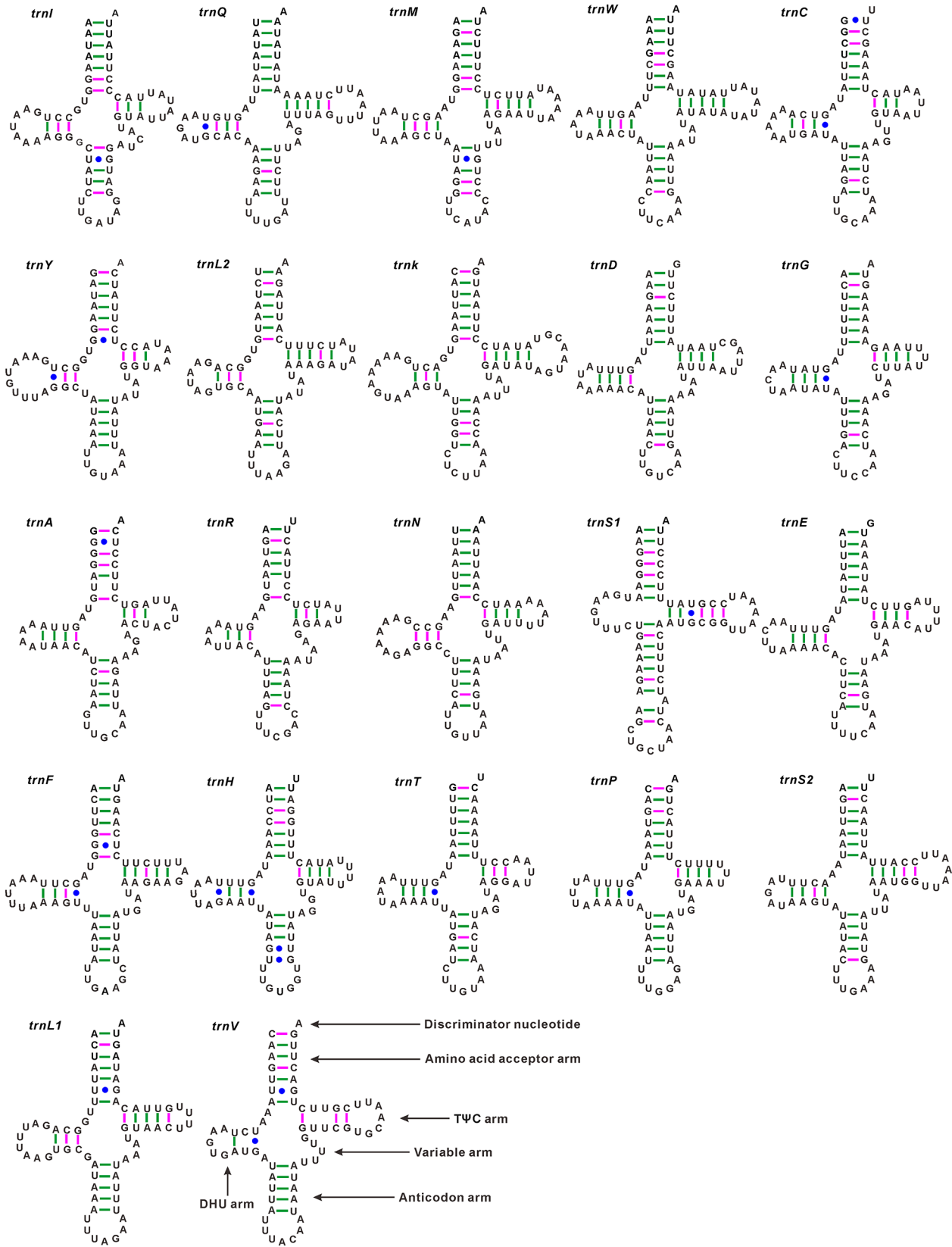
Fig 5. Putative secondary structures of the 22 tRNA genes identified in the mitochondrial genome of *Corizus tetraspilus*. All tRNA genes are shown in the order of occurrence in the mitochondrial genome starting from *trnI*. Bars indicate Watson–Crick base pairings, and dots between G and U pairs mark canonical base pairings in tRNA.

doi:10.1371/journal.pone.0129003.g005

among different insects, and no consistent structure has been found for this region [26]. In *C. tetraspilus*, the possible secondary structure of this region, predicted by the Mfold Web Server [27], consists of a long stem and a short terminal loop, which is similar to that in *Stenopirates* sp. [39] and *Chauliops fallax* [37]. The helices (H1047, H1068, H1074 and H1113) are highly variable, and may yield multiple possible secondary structures due to its high A+T bias and several non-canonical base pairs as observed in other insects [25, 26, 37, 39, 48]. However, the helix H1047 is highly conserved in both sequence and structure among Coreoidea. The helix H1068 has been found in some insects [25, 26, 37, 53], but this helix seems not to be present in the *rrnS* of *C. tetraspilus*, which is similar to those in *Stenopirates* sp. [39] and *Cavariella salicicola* [48].
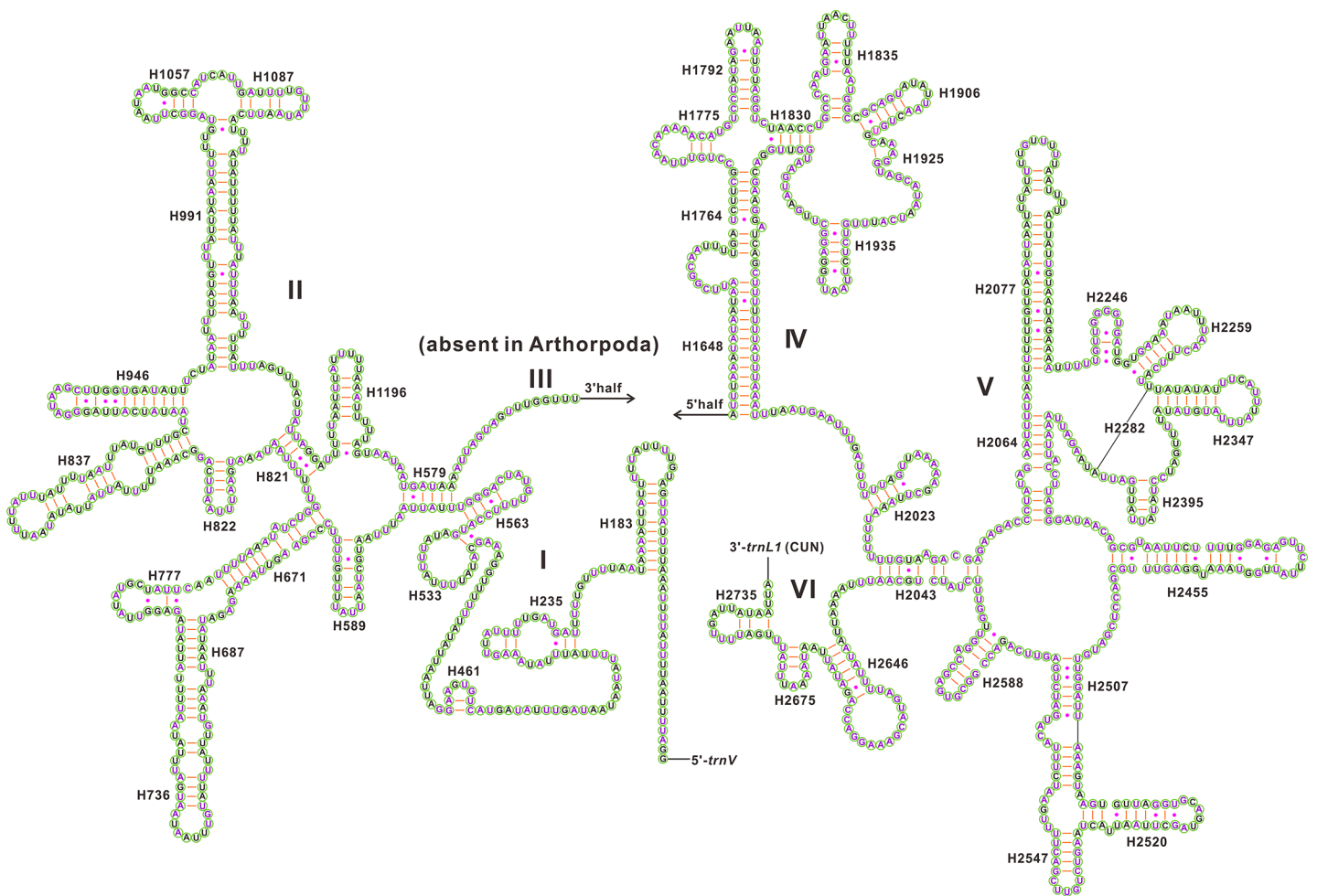
Fig 6. Predicted secondary structure for the *rrnL* in the mitochondrial genome of *Corizus tetraspilus*. The nucleotides showing 100% identity among sequenced Coreoidea species are marker with purple color. Inferred Watson-Crick bonds are illustrated by lines, whereas GU bonds are illustrated by dots.

doi:10.1371/journal.pone.0129003.g006
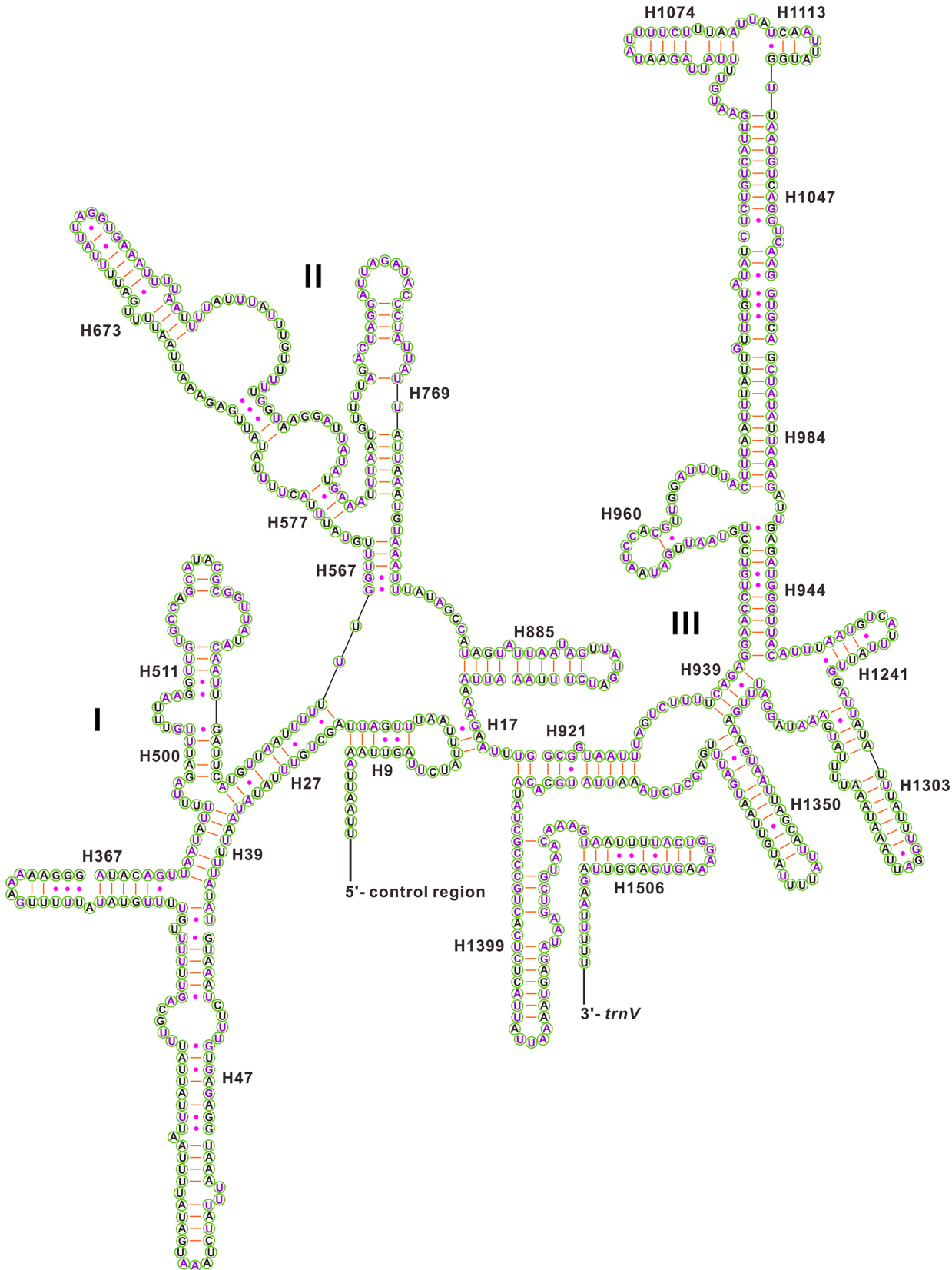
Fig 7. Predicted secondary structure for the *rrnS* in the mitochondrial genome of *Corizus tetraspilus*. The nucleotides showing 100% identity among sequenced Coreoidea species are marker with purple color. Inferred Watson-Crick bonds are illustrated by lines, whereas GU bonds are illustrated by dots.

doi:10.1371/journal.pone.0129003.g007

Non-Coding Regions

The largest non-coding region (440 bp) in the *C. tetraspilus* mitogenome is flanked by *rrnS* and *trnI–trnQ–trnM* gene cluster (Fig 1, Table 2), and can be identified as the putative mitochondrial control region based on the conserved position compared to other insect mitogenomes. The A+T content of this region is 60.68%, which is much lower than that of the entire mitogenome, and is lowest among all the four sequenced mitochondrial control regions (S3 Fig). Although the insect mitochondrial control region is typically characterized by high A+T content, low A+T content in this region has been found in many heteropterans [13]. Furthermore, the control region of *C. tetraspilus* harbors more Ts than As (AT-skew = –0.12), which is opposite to that of other Coreoidea species (S3 Fig).

The length of control regions in the four completely sequenced Coreoidea mitogenomes is highly variable, ranging from 440 bp in *C. tetraspilus* to 1,991 bp in *H. longirostris*. Generally, the putative control regions of the arthropods have any or all of these four motifs: a long sequence of thymines, tandem repeats, a subregion of even higher A+T content, and stem-loop structures [54]. However, neither tandem repeats nor long T-stretches are present in Coreoidea control regions, with the exception of *Riptortus pedestris*. Although the four control regions could form several stem-loop structures, no conserved block has been found, making it difficult to identify any putatively functional motifs. No typical subregions with higher A+T content is present in the control region of *C. tetraspilus*, but a GC-rich region (G+C% = 76.19%) has been found at the 5'-end of the control region. A similar GC-rich region is also present in three other Coreoidea species, with G+C content ranging from 54.83% in *R. pedestris* to 82.60% in *S. subviridis*.

In addition to the putative control region, 31 nucleotides are dispersed in eight intergenic spacers, ranging in size from 1 to 18 bp (Fig 1, Table 2). The majority of intergenic spacer sequences are short (1–3 bp). The longest intergenic spacer sequence (18 bp) is located between *trnS2* (UCN) and *nad1* (Table 2). This intergenic spacer is also detected in other Coreoidea species. Similar non-coding sequences are present at this position in other insect orders [26], and these sequences have been shown to be binding site of a transcription termination factor (DmTTF) [55]. All of the sequences observed in the Coreoidea mitogenomes are highly conserved, and have a sequence of identical length (7 bp) and with significant similarity to the DmTTF binding site (S6 Fig).

In *Stictopleurus subviridis* and *R. pedestris*, another large non-coding region is found between *trnI* and *trnQ* [13, 56]. However, this is not true for *C. tetraspilus* and *H. longirostris*, where *trnQ* overlaps 3 nucleotides with *trnI* on the opposite strand, as found in most hemipteran mitogenomes [47, 48, 57, 58].

Phylogenetic Relationships

Phylogenetic analyses based on the two datasets (PCG and PCGRNA) using two methods (BI and ML) result in almost identical tree topology (Fig 8, S7–S9 Figs). Nodal supports are generally higher in BI tree than those in ML tree generated from the same dataset, as has been revealed by previous studies [59–61]. The only topological incongruence between BI and ML trees based on sequences of 13 PCGs is the phylogenetic relationship among three species within the family Pentatomidae. In BI tree, *Nezara viridula* has a closer relationship with *Halyomorpha halys* with high supports (PP = 0.99, Fig 8), whereas a sister-species relationship between *N. viridula* and *Dolycoris baccarum* is recovered in ML tree with relatively low support (BS = 67, S7 Fig). Phylogenetic analyses using the PCGRNA dataset reduce support values in some nodes and the monophyly of the family Malcidae is not recovered in both BI and ML trees (S8 and S9 Figs), suggesting that RNA data might be unsuitable for reconstructing the evolutionary relationships within Pentatomomorpha.
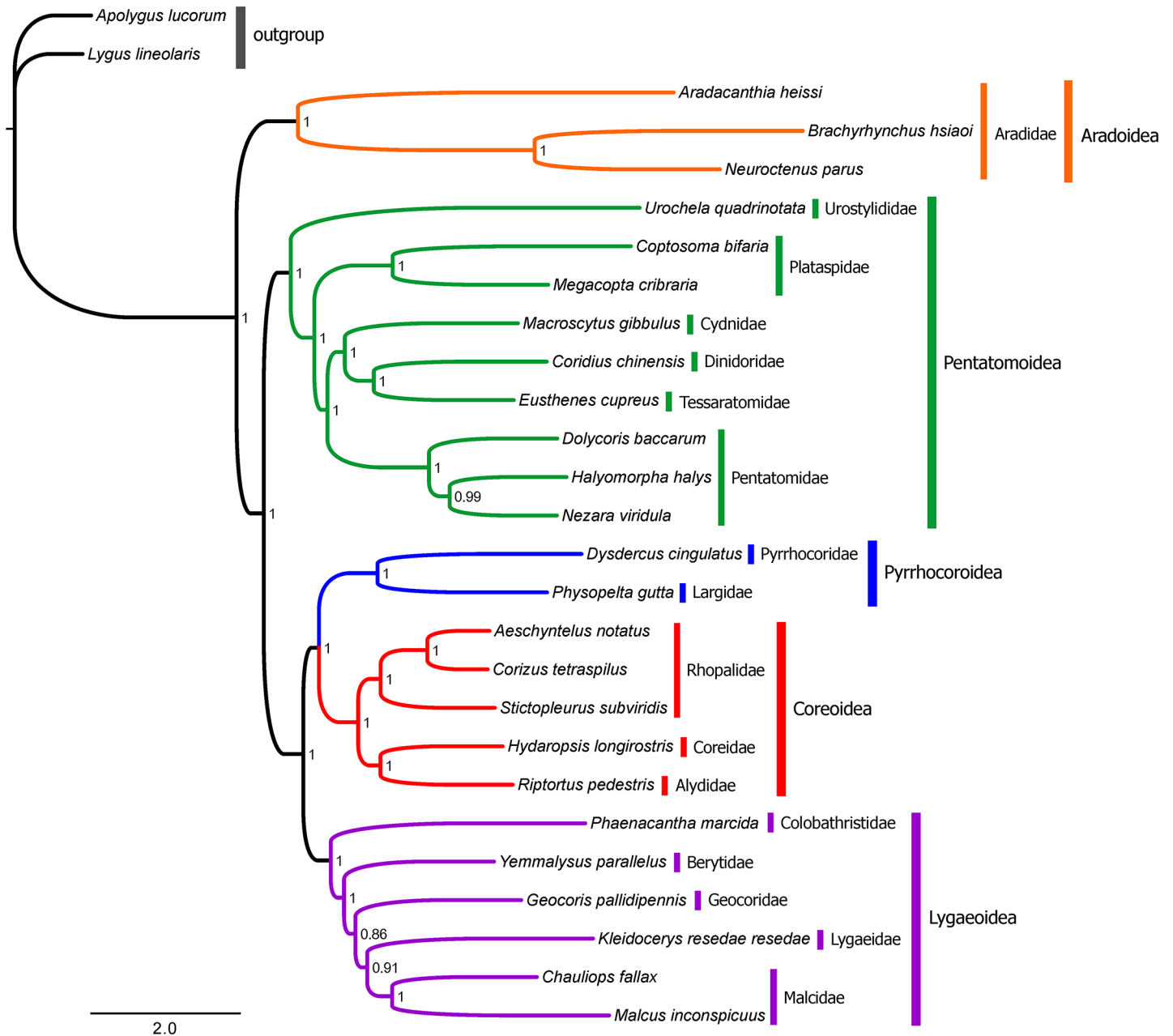
Fig 8. Bayesian phylogenetic relationships among five Pentatomomorpha superfamilies. Phylogenetic analysis is based on the concatenated nucleotide sequences of 13 mitochondrial protein-coding genes. Numbers on branches are Bayesian posterior probabilities.

Our results consistently recover all the superfamilies (Aradoidea, Pentatomoidea, Pyrrhocoroidea, Lygaeoidea, and Coreoidea) established previously in Pentatomomorpha as monophyletic groups with high supports (PP = 1.0, BS = 89–100; Fig 8, S7–S9 Figs). Our results also confirm the hypothesis that Aradoidea and the Trichophora are sister groups, as indicated in previous analyses based on the morphological and molecular data [13–17]. Furthermore, the sister-group relationship of Pentatomoidea and the remainder of the Trichophora is also recognized, which is congruent with previous studies [13, 14, 16, 17].

Incongruent phylogenetic relationships within Eutrichophora have been frequently observed in previous molecular studies [13, 16, 37, 62, 63]. In Eutrichophora, our study recognizes a phylogeny of (Lygaeoidea + (Pyrrhocoroidea + Coreoidea)) consistently supported by both BI and ML analyses (PP = 0.98–1.0, BS = 52–100; Fig 8, S7–S9 Figs), which is consistent with traditional taxonomic hypotheses based on morphology [17] and molecular phylogenetic studies [16]. Especially, this relationship is also recognized by previous studies base on mitogenome data [37, 62]. However, our results are different from that of [13] where the sister-relationship between Lygaeoidea and Coreoidea was revealed based on mitogenomic data. This conflict relationship within Eutrichophora may be due to different taxa sampling and analytical methods. A total of 13 taxa from Pentatomomorpha were used in [13], while 25 species from Pentatomomorpha are included in the present study. The number of species included in Eutrichophora has increased from seven species analyzed by [13] to 13 species used in our study. In [13] all alignments were performed with Clustal W, whereas in the present study rRNA genes are aligned with MAFFT (Q-INS-i method) which has been shown to be more accurate than other programs due to considering the secondary structures of rRNA [29]. For resulting alignment of each gene, poorly aligned positions and divergent regions are removed using GBlocks in our study, but not removed in [13]. In addition, we use PartitionFinder to find both the best partitioning strategy and models of substitution for each partition in Bayesian and ML analyses, whereas in [13] phylogenetic analyses were conducted with a GTR+I+G model without data partitions. The partitioning strategy might optimize the information from the genes and codon positions, which markedly improves phylogenetic resolution in recent studies [60, 64]. Although the present study based on the limited taxa is difficult to well infer the family level relationships within each superfamilies, it still has important implications for the usefulness of mitogenome sequences in evolutionary and phylogenetic studies of Pentatomomorpha.

Supporting Information

S1 Fig. Substitution saturation of 13 protein-coding genes (PCGs), 2 rRNA genes (*rrnL* and *rrnS*) and 22 RNA genes. Transitions and transversions plotted against the F84 distance. (A) first codon positions of 13 PCGs; (B) second codon positions of 13 PCGs; (C) third codon positions of 13 PCGs; (D) all sites of 13 PCGs; (E) all sites of *rrnL*; (F) all sites of *rrnS*; and (G) all sites of tRNAs.
(PDF)

S2 Fig. Alignment of the three longest gene overlaps among the mitochondrial genomes of five Coreoidea species.
(TIF)

S3 Fig. Nucleotide composition of mitochondrial genomes of five Coreoidea species.
(TIF)

S4 Fig. Evaluation of codon bias in the mitochondrial genomes of five Coreoidea species. Species are abbreviated as following: AN, *Aeschyntelus notatus*; CT, *Corizus tetraspilus*; HL, *Hydaropsis longirostris*; RP, *Riptortus pedestris*; SS, *Stictopleurus subviridis*.
(TIF)

S5 Fig. Alignment of the 22 mitochondrial tRNA genes in five Coreoidea species. See S4 Fig for the full names of species.
(TIF)

S6 Fig. Sequence alignments of non-coding region (between *trnS2* and *nad1*) between five Coreoidea species and *Drosophila melanogaster*.
(TIF)

S7 Fig. Maximum likelihood tree among five Pentatomomorpha superfamilies inferred from the concatenated nucleotide sequences of 13 mitochondrial protein-coding genes. Numbers on branches are bootstrap support values.
(PDF)

S8 Fig. Bayesian phylogenetic tree among five Pentatomomorpha superfamilies inferred from the concatenated nucleotide sequences of 13 mitochondrial protein-coding genes and 24 RNA genes. Numbers on branches are Bayesian posterior probabilities.
(PDF)

S9 Fig. Maximum likelihood tree among five Pentatomomorpha superfamilies inferred from the concatenated nucleotide sequences of 13 mitochondrial protein-coding genes and 24 RNA genes. Numbers on branches are bootstrap support values.
(PDF)

S1 Table. Primers used in this study.
(DOCX)

S2 Table. The best partitioning scheme selected by PartitionFinder for the concatenated nucleotide sequences of 13 protein-coding genes.
(DOCX)

S3 Table. Start and stop codons of mitochondrial protein-coding genes of Coreoidea.
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: MLY. Performed the experiments: MLY QLZ ZLG JW. Analyzed the data: MLY QLZ ZLG. Contributed reagents/materials/analysis tools: MLY YYS. Wrote the paper: MLY QLZ. Collected the samples: MLY QLZ.

References

1. Boore JL. Animal mitochondrial genomes. Nucleic Acids Res. 1999; 27: 1767–1780. doi: 10.1093/nar/27.8.1767 PMID: 10101183.

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

3. Wolstenholme DR. Animal mitochondrial DNA: structure and evolution. Int Rev Cytol. 1992; 141: 173–216. doi: 10.1016/S0074-7696(08)62066-5 PMID: 1452431.

4. Zhang DX, Hewitt GM. Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. Biochem Syst Ecol. 1997; 25: 99–120. doi: 10.1016/S0305-1978(96)00042-7

5. Boore JL. The use of genome-level characters for phylogenetic reconstruction. Trends Ecol Evol. 2006; 21: 439–446. doi: 10.1016/j.tree.2006.05.009 PMID: 16762445.

6. Yuan ML, Wei DD, Wang BJ, Dou W, Wang JJ. The complete mitochondrial genome of the citrus red mite *Panonychus citri* (Acari: Tetranychidae): high genome rearrangement and extremely truncated tRNAs. BMC Genomics. 2010; 11: 597. doi: 10.1186/1471-2164-11-597 PMID: 20969792

7. Masta SE. Mitochondrial rRNA secondary structures and genome arrangements distinguish chelicerates: comparisons with a harvestman (Arachnida: Opiliones: *Phalangium opilio*). Gene. 2010; 449: 9–21. doi: 10.1016/j.gene.2009.09.009 PMID: 19800399.

8. Dowton M, Castro LR, Austin AD. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. Invertebr Syst. 2002; 16: 345–356. doi: 10.1071/IS02003

9. Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT. 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. 2006; 37: 545–579. doi: 10.1146/annurev.ecolsys.37.091305.110018

10. Avise JC. Phylogeography: retrospect and prospect. Journal of Biogeography. 2009; 36: 3–15. doi: 10.1111/j.1365-2699.2008.02032.x

11. Wang IJ. Recognizing the temporal distinctions between landscape genetics and phylogeography. Mol Ecol. 2010; 19: 2605–2608. doi: 10.1111/j.1365-294X.2010.04715.x PMID: 20561197.

12. Weirauch C, Schuh RT. Systematics and evolution of Heteroptera: 25 years of progress. Annu Rev Entomol. 2011; 56: 487–510. doi: 10.1146/annurev-ento-120709-144833 PMID: 20822450.

13. Hua JM, Li M, Dong PZ, Cui Y, Xie Q, Bu WJ. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). BMC Genomics. 2008; 9: 610. doi: 10.1186/1471-2164-9-610 PMID: 19091056.

14. Xie Q, Bu WJ, Zheng LY. The Bayesian phylogenetic analysis of the 18S rRNA sequences from the main lineages of Trichophora (Insecta: Heteroptera: Pentatomomorpha). Mol Phylogenet Evol. 2005; 34: 448–451. doi: 10.1016/j.ympev.2004.10.015 PMID: 15619455.

15. Li HM, Deng RQ, Wang JW, Chen ZY, Jia FL, Wang XZ. A preliminary phylogeny of the Pentatomomorpha (Hemiptera: Heteroptera) based on nuclear 18S rDNA and mitochondrial DNA sequences. Mol Phylogenet Evol. 2005; 37: 313–326. doi: 10.1016/j.ympev.2005.07.013 PMID: 16137895.

16. Tian XX, Xie Q, Li M, Gao CQ, Cui Y, Xi L, et al. Phylogeny of pentatomomorphan bugs (Hemiptera-Heteroptera: Pentatomomorpha) based on six *Hox* gene fragments. Zootaxa. 2011; 2888: 57–68.

17. Henry TJ. Phylogenetic analysis of family groups within the infraorder Pentatomomorpha (Hemiptera: Heteroptera), with emphasis on the Lygaeoidea. Ann Entomol Soc Am. 1997; 90: 275–301. doi: http://dx.doi.org/10.1093/aesa/90.3.275

18. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Sym Ser. 1999; 41: 95–98.

19. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30: 2725–9. doi: 10.1093/molbev/mst197 PMID: 24132122.

20. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997; 25: 0955–964. PMID: 9023104.

21. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25: 1451–1452. doi: 10.1093/bioinformatics/btp187 PMID: 19346325.

22. Perna NT, Kocher TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol. 1995; 41: 353–358. PMID: 7563121.

23. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999; 27: 573. PMID: 9862982.

24. Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, et al. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics. 2002; 3: 2. doi: 10.1186/1471-2105-3-2 PMID: 11869452.

25. Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. Insect Mol Biol. 2006; 15: 657–686. doi: 10.1111/j.1365-2583.2006.00689.x PMID: 17069639.

26. Cameron SL, Whiting MF. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. Gene. 2008; 408: 112–23. doi: 10.1016/j.gene.2007.10.023 PMID: 18065166.

27. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003; 31: 3406–3415. doi: 10.1093/nar/gkg595 PMID: 12824337.

28. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 2010; 38: W7–13. doi: 10.1093/nar/gkq291 PMID: 20435676.

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

30. Xia X. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Mol Biol Evol. 2013; 30: 1720–1728. doi: 10.1093/molbev/mst064 PMID: 23564938.

31. Lanfear R, Calcott B, Ho SY, Guindon S. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 2012; 29: 1695–701. doi: 10.1093/molbev/mss020 PMID: 22319168.

32. Miller MA, Pfeiffer W, Schwartz T, editors. Creating the CIPRES Science Gateway for inference of large phylogenetic trees2010 14–14 Nov. 2010: Gateway Computing Environments Workshop (GCE), 2010.

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

34. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012; 61: 539–42. doi: 10.1093/sysbio/sys029 PMID: 22357727.

35. Hershberg R, Petrov DA. Selection on Codon Bias. Annual Review of Genetics. 2008; 42: 287–299. doi: 10.1146/annurev.genet.42.110807.091442 PMID: 18983258.

36. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet. 2011; 12: 32–42. doi: 10.1038/nrg2899 PMID: 21102527

37. Li T, Gao C, Cui Y, Xie Q, Bu W. The complete mitochondrial genome of the stalk-eyed bug *Chauliops fallax* Scott, and the monophyly of Malcidae (Hemiptera: Heteroptera). PLoS ONE. 2013; 8: e55381. doi: 10.1371/journal.pone.0055381 PMID: 23390534.

38. Li H, Liu HY, Song F, Shi AM, Zhou XG, Cai WZ. Comparative mitogenomic analysis of damsel bugs representing three tribes in the family Nabidae (Insecta: Hemiptera). PLoS ONE. 2012; 7: e45925. doi: 10.1371/journal.pone.0045925 PMID: 23029320.

39. Li H, Liu H, Shi AM, Stys P, Zhou XG, Cai WZ. The complete mitochondrial genome and novel gene arrangement of the unique-headed bug *Stenopirates* sp. (Hemiptera: Enicocephalidae). PLoS ONE. 2012; 7: e29419. doi: 10.1371/journal.pone.0029419 PMID: 22235294.

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

41. Lavrov DV, Boore JL, Brown WM. Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: Duplication and nonrandom loss. Mol Biol Evol. 2002; 19: 163–169. PMID: 11801744.

42. Raupach MJ, Hendrich L, Küchler SM, Deister F, Morinière J, Gossner MM. Building-up of a DNA barcode library for true bugs (insecta: hemiptera: heteroptera) of Germany reveals taxonomic uncertainties and surprises. PLoS ONE. 2014; 9: e106940. doi: 10.1371/journal.pone.0106940 PMID: 25203616.

43. Park D-S, Foottit R, Maw E, Hebert PDN. Barcoding Bugs: DNA-Based Identification of the True Bugs (Insecta: Hemiptera: Heteroptera). PLoS ONE. 2011; 6: e18749. doi: 10.1371/journal.pone.0018749 PMID: 21526211.

44. Jung S, Duwal RK, Lee S. COI barcoding of true bugs (Insecta, Heteroptera). Molecular Ecology Resources. 2011; 11: 266–270. doi: 10.1111/j.1755-0998.2010.02945.x PMID: 21429132.

45. Wang Y, Li H, Wang P, Song F, Cai WZ. Comparative mitogenomics of plant bugs (Hemiptera: Miridae): identifying the AGG codon reassignments between Serine and Lysine. PLoS ONE. 2014; 9: e101375. doi: 10.1371/journal.pone.0101375 PMID: 24988409.

46. Wang P, Li H, Wang Y, Zhang JH, Dai X, Chang J, et al. The mitochondrial genome of the plant bug *Apolygus lucorum* (Hemiptera: Miridae): presently known as the smallest in Heteroptera. Insect Sci. 2014; 21: 159–73. doi: 10.1111/1744-7917.12029 PMID: 23956187.

47. Zhang QL, Yuan ML, Shen YY. The complete mitochondrial genome of *Dolycoris baccarum* (Insecta: Hemiptera: Pentatomidae). Mitochondrial DNA. 2013; 24: 469–71. doi: 10.3109/19401736.2013.766182 PMID: 23391217.

48. Wang Y, Huang XL, Qiao GX. Comparative analysis of mitochondrial genomes of five aphid species (Hemiptera: aphididae) and phylogenetic implications. PLoS ONE. 2013; 8: e77511. doi: 10.1371/journal.pone.0077511 PMID: 24147014.

49. Lavrov DV, Brown WM, Boore JL. A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. Proc Natl Acad Sci. 2000; 97: 13738–13742. doi: 10.1073/pnas.250402997 PMID: 11095730.

50. Masta SE, Boore JL. The complete mitochondrial genome sequence of the spider *Habronattus oregonensis* reveals rearranged and extremely truncated tRNAs. Mol Biol Evol. 2004; 21: 893. doi: 10.1093/molbev/msh096 PMID: 15014167.

51. Boore JL. Complete mitochondrial genome sequence of the polychaete annelid *Platynereis dumerilii*. Mol Biol Evol. 2001; 18: 1413–1416. PMID: 11420379.

52. Boore JL. The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda). BMC Genomics. 2006; 7: 182. doi: 10.1186/1471-2164-7-182 PMID: 16854241.

53. Page RDM. Comparative analysis of secondary structure of insect mitochondrial small subunit ribosomal RNA using maximum weighted matching. Nucleic Acids Res. 2000; 28: 3839–3845. doi: 10.1093/nar/28.20.3839 PMID: 11024161

54. Cook C. The complete mitochondrial genome of the stomatopod crustacean Squilla mantis. BMC Genomics. 2005; 6: 105. doi: 10.1186/1471-2164-6-105 PMID: 16091132.

55. Roberti M, Polosa PL, Bruni F, Musicco C, Gadaleta MN, Cantatore P. DmTTF, a novel mitochondrial transcription termination factor that recognises two sequences of *Drosophila melanogaster* mitochondrial DNA. Nucleic Acids Res. 2003; 31: 1597–1604. doi: 10.1093/Nar/Gkg272 PMID: 12626700

56. Hua JM, Dong PZ, Li M, Cui Y, Zhu WB, Xie Q, et al. The analysis of mitochondrial genome of *Stictopleurus subviridis* Hsiao (Insecta: Hemiptera-Heteroptera: Rhopalidae). Acta Zootax Sin. 2009; 34: 1–9.

57. Li H, Liu HY, Cao LM, Shi AM, Yang HL, Cai WZ. The complete mitochondrial genome of the damsel bug *Alloeorhynchus bakeri* (Hemiptera: Nabidae). Int J Biol Sci. 2012; 8: 93–107. doi: 10.7150/ijbs.8.93 PMID: 22211108.

58. Zhang QL, Guo ZL, Yuan ML. The complete mitochondrial genome of *Poratrioza sinica* (Insecta: Hemiptera: Psyllidae). Mitochondrial DNA. 2015: In press. doi: 10.3109/19401736.2014.913164

59. Mao M, Gibson T, Dowton M. Evolutionary dynamics of the mitochondrial genome in the Evaniomorpha (Hymenoptera)—A group with an intermediate rate of gene rearrangement. Genome Biol Evol. 2014; 6: 1862–1874. doi: 10.1093/gbe/evu145 PMID: 25115010

60. Wei SJ, Li Q, van Achterberg K, Chen XX. Two mitochondrial genomes from the families Bethylidae and Mutillidae: independent rearrangement of protein-coding genes and higher-level phylogeny of the Hymenoptera. Mol Phylogenet Evol. 2014; 77: 1–10. doi: 10.1016/j.ympev.2014.03.023 PMID: 24704304.

61. Cameron SL, Lo N, Bourguignon T, Svenson GJ, Evans TA. A mitochondrial genome phylogeny of termites (Blattodea: Termitoidae): Robust support for interfamilial relationships and molecular synapomorphies define major clades. Mol Phylogenet Evol. 2012; 65: 163–173. doi: 10.1016/j.ympev.2012.05.034 PMID: 22683563.

62. Kocher A, Kamilari M, Lhuillier E, Coissac E, Péneau J, Chave J, et al. Shotgun assembly of the assassin bug *Brontostoma colossus* mitochondrial genome (Heteroptera, Reduviidae). Gene. 2014; 552: 184–194. doi: 10.1016/j.gene.2014.09.033 PMID: 25240790.

63. Song N, Liang AP, Bu CP. A molecular phylogeny of Hemiptera inferred from mitochondrial genome sequences. PLoS ONE. 2012; 7: e48778. doi: 10.1371/journal.pone.0048778 PMID: 23144967.

64. Leavitt JR, Hiatt KD, Whiting MF, Song H. Searching for the optimal data partitioning strategy in mitochondrial phylogenomics: A phylogeny of Acridoidea (Insecta: Orthoptera: Caelifera) as a case study. Mol Phylogenet Evol. 2013; 67: 494–508. doi: 10.1016/j.ympev.2013.02.019 PMID: 23454468.

65. Wang Y, Li H, Xun H, Cai W. Complete mitochondrial genome sequence of the plant bug *Adelphocoris fasciaticollis* (Hemiptera: Heteroptera: Miridae). Mitochondrial DNA. 2015: In press. doi: 10.3109/19401736.2014.880898 PMID: 24495137.

66. Shi AM, Li H, Bai XS, Dai X, Chang J, Guilbert E, et al. The complete mitochondrial genome of the flat bug *Aradacanthia heissi* (Hemiptera: Aradidae). Zootaxa. 2012; 3238: 23–38.

67. Li H, Shi A, Song F, Cai W. Complete mitochondrial genome of the flat bug *Brachyrhynchus hsiaoi* (Hemiptera: Aradidae). Mitochondrial DNA. 2015: In press. doi: 10.3109/19401736.2013.867437 PMID: 24438289.

68. Li T, Yi W, Zhang H, Xie Q, Bu W. Complete mitochondrial genome of the birch catkin bug *Kleidocerys resedae resedae*, as the first representative from the family Lygaeidae (Hemiptera: Heteroptera: Lygaeoidea). Mitochondrial DNA. 2015: In press. doi: 10.3109/19401736.2014.908372 PMID: 24725058

69. Liu L, Li H, Song F, Song W, Dai X, Chang J, et al. The mitochondrial genome of *Coridius chinensis* (Hemiptera: Dinidoridae). Zootaxa. 2012; 3537: 29–40.

70. Lee W, Kang J, Jung C, Hoelmer K, Lee S. Complete mitochondrial genome of brown marmorated stink bug *Halyomorpha halys* (Hemiptera: Pentatomidae), and phylogenetic relationships of hemipteran suborders. Mol Cells. 2009; 28: 155–165. doi: 10.1007/s10059-009-0125-9 PMID: 19756390.

71. Song W, Li H, Song F, Liu L, Wang P, Xun HZ, et al. The complete mitochondrial genome of a tessaratomid bug, *Eusthenes cupreus* (Hemiptera: Heteroptera: Pentatomomorpha: Tessaratomidae). Zootaxa. 2013; 3620: 260–272. doi: http://dx.doi.org/10.11646/zootaxa.3620.2.4

72. Dai YT, Li H, Jiang P, Song F, Ye Z, Yuan XQ, et al. Sequence and organization of the mitochondrial genome of an urostylidid bug, *Urochela quadrinotata* Reuter (Hemiptera: Urostylididae). Entomotaxonomia. 2012; 34: 613–623.