

Comparative Mitogenomic Analysis of the *Eurydema* Genus in the Context of Representative Pentatomidae (Hemiptera: Heteroptera) Taxa

Wanqing Zhao,¹ Qing Zhao,² Min Li,³ Jiufeng Wei,² Xianhong Zhang,^{2,*} and Hufang Zhang^{1,4,*}

¹Department of Biology, Xinzhou Teachers University, Xinzhou 034000, China, ²Department of Entomology, Shanxi Agricultural University, Taigu 030800, China, ³Department of Biology, Taiyuan Normal University, Taiyuan 030619, China, and ⁴Corresponding author, e-mail: zh_hufang@sohu.com

*These authors contributed equally to this article.

Subject Editor: Margaret Allen

Received 4 September 2019; Editorial decision 25 November 2019

Abstract

The family Pentatomidae, the largest within the superfamily Pentatomoidae, comprises about 5,000 species; many of which are economically important pests. Although the phylogeny of Pentatomidae species has been studied using various molecular markers, their phylogenetic relationships remain controversial. Recently, mitochondrial genomes (mitogenomes) have been extensively employed to examine the phylogenetics and evolution of different insects, and in this study, we sequenced complete/near-complete mitochondrial genomes from five shield bug species of *Eurydema* to gain a better understanding of phylogenetic relationships in the Pentatomidae. The five mitogenomes ranged in length from 15,500 to 16,752 bp and comprised 13 protein-coding genes (PCGs), 22 transfer RNAs (tRNAs), 2 ribosomal RNAs (rRNAs), and a control region. We compared mitogenomic characteristics of the Pentatomidae and constructed phylogenetic trees using Bayesian inference and maximum likelihood methods. Our results showed that gene arrangements, base composition, start/stop codons, gene overlaps, and RNA structures were conserved within the Pentatomidae and that congeneric species shared more characteristics. Saturation and heterogeneity analyses revealed that our PCGs and PCGRNA datasets were valid for phylogenetic analysis. Phylogenetic analyses showed consistent topologies based on BI and ML methods. These analyses strongly supported that *Eurydema* species belong to the tribe Strachiini, and formed a sister group with Pentatomini. The relationships among *Eurydema* species were shown to be consistent with their morphological features. (Strachiini + Pentatomini) was found to be a stable sibling of the clade comprising Cappaeini, Graphosomini, and Carpocorini. Furthermore, our results indicated that *Graphosoma rubrolineatum* (Heteroptera: Pentatomidae) belongs to the Pentatominae and not the Podopinae.

Key words: Pentatomidae, *Eurydema*, mitogenome, phylogenetic relationship

The Pentatomidae, or stink bugs, are an important group of Heteroptera and the largest family within the superfamily Pentatomoidae (Hemiptera: Heteroptera). Worldwide, this family consists of ~5,000 known species belonging to >800 genera and 10 subfamilies (Wang 2010, Zhao et al. 2018). In China, Pentatomidae species belong to four subfamilies: Asopinae, Pentatominae, Phyllocephalinae, and Podopinae (Rider et al. 2002; Rider and Zheng 2002, 2005). In previous studies, high-level relationships within the Heteroptera have been a primary focus, whereas there has been comparatively limited research conducted on the phylogenetic relationships of subfamilies and tribes. To date, few molecular markers or morphological characteristics have been applied

in analyzing the phylogeny of the Pentatomidae, and accordingly, the phylogenetic relationships among subfamilies and tribes remain controversial (Grazia et al. 2008, Yuan et al. 2015). Yang (1962) suggested a classification system of nine tribes, in which *Eurydema* was placed within the Palomenini. However, Rider (2015) subsequently classified *Eurydema* within the Strachiini on the website of Pentatomioidea Home Page (<http://www.ndsu.nodak.edu/ndsu/rider/Pentatomioidea/>). In addition, on the basis of morphological data, some researchers have considered that the genus *Graphosoma* belongs to the tribe Graphosomini of Pentatominae (Hsiao 1977, Liu 2007), whereas others have upgraded it to the Podopinae (Rider 2002, 2015; Xing 2013).

In China, species within the genus *Eurydema* Laporte (Heteroptera: Pentatomidae) are known as agricultural pests of cruciferous plants, with outbreaks occurring from June to August. Whereas some of these *Eurydema* species are widespread across the Palaearctic (e.g., *Eurydema dominulus* and *Eurydema oleracea*), others are endemic (e.g., *Eurydema maracandica* and *Eurydema liturifera*) (Yang 1962, Hsiao 1977, Rider 2006, Henry 2009). Although in general, phenotypic variation within the species of the genus *Eurydema* with respect to color has aided species identification compared with true bugs, it has been of limited utility in distinguishing closely related species. Additionally, multiple intraspecific phenotypes can be found, making precise identification difficult. In order to overcome these taxonomic difficulties, some authors have examined the utility of DNA barcoding and the characteristics of male genitalia (Jung et al. 2011; Akhter et al. 2015; Zhao et al. 2017a,b). However, mitogenome analysis may offer a different perspective with regards to determining the phylogeny and diversification of the genus *Eurydema* and has a high probability of resolving the evolution of this genus (Breeschoten et al. 2016).

Insect mitogenome is double-stranded circular DNA molecules with length of 14–17 kb, and typically contains 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), and 2 ribosomal RNA genes (rRNAs) (Wolstenholme 1992, Franoso et al. 2016, Yang et al. 2018). Mitogenome harbors a number of key molecular markers that can be used to resolve phylogenetic relationships at multiple taxonomic levels and are characterized by a number of informative features, including, length variation, altered RNA secondary structures, rearranged genes, and hyper variable control regions (Downton et al. 2002, Cameron et al. 2010, Colinet et al. 2017, Djoumad et al. 2017). Owing to their relatively conserved gene content and organization, rapid evolutionary rate, high genetic information content, and abundant genome-level features, insect mitogenomes have been widely used in phylogenetic analyses, species identification, and evolution at different taxonomic levels (Ma et al. 2012, Hegedusova et al. 2014, Liu 2017, Wu et al. 2019).

At present, 167 complete or near-complete mitogenomes of heteropteran species are available in GenBank, and there have been numerous studies that have used these mitogenomes (Fischer et al.

2000; Weirauch and Schuh 2010; Li et al. 2012, 2017; Wang et al. 2016). However, only 14 mitogenomes have been sequenced for Pentatomidae to date. Therefore, sequencing additional mitogenomes from species within the Pentatomoida is needed.

Herein, we sequenced the mitogenomes of five *Eurydema* species and compared the mitogenome features of the *Eurydema* with those of the 14 previously sequenced Pentatomidae mitogenomes. We analyzed the A+T content and AT-skew of different partitions, codon usage, and nucleotide substitution of 13 PCGs, and the secondary structure of RNA genes. The applicability of two datasets containing the 13 PCGs and all 37 genes for phylogenetic analysis were evaluated by heterogeneity and substitution saturation analysis. Finally, we examined phylogenetic relationships using Bayesian inference (BI) and maximum likelihood (ML) methods to confirm the main tribes of Pentatominae and species within the genus *Eurydema*.

Materials and Methods

Taxon Sampling and Mitogenome Sequencing

The specimens used in this study were collected in China between 2013 and 2016, and the detailed information listed in Table 1. All samples were preserved in 100% ethanol and stored at –20°C. Total genomic DNA was extracted from leg muscles using a ONE-4-ALL Genomic DNA Mini-Prep Kit (BS88504; Sangon, Shanghai, China). The mitogenomes were sequenced using the whole-genome shotgun method on an Illumina Miseq platform (Personalbio, Shanghai, China). After filtering low-quality and adapter contaminated reads, A5-miseq version 20150522 (Coil et al. 2015) was used for contig assembly.

Sequence Annotation and Analysis

The five newly sequenced mitogenomes were annotated using Geneious 10.1.3 (Kearse et al. 2012) via comparison with the previously published sequence of the *E. maracandica* mitogenome (Zhao et al. 2017a). Annotations of 13 protein-coding regions were edited manually by predicting open reading frames using the invertebrate mitochondrial code. tRNA genes were identified based on

Table 1. Sequence information included in the present study

Species	Common name	Accession number	Locality	Collection date
<i>Aradacanthia heissi</i>	Flat bug	HQ441233	No record in GenBank	No record in GenBank
<i>Urochela quadrinotata</i>	Four-spotted shield bug	NC_020144	No record in GenBank	No record in GenBank
<i>Megacopta cribraria</i>	Bean plataspid	NC_015342	No record in GenBank	No record in GenBank
<i>Coptosoma bifaria</i>	Double-row turtle bug	EU427334	No record in GenBank	No record in GenBank
<i>Macroscytus gibbulus</i>	Horizontal leather burrower bug	NC_012457	No record in GenBank	No record in GenBank
<i>Eusthenes cupreus</i>	Discolor giant bugs	NC_022449	No record in GenBank	No record in GenBank
<i>Coridius chinensis</i>	Melon black bug	JQ739179	No record in GenBank	No record in GenBank
<i>Halyomorpha halys</i>	Brown marmorated stink bug	NC_013272	No record in GenBank	No record in GenBank
<i>Pentatomidae</i> sp.	Stink bug	KM244699	No record in GenBank	No record in GenBank
<i>Graphosoma rubrolineatum</i>	Striped shield bug	NC_033875	No record in GenBank	No record in GenBank
<i>Rubiconia intermedia</i>	Bead bug	KP207596	No record in GenBank	No record in GenBank
<i>Dolycoris baccarum</i>	Sloe bug	KC460537	No record in GenBank	No record in GenBank
<i>Nezara viridula</i>	Southern green stink bug	EF208087	No record in GenBank	No record in GenBank
<i>Eurydema gebleri</i>	Cross grain cabbage bug	NC_027489	No record in GenBank	No record in GenBank
<i>Eurydema maracandica</i>	Xinjiang cabbage bug	MF135553	No record in GenBank	No record in GenBank
<i>Eurydema dominulus</i>	Cabbage bug	MG584833	Linyi, Shandong	2 Aug. 2014
<i>Eurydema ventralis</i>	Border-radius cabbage bug	MG584837	Altay, Xinjiang	1 Aug. 2011
<i>Eurydema qinlingensis</i>	Qinling cabbage bug	MG584836	Ankang, Shaanxi	17 July 2016
<i>Eurydema oleracea</i>	Blue cabbage bug	MG584835	Tarbagatay, Xinjiang	8 Aug. 2016
<i>Eurydema liturifera</i>	India-Burma cabbage bug	MG584834	Pu'er, Yunnan	1 Aug. 2006

their characteristic cloverleaf secondary structure using the MITOS web server (Bernt et al. 2013). rRNA genes were identified by comparing nucleotide sequences with those of previously reported mitogenomes.

Along with the complete mitogenomes in the present study, 15 complete Heteroptera mitogenomes were downloaded from the NCBI GenBank (Table 1). The mitogenome of *Aradacanthia heissi* (Aradidae) was used as an outgroup. The sequences of 13 PCGs, 22 tRNAs and two rRNAs were extracted from each mitogenome. PCGs were aligned according to codon-based multiple alignments using the MUSCLE algorithm in MEGA 7.0 (Kumar et al. 2016). Genes for RNAs were aligned with MUSCLE using default parameters in MEGA 7.0. Alignments of individual genes were concatenated in SequenceMatrix (Vaidya et al. 2011) according to the following two matrices for phylogenetic analyses: i) the PCGs matrix, including all three codon positions of the 13 PCG genes (13PCG); ii) the PCGRNA matrix, including 13 PCG genes, 22 tRNA genes, and 2 rRNA genes. The codon usage of PCGs and nucleotide composition of different regions were analyzed using MEGA 7.0. The AT- and GC-skew values were calculated as follows: AT skew = $(A-T)/(A+T)$ and GC skew = $(G-C)/(G+C)$.

Phylogenetic Analyses

Phylogenetic analyses were conducted using BI. The heterogeneity of sequence divergence within datasets was analyzed using AliGROOVE (Kück et al. 2014). In order to avoid incorrect phylogenetic inferences, Xia's saturation index (*Iss*) was estimated and compared to the critical values (*Iss.c*) using DAMBE (Xia 2018). BI analysis was performed using MrBayes3.2 (Ronquist et al. 2012) under a GTR + G + I model selected by jModeltest (Darriba et al. 2012), in which 1,000,000 generations were run, with four simultaneous Markov chains (one cold chain and three heated chains), and trees were sampled after every 1,000 generations. The first 25% of generations were discarded as burn-in when average standard deviation of split frequencies were >0.01 . The ML trees were constructed using RAxML (Alexandros 2014) under the GTR+GAMMA model, and the node support values were assessed by bootstrap re-sampling (BP) calculated using 1,000 replicates. Phylogenetic informativeness (PI) of data partitions based on the tree constructed using the combined data of PCGs was measured using PhyDesign (López-Giráldez and Townsend 2011).

Results

Genome Organization

We sequenced the complete mitogenomes of *E. dominulus*, *Eurydema ventralis*, *E. oleracea*, and *Eurydema qinlingensis*, and the near-complete mitogenome of *E. liturifera*, which was lacking part of the control region. The mitogenome sequences ranged in size from 15,500 bp (*E. ventralis*) to 16,752 bp (*E. dominulus*), whereas the size of the near-complete mitogenome of *E. liturifera* was 15,585 bp. The failure to amplify part of the control region in *E. liturifera* may be due to its high variability and/or extreme length. The five annotated mitogenomes have been deposited in GenBank with accession numbers as per Table 1. The determined sequences are double-stranded circular molecules that in common with most other heteropterans mitogenomes, encode 37 genes (13 PCGs, 22 tRNAs, and 2 rRNAs). The majority of these genes (23) are encoded by the heavy strand (J-strand), with the remaining 14 genes (*trnQ*, *trnC*, *trnY*, *trnF*, *nad5*, *trnH*, *nad4*, *nad4L*, *trnP*, *nad1*, *trnL1(CUN)*, *trnV*, 16S *rRNA*, and 12S *rRNA*) being located on the light strand (N-strand).

The organizations of the five mitogenomes are shown in Supp file (online only).

Although we found little variation in the length of the PCGs, tRNAs and rRNAs of the 14 sequenced Pentatominae mitogenomes, the lengths of control and intergenic regions showed more pronounced variation. Within the Pentatominae mitogenomes, the total length of PCGs, tRNAs, rRNAs, and the control region ranged from 11,004 to 11,027 bp; 1,459 to 1,506 bp; 2,060 to 2,107 bp; and 543 to 2,190 bp, respectively. The Pentatominae mitogenomes

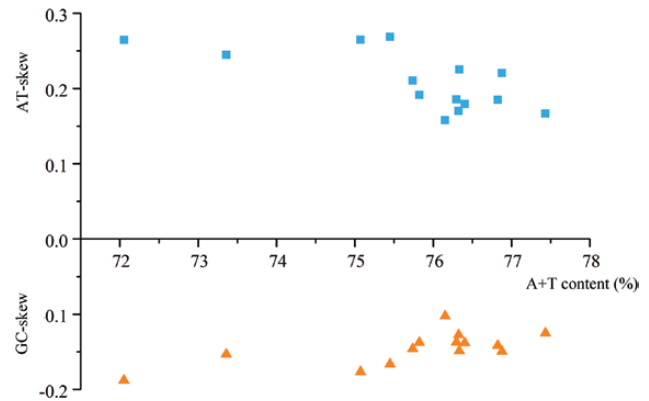


Fig. 1. AT skew and GC-skew of Pentatomidae mitogenomes.

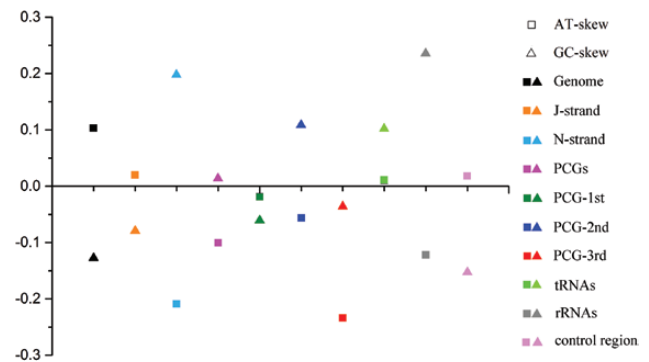


Fig. 2. AT skew and GC skew of different databases.

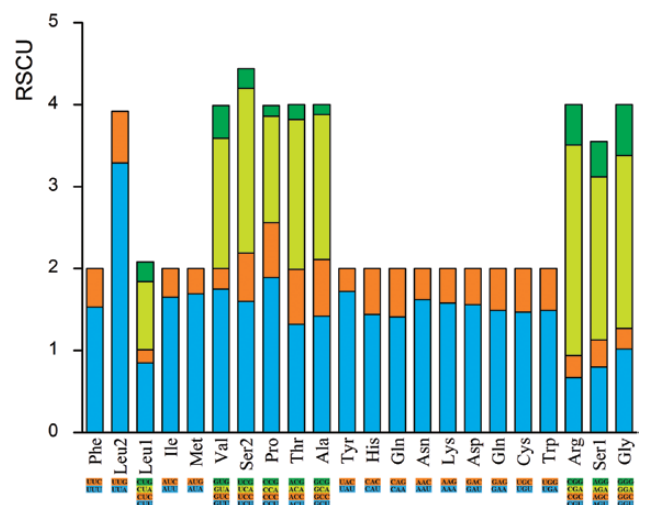


Fig. 3. The codon usage of PCGs in Pentatomidae mitogenomes.

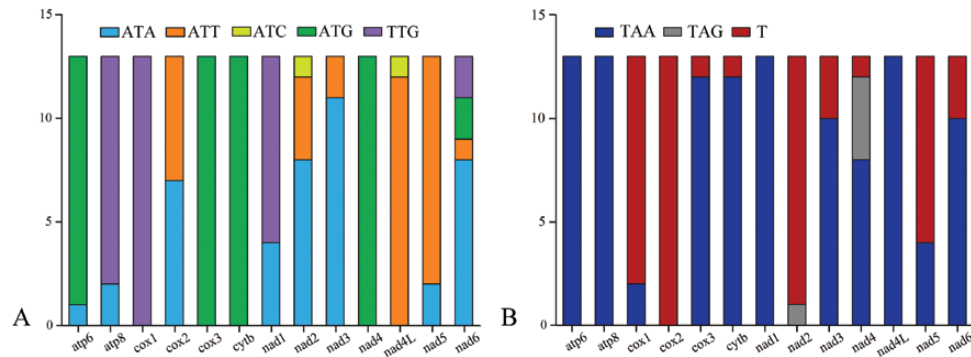


Fig. 4. Start (A) and stop (B) codons usage of Pentatomidae protein-coding genes.

have compact arrangements, and both gene overlaps and intergenic spacers were observed, with the former ranging from 1 to 8 bp with a total length of 22–32 bp. The longest overlap was that of the gene pair *trnW/nad1* (8 bp), and we found that overlaps for the gene pair *trnC/trnW*, *atp8/atp6*, and *nad4/nad4L* are completely conserved in the *Eurydema* mitogenomes. Intergenic spacers ranged in length from 1 to 31 bp, with a total length of 105–167 bp, among which the spacer between the gene pair *trnS2/nad1* (27–29 bp) was found to be the longest intergenic region.

Nucleotide Composition and Codon Usage

Data relating to the nucleotide composition, whole mitogenome AT- and GC-skew values, J-strands, N-strands, PCGs, PCG-1st, PCG-2nd, PCG-3rd, tRNAs, rRNAs, and control regions of the Pentatominae mitogenomes are presented in Figs. 1 and 2. Overall, the A+T content of these mitogenomes is significantly higher than that of G+C content, and strongly skewed toward A and T. For the three codon positions of protein-coding genes, PCG-3rd (79.9%) has the highest content of A+T, and PCG-1st (72.32%) the lowest. The A+T content of the N-strand is slightly higher than that of the J-strand, with an average A+T content of 75.59%. Among the 13 PCGs, *atp8* has the highest A+T content (83.1%), whereas *cox1* has the lowest (70.3%).

For all sequenced Pentatominae mitogenomes, the AT-skew is greater than 0 (ranging from 0.1578 to 0.2645), whereas the GC-skews are all less than 0 (from -0.1877 to -0.1023). Whole sequences show moderate A/C skew, J-strands show slight A/C skew, N-strands show obvious T/G skew, PCGs show moderate A skew and slight C skew, PCGs-1st shows slight T/G skew, PCGs-2nd shows slight T skew and moderate G skew, PCGs-3rd shows obvious T skew and slight C skew, tRNAs show slight Askew and moderate A skew, and rRNAs show moderate T skew and obvious G skew.

The results of statistical analysis of the relative synonymous codon usage (RSCU) in mitogenome PCGs are shown in Fig. 3. The most frequently used codons are UUA (Leu), CGA (Arg), and GGA (Gly), whereas the two most infrequently used codons are GCG (Ala) and CCG (Pro). Synonymous codons ending with an A or U are more prevalent than those ending in a G or C. For instance, UUU (RSCU = 1.53) is more common than UUC (RSCU = 0.47) for Phe, CAU (RSCU = 1.44) is more common than CAC (RSCU = 0.56) for His, and AUU (RSCU = 1.65) is more common than AUC (RSCU = 0.35) for Ile. Although the amino acid compositions of the Pentatominae mitogenome PCGs are similar, we found that their proportions are not the same. Leu was observed to be the most commonly represented amino acid, followed by Phe, Ile, and Met, and

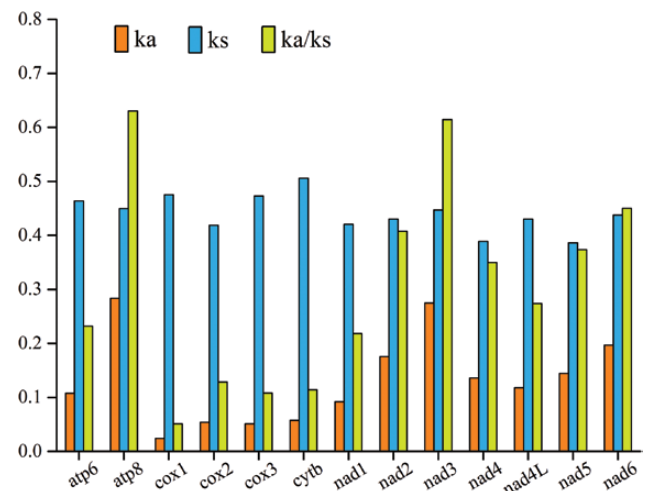


Fig. 5. The Ka, Ks, and Ka/Ks values of protein-coding genes.

therefore, the codons corresponding to these common amino acids also have relatively high proportions.

Protein Coding Genes

Among the nucleotide sequences of the PCGs in Pentatominae mitogenome, we found that *cox1* had the highest conserved sites (66.17%), whereas that for *atp8* (27.78%) is the lowest. In terms of amino acid sequences, the proportion of conserved sites in the *cox1* (86.35%) and *cytb* (74.14%) genes was observed to be higher than those in other genes, whereas the proportion in *nad2* (33.03%) was the lowest. Conservation of PCGs is also reflected in the usage of start and stop codons. The results of statistical analysis of Pentatominae mitogenome start/stop codons are shown in Fig. 4. Although most PCGs were found to initiate with a typical start codon (ATN), *atp8*, *cox1*, and *nad1* invariably begin with TTG. Furthermore, all PCGs ended with complete termination codon TAA/TAG, except *cox1*, *cox2*, *nad3*, and *nad5* ended with truncated stop codon (T). Calculation of the rates of nonsynonymous (Ka) and synonymous (Ks) nucleotide substitutions revealed that the Ka/Ks ratios for all 13 PCGs were considerably <1 (<0.63), thereby indicating evolution under purifying selection (Fig. 5). All mitochondrial PCGs can be used to examine phylogenetic relationships. The highest Ka/Ks value was found in *atp8*, whereas *cox1* appears to have the lowest evolutionary rate. *atp8* can be used to analyze intraspecific relationships, whereas the four genes *cox1*, *cox2*, *cox3*, and *cytb* is useful for DNA barcoding.

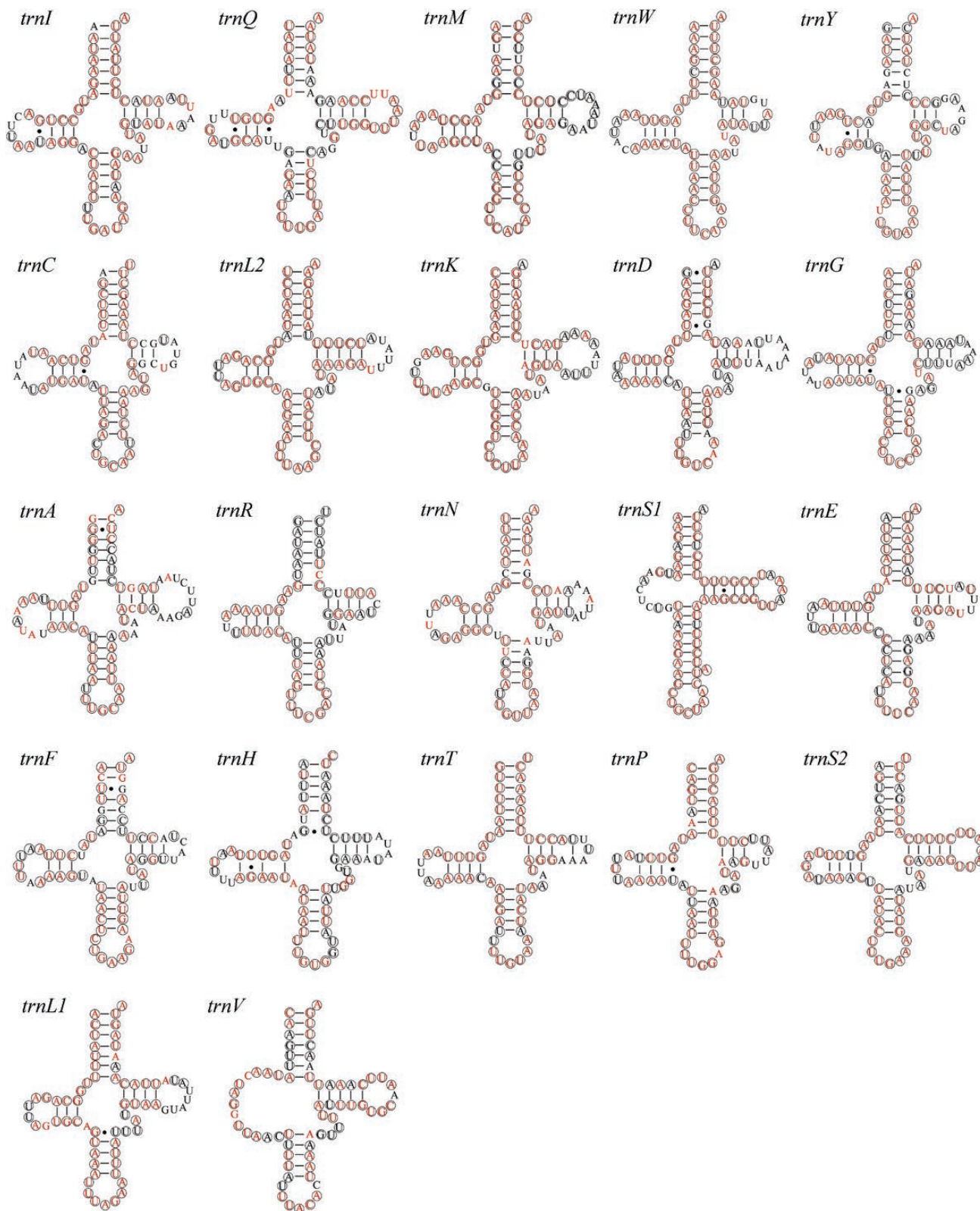


Fig. 6. Potential secondary structure of tRNA in *Eurydema dominulus*. The conserved sites within *Eurydema* were labeled with circles, and conserved sites within Pentatominae were marked with red color.

Transfer RNA and Ribosomal RNA Genes

In common with other Pentatominae mitogenomes, we found that the five sequenced *Eurydema* mitogenomes contain 22 tRNA genes, with lengths ranging from 62 to 72 bp. Among these, 20 were predicted

to fold into the typical cloverleaf secondary structure, whereas the remaining two tRNAs lack the dihydrouridine (DHU) stem and forma loop. The predicted secondary structures of the 22 tRNAs of *E. dominulus* are shown in Fig. 6. The DHU stem, anticodon

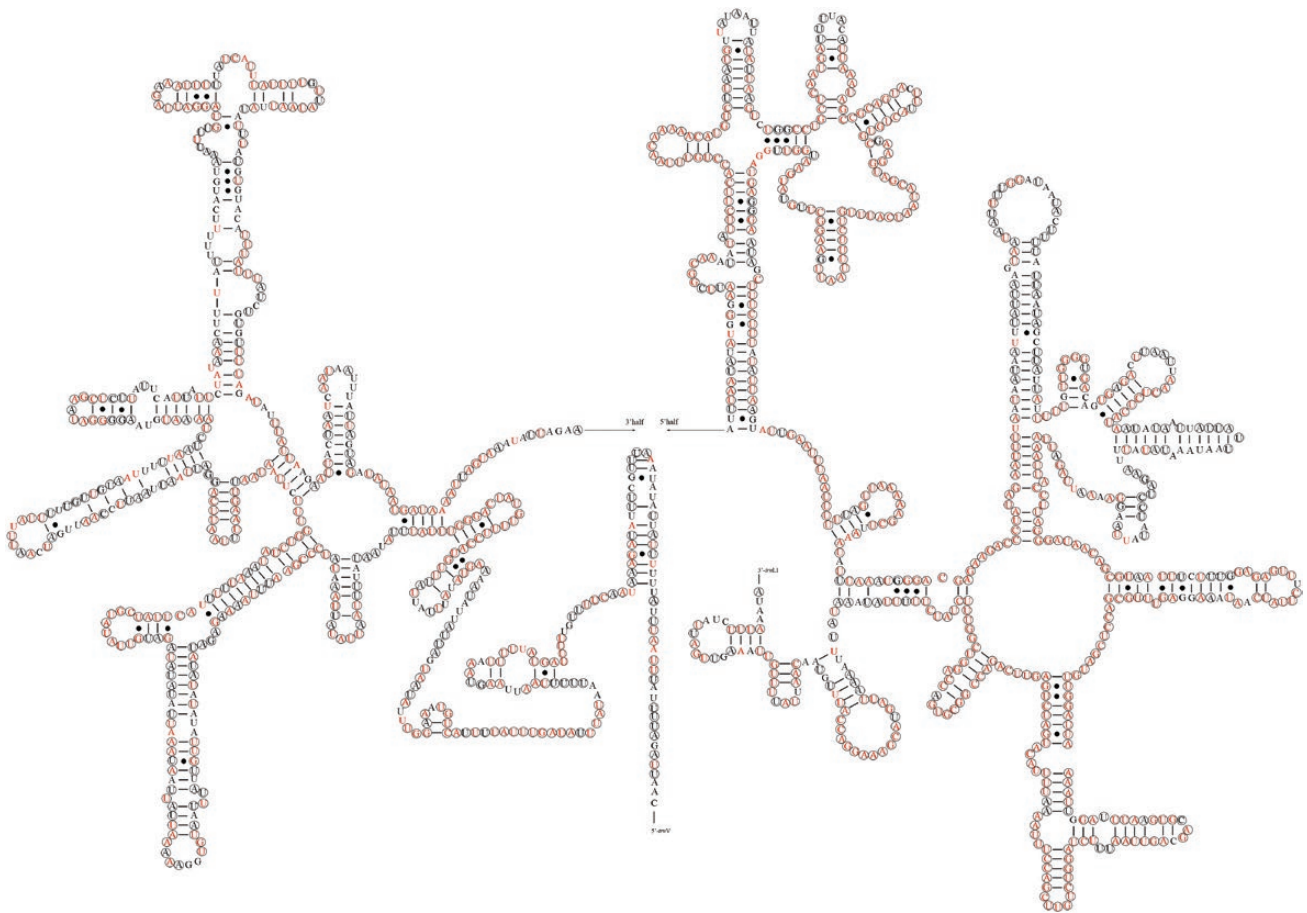


Fig. 7. Potential secondary structure of 16S rRNA in *Eurydema dominulus*. The conserved sites within *Eurydema* were labeled with circles, and conserved sites within Pentatominae were marked with red color.

stem, TΨC stem, and amino acid acceptor stems of *trnL1*(CUN), *trnL2*(UUR), *trnS1*(AGN), and *trnS2*(UCN) show 100% identity within *Eurydema*. Similarly, the anticodon loops in *trnW*, *trnK*, *trnL2*, and *trnF* show 100% identity within the Pentatominae. We detected a total of seven pairs of conserved G-U mismatches within the Pentatominae, which contribute to maintaining the stability of tRNA secondary structures. Additionally, A-G in the DHU stem of *trnH* and C-U in the anticodon stem of *trnM* are also conserved across *Eurydema*.

Mitogenomes of the Pentatominae are characterized by large rRNA (16S rRNA), located between *trnL1*(CUN) and *trnV*, and small rRNA (12S rRNA) inserted between *trnV* and the control region. The predicted secondary structures of the *E. dominulus* rRNAs are depicted in Figs. 7 and 8. The stem region structure of rRNAs is more highly conserved than the loop structure. The 16S rRNA contained five domains (the third domain is absent, as in other arthropods) with 44 predicted stems, containing 54.38% conserved sites within the Pentatominae. The 12S rRNA contains three domains with 26 predicted stems, containing 53.42% conserved sites within the Pentatominae.

Control Region

Compared with other regions, the control regions of the Pentatominae mitogenomes between the 12S rRNA and *trnM* exhibited more variation in length, ranging from 543 to 2,190 bp. It is obvious that the length variation among them contribute to the

total length differences of their mitogenomes. This region also harbored the highest A + T content (avg. 76.64%) in the mitogenomes, whereas AT skewness and GC skewness were significantly different, which was also observed in other insect mitogenomes from Heteroptera.

The comparison of tandem repeats in control regions of *Eurydema* mitogenomes was shown in Fig. 9. The length and copies of repeat units were different among five species. However, the first repeat unit of 18bp was found in all regions except *E. qinlingensis*.

Phylogenetic Analysis

PCGs and PCGRNA datasets were used to construct phylogenetic trees using BI and ML methods. For both datasets, *Iss* values were lower than those of *Iss.c*: PCGs [*Iss*(0.272) < *Iss.c* (0.851)] and PCGRNA [*Iss* (0.516) < *Iss.c* (1.215)]. In order to evaluate sequence variation heterogeneity, the two datasets were analyzed using AliGROOVE, and we found that there was low heterogeneity in the sequence composition of both datasets (Fig. 10). Overall, the saturation and heterogeneity analyses indicated that the two datasets were suitable for further phylogenetic investigation.

The final alignments contained 11,609 and 16,579 sites in the PCGs matrix and PCGRNA matrix, respectively. The monophyly of each family and tribe were generally well supported in both BI and ML trees (Fig. 11); most posterior probabilities (PP) are 1.00 and bootstrap pseudoreplicates (BP) are 100. The topology of the Pentatomoidea was as follows: (Urostylididae + (Plataspidae

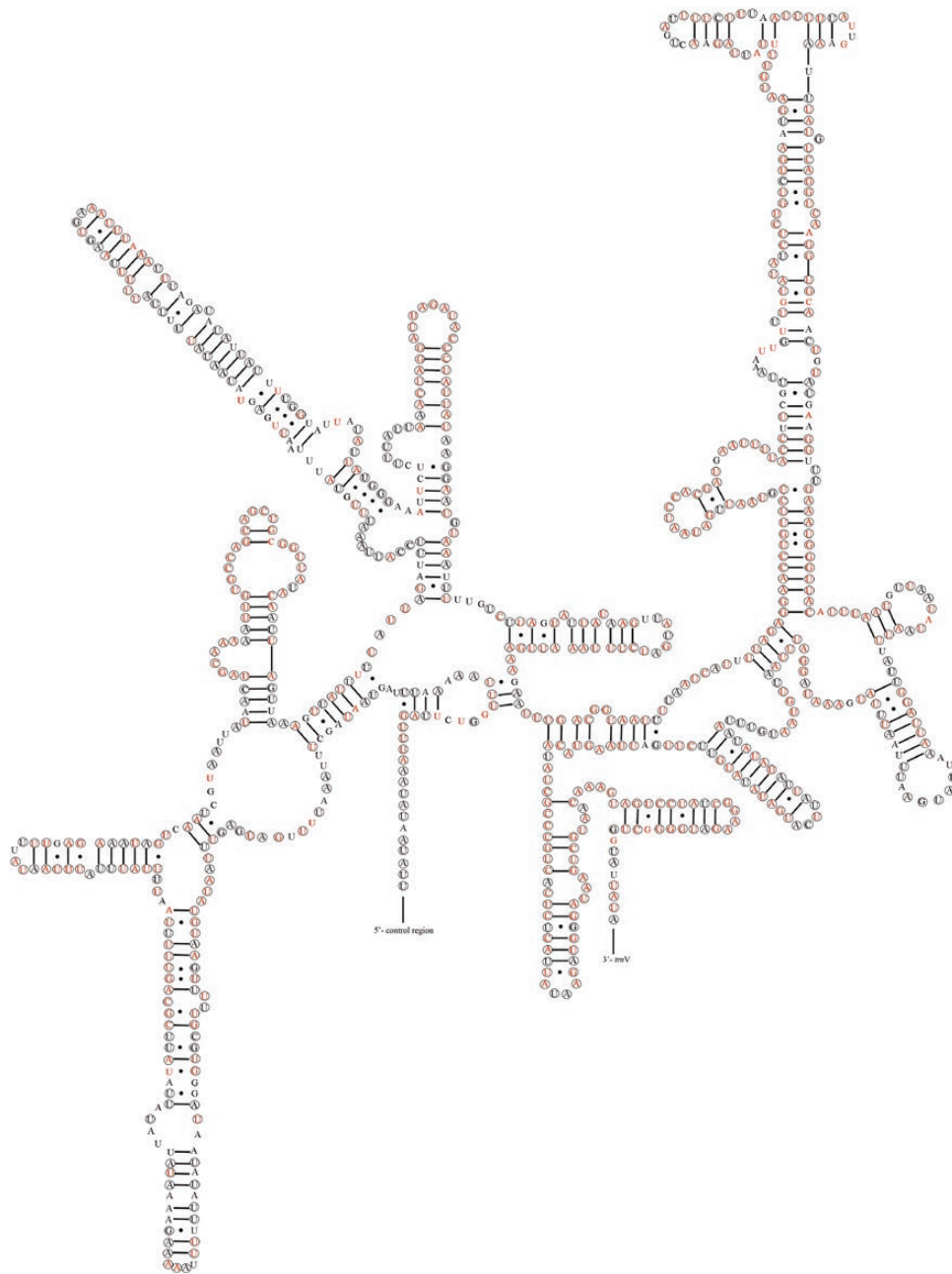


Fig. 8. Potential secondary structure of 12S rRNA in *Eurydema dominulus*. The conserved sites within *Eurydema* were labeled with circles, and conserved sites within Pentatominae were marked with red color.

+((Cydnidae +(Dinidoridae +Tessaratomidae)) +Pentatomidae))). The phylogenetic analyses performed in the present study indicated that *Eurydema* species belong to Strachiini, which form a sister group with Pentatomini (PP= 1.00/1.00 and BP = 100/100). *Graphosoma rubrolineatum* was resolved as a sister to Carporcorini, suggesting that this species belongs to the Graphosomini tribe of Pentatominae (PP = 0.71/1.00 and BP = 86/94). The stable clade (Strachiini + Pentatomini) was identified as a sibling to the clade comprising Cappaeini, Graphosomini, and Carporcorini (PP = 1.00/1.00 and BP = 100/100). In addition, *E. qinlingensis* was found to be the first diverging clade within *Eurydema*, whereas the others divided into two clades. *E. ventralis*, *E. liturifera*, and *E. oleracea* formed the clade I, and *E. maracandica*, *E. gebleri*, and *E. dominulus* formed the clade II (PP = 1.00/1.00 and BP = 100/100).

Phylogenetic Informativeness

We calculated the net and per site PI values along the root-to-tip axis for each gene and the three codon positions of PCGs (Fig. 12). PI analysis measures the power of a set of characters to resolve branching order in a phylogenetic tree. The PI curves for PCGs were found to be similar in shape, with a steady increase from the root, to a maximum close to the tip at a hierarchical level that defines the phylogenetic relationships, and then dropped rapidly. For PCGs, *atp8* had the lowest net value (0.77), whereas *nad5* had the highest per site PI value (103.57). The third codon positions had the highest PI along the entire root-to-tip axis both for net and per site values. The PI curves for functional RNAs were similar in shape, with a peak near the tip. For rRNAs, we obtained contrasting net and per site PI results. Specifically, the per site PI of 16S rRNA (70.09) was

higher than that 12S rRNA (51.93), whereas the net PI of 12S rRNA (1.43) was slightly higher than that of 16S rRNA (1.32). For tRNAs, *trnM* and *trnI* had the highest net PI values, whereas *trnS1* had the highest per site PI. The high PI values of PCGs and rRNAs resolved tribe relationships within Pentatominae, whereas tRNAs corresponded to generic relationships.

Discussion

In this study, we describe the mitochondrial genomes of five *Eurydema* species within the family Pentatomidae. Comparative analysis using the previously sequenced mitogenomes of 14 Pentatomidae species revealed relatively low conservation in the control region, the length of which varied from 543 to 2,190 bp, and is the results of the number of repeating units copies. Furthermore, species invariably show different repeat unit even at the genus level. In contrast, we detected relatively limited variation with respect to the length of PCGs, tRNAs, and rRNAs, with PCGs showing the least variation (23 bp). Given their stable secondary structures, tRNAs and rRNAs also tend to contain conserved base pairs, such as A-G in the DHU stem and C-U in the anticodon stem (Hughes 1996).

We found that the A+T contents of Pentatomidae mitogenomes (72.05–77.43%) is significantly higher than that of G+C as other heteropteran species and is lower than the Lepidoptera mitogenomes (77.84–81.59%) (Hua et al. 2008, Wang et al. 2016, Li et al. 2017).

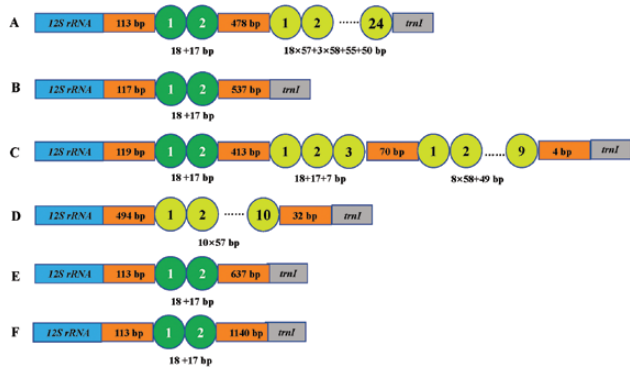


Fig. 9. Organization of the control region in *Eurydema* mitochondrial genomes. (A) *Eurydema dominulus*; (B) *Eurydema maracandica*; (C) *Eurydema oleracea*; (D) *Eurydema qinlingensis*; (E) *Eurydema ventralis*; and (F) *Eurydema gebleri*.

Among the 13 PCGs, we found that the *atp8* and *cox1* genes have the highest and lowest A+T content, respectively, with content differing by more than 10%. Coincidentally, the *atp8* and *cox1* genes were found to have the highest (0.63) and lowest (0.05) Ka/Ks values, respectively. These results indicated that the evolutionary rate of proteins are related to their nucleotide composition, which is consistent with the view of Du et al. (2018). G+C content influences amino acid compositions, and then changes the proteins' evolution (Jordan et al. 2005).

Most PCGs in heteropteran mitogenomes start with an ATN codon, although *cox1* begins with a TTG start codon (Hua et al. 2008, Yuan et al. 2015). Among Pentatomidae PCGs, TTG is the most frequently used start codon in *atp8* and *nad1*. Furthermore, among the synonymous codons encoding the same amino acid in Pentatomidae, those ending with an A or U are more frequently than those ending in a G or C, and also observed in other heteropteran species (Wang et al. 2015).

In general, a canonical clover-leaf secondary structure is predicted for tRNA genes; however, *trnS1* and *trnV* lack the DHU arm and form an aberrant loop structure (Johnson et al. 2018). In addition to the DHU arm, the TΨC stem also varies in length within different tRNAs (Navajas et al. 2002). Typically, the stem region structure of RNAs is more highly conserved than the loops, particularly the anticodon arm, and in line with expectations, we found that the anticodon arms in *trnW*, *trnK*, *trnL2* and *trnF* within Pentatomidae show 100% identity. For rRNAs, more than half of the sites are conserved within the Pentatomidae, with most of the conserved sites being located in stem regions.

The PCGs and PCGRNA databases used in the present study were evaluated for saturation and heterogeneity, and both were found to meet the criteria of phylogenetic analysis. In the ML and BI trees, the monophyly of different families within the Pentatomioidea were strongly supported based on the different datasets. The basal position of Urostyliidae and the innermost position of Pentatomidae are consistent with previous hypotheses based on morphological data (Gapud 1991, Grazia et al. 2008) and molecular phylogenetic analyses (Grazia et al. 2008, Lis et al. 2012, Kocher et al. 2014). The Dinidoridae and Tessaratomidae formed a sister clade with nodal supports 100 for ML and 1.00 for BI. The close relationship of these two families was also suggested based on morphology (Henry 1997, Kment and Vilimova 2010) and molecular phylogenies (Grazia et al. 2008, Yuan et al. 2015).

The monophyly of tribes Strachiini and Pentatomini was consistently supported in ML and BI analyses with strong nodal support. The sister relationship of these two tribes is also suggested by

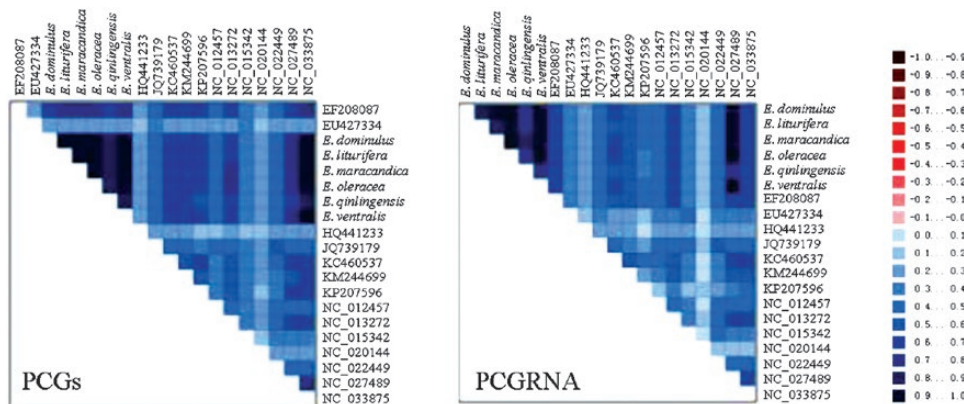


Fig. 10. Heterogeneity analysis of PCGs and PCGRNA datasets.

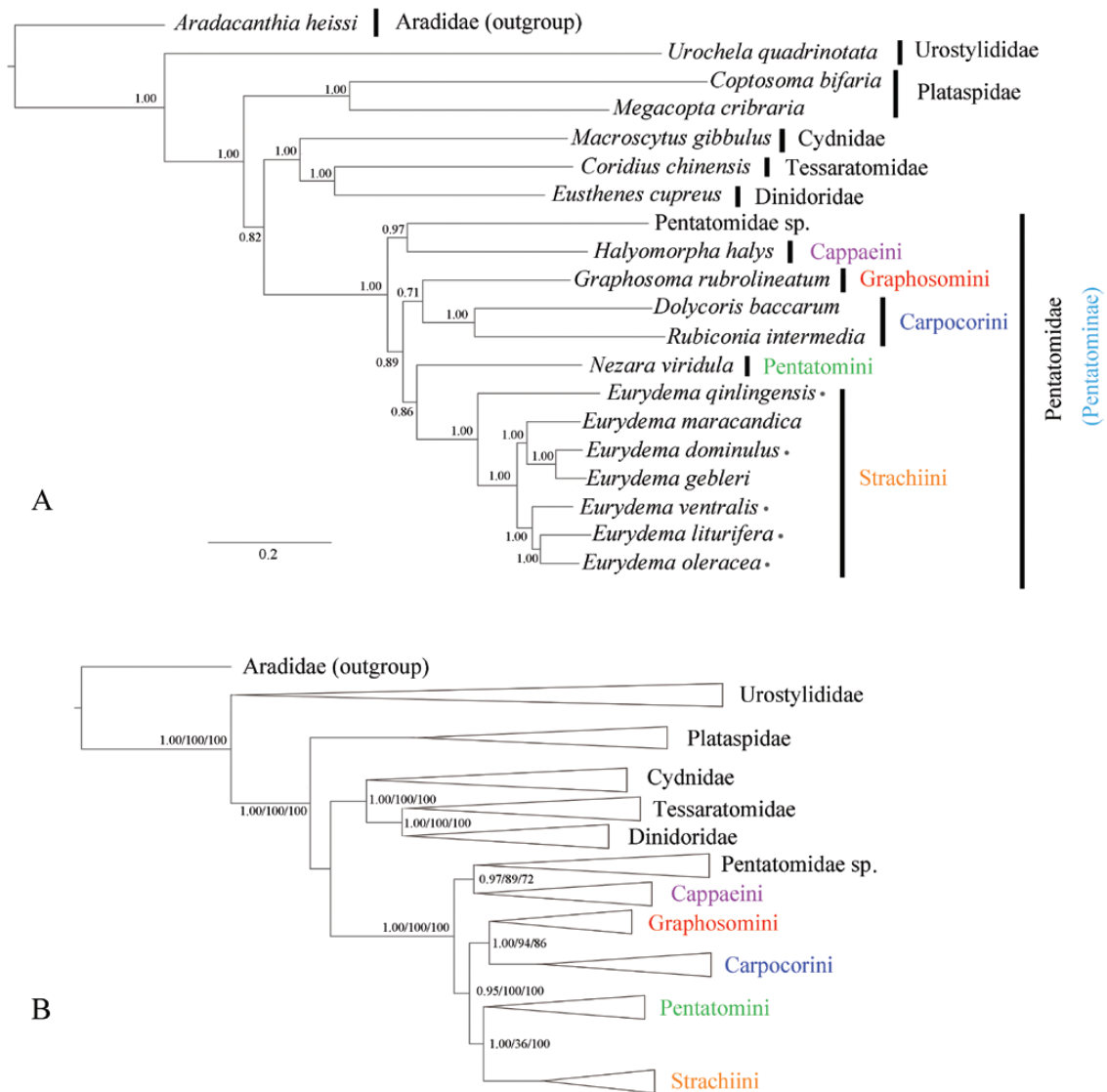


Fig. 11. Phylogenetic trees based on Pentatomidae mitogenomes. (A) The topology from the analysis of BI-PCGs. (B) The schematic version of trees from the analysis of BI-PCGRNA, ML-PCGs and ML-PCGRNA.

morphological data (Yang 1962). In the past, the taxonomic status of the genus *Graphosoma* has tended to be somewhat ambiguous. On the basis of morphological analysis, most researchers adopt the view of that *Graphosoma* belongs to Podopinae (Rider 2006, Durak and Kalender 2009, Yuan et al. 2015). Nevertheless, many specialists hold the opinion that *Graphosoma* should be assigned to the tribe Graphosomini of Pentatominae. Although the phylogenetic analysis conducted in the present study was based a limited number of mitogenomes, our finding revealed that *G. rubrolineatum* is sister to Carpororini, indicating that *Graphosoma* is a genus within the Pentatominae and not the Podopinae.

Our phylogenetic analyses also provided strong support for the assumed close relationship among *E. dominulus*, *E. gebleri*, and *E. maracandica*, and *E. ventralis*, *E. liturifera*, and *E. oleracea*, which were consistent with their morphological features. For instance, *E. dominulus*, *E. gebleri*, and *E. maracandica* were doubted to be identified due to unobvious morphological and structural differences, whereas abdominal spot variation in *E. ventralis*, *E. liturifera*, and *E. oleracea* is more obvious than that in the other three species.

In order to examine the contribution of different gene types, we assessed the PI of genes and partitions. We found that the PCGs are more variable than the RNA genes and that the phylogenetic information of the third position of the PCG codons is notably higher than that of either the first or second position. These results, therefore, indicated that the PCGs and particularly the third position of these genes constitute highly informative phylogenetic characters for the resolution of phylogenetic relationships (Nie et al. 2018).

In the present study, five mitogenomes from the Pentatomidae were added, the available mitogenomes for this group are still limited given the species richness. We accordingly emphasize the necessity for further sequencing of mitogenomes in the family Pentatomidae, which will contribute to the resolution of phylogenetic relationships.

Supplementary Data

Supplementary data are available at *Journal of Insect Science* online.

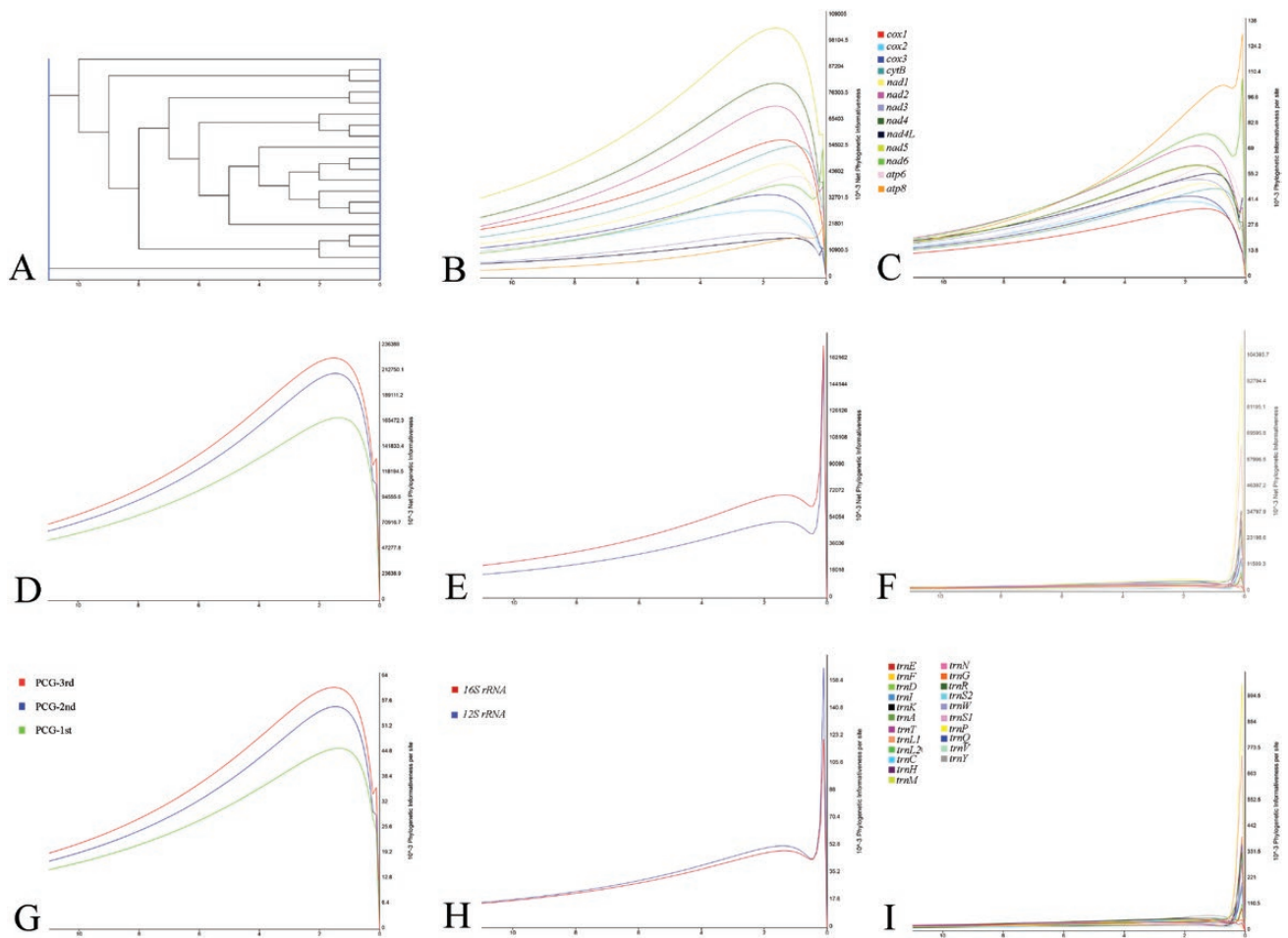


Fig. 12. The phylogenetic information of mitochondrial genes/characters.

Acknowledgments

This project was supported by the National Science Foundation Project of China (No. 31501876, 31872272), and supported by Scientific Research Project of Xinzhou Teachers University (2018KY04).

References Cited

- Akhter, N., S. T. Ahmad, and M. N. Azim. 2015. On the biology of agriculturally important pentatomid pests *Dolycoris indicus* Stål and *Eurydema Pulchrum* Westwood (Hemiptera: Pentatomidae). *N. Y. Sci. J.* 8(2): 89–91.
- Alexandros, S. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Bernt, M., A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritsch, J. Pütz, M. Middendorf, and P. F. Stadler. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* 69: 313–319.
- Breeschoten, T., C. Doorenweerd, S. Tarasov, and A. P. Vogler. 2016. Phylogenetics and biogeography of the dung beetle genus *Onthophagus* inferred from mitochondrial genomes. *Mol. Phylogenet. Evol.* 105: 86–95.
- Cameron, S. L., C. L. Lambkin, S. C. Barker, and M. F. Whiting. 2010. A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. *Syst. Entomol.* 32: 40–59.
- Coil, D., G. Jospin, and A. E. Darling. 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* 31: 587–589.
- Colinet, H., D. Renault, and D. Roussel. 2017. Cold acclimation allows *Drosophila* flies to maintain mitochondrial functioning under cold stress. *Insect Biochem. Mol. Biol.* 80: 52–60.
- Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods.* 9: 772.
- Djoumad, A., A. Nisole, R. Zahiri, L. Freschi, S. Picq, D. E. Gundersen-Rindal, M. E. Sparks, K. Dewar, D. Stewart, H. Maaroufi, et al. 2017. Comparative analysis of mitochondrial genomes of geographic variants of the gypsy moth, *Lymantria dispar*, reveals a previously undescribed genotypic entity. *Sci. Rep.* 7: 14245.
- Dowton, M., L. R. Castro, and A. D. Austin. 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome ‘morphology’. *Invertebr. Syst.* 16: 345–356.
- Du, M. Z., S. Liu, Z. Zeng, L. A. Alemayehu, W. Wei, and F. B. Guo. 2018. Amino acid compositions contribute to the proteins’ evolution under the influence of their abundances and genomic GC content. *Sci. Rep.* 8: 7382.
- Durak, D., and Y. Kalender. 2009. Fine structure and chemical analysis of the metathoracic scent gland secretion in *Graphosoma lineatum* (Linnaeus, 1758) (Heteroptera, Pentatomidae). *C. R. Biol.* 332: 34–42.
- Fischer, C., M. Mahner, and E. Wachmann. 2000. The rhabdom structure in the ommatidia of the Heteroptera (Insecta), and its phylogenetic significance. *Zoomorphology* 120: 1–13.
- Françoso, E., F. Gomes, and M. C. Arias. 2016. A protocol for isolating insect mitochondrial genomes: a case study of NUMT in *Melipona flavolineata* (Hymenoptera: Apidae). *Mitochondrial DNA A. DNA Mapp. Seq. Anal.* 27: 2401–2404.

- Gapud, V. 1991. A generic revision of the subfamily Asopinae with consideration of its phylogenetic position in the family Pentatomidae and superfamily Pentatomoidea (Hemiptera-Heteroptera). *Philippine Entomol.* 8: 865–961.
- Grazia, J., R. T. Schuh, and W. C. Wheeler. 2008. Phylogenetic relationships of family groups in Pentatomoidea based on morphology and DNA sequences (Insecta: Heteroptera). *Cladistics* 24: 932–976.
- Hegeudusova, E., B. Břejova, L. Tomaska, M. Šipiczki, and J. Nosek. 2014. Mitochondrial genome of the basidiomycetous yeast *Jaminaea angkorensis*. *Curr. Genet.* 60: 49–59.
- Henry, T. J. 1997. Phylogenetic analysis of family groups within the infra-order Pentatomomorpha (Hemiptera: Heteroptera), with emphasis on the Lygaeoidea. *Ann. Entomol. Soc. Am.* 90: 275–301.
- Henry, T. J. 2009. Biodiversity of heteroptera. *Insect biodivers. Sci. Soc.* 1: 223–263.
- Hsiao, T.Y. 1977. A handbook for the determination of the Chinese Heteroptera. Vol. I. Science Press, Beijing, China, pp. 140–141.
- Hua, J., M. Li, P. Dong, Y. Cui, Q. Xie, and W. Bu. 2008. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). *BMC Genomics.* 9: 610.
- Hughes, J. M. 1996. Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. *J. Mol. Biol.* 259: 645–654.
- Johnson, K. P., C. H. Dietrich, F. Friedrich, R. G. Beutel, B. Wipfler, R. S. Peters, J. M. Allen, M. Petersen, A. Donath, K. K. O. Walden, et al. 2018. Phylogenomics and the evolution of hemipteroid insects. *Proc. Natl. Acad. Sci. USA.* 115: 12775–12780.
- Jordan, I. K., F. A. Kondrashov, I. A. Adzhubei, Y. I. Wolf, E. V. Koonin, A. S. Kondrashov, and S. Sunyaev. 2005. A universal trend of amino acid gain and loss in protein evolution. *Nature.* 433: 633–638.
- Jung, S., I. M. Kerzhner, and S. Lee. 2011. Species of the genus *Eurydema* (Hemiptera: Heteroptera: Pentatomidae) in Far East Asia: an integrated approach using morphological, molecular, and data crossing analyses for taxonomy. *Zootaxa.* 2917: 48–58.
- Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28: 1647–1649.
- Kment, P., and J. Vilimova. 2010. Thoracic scent efferent system of the Tessaratomidae sensu lato (Hemiptera: Heteroptera: Pentatomoidea) with implication to the phylogeny of the family. *Zootaxa* 2363(1): 1–59.
- Kocher, A., M. Kamilari, E. Lhuillier, E. Coissac, J. Péneau, J. Chave, and J. Murienne. 2014. Shotgun assembly of the assassin bug *Brontostoma colossus* mitochondrial genome (Heteroptera, Reduviidae). *Gene.* 552: 184–194.
- Kück, P., S. A. Meid, C. Groß, J. W. Wägele, and B. Misof. 2014. AliGROOVE—visualization of heterogeneous sequence divergence within multiple sequence alignments and detection of inflated branch support. *BMC Bioinformatics.* 15: 294.
- Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33: 1870–1874.
- Li, M., Y. Tian, Y. Zhao, and W. Bu. 2012. Higher level phylogeny and the first divergence time estimation of Heteroptera (Insecta: Hemiptera) based on multiple genes. *PLoS One.* 7: e32152.
- Li, H., J. M. Leavengood, E. G. Chapman, D. Burkhardt, F. Song, P. Jiang, J. Liu, X. Zhou, and W. Cai. 2017. Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. *Proc. R. Soc. B Biol. Sci.* 284: 20171223.
- Lis, J. A., P. Lis, D. J. Ziaja, and A. Kocorek. 2012. Systematic position of Dinidoridae within the superfamily Pentatomoidea (Hemiptera: Heteroptera) revealed by the Bayesian phylogenetic analysis of the mitochondrial 12S and 16S rDNA sequences. *Zootaxa* 3423: 61–68.
- Liu, H., Z. Li, S. Li, and W. Cai. 2007. Application of Mitochondrial DNA sequences in molecular systematics of true bugs (Hemiptera: Heteroptera). *Entomotaxonomia.* 29(4): 265–274.
- Liu, X., H. Li, Y. Cai, F. Song, J. J. Wilson, and W. Cai. 2017. Conserved gene arrangement in the mitochondrial genomes of barklouse families Stenopsocidae and Psocidae. *Front. Agric. Sci. Eng.* 4(3): 358–365.
- López-Giráldez, F., and J. P. Townsend. 2011. PhyDesign: an online application for profiling phylogenetic informativeness. *BMC Evol. Biol.* 11: 152.
- Ma, C., P. Yang, F. Jiang, M. P. Chapuis, Y. Shali, G. A. Sword, and L. Kang. 2012. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Mol. Ecol.* 21: 4344–4358.
- Navajas, M., Y. Le Conte, M. Solignac, S. Cros-Arteil, and J. M. Cornuet. 2002. The complete sequence of the mitochondrial genome of the honeybee ectoparasite mite Varroa destructor (Acari: Mesostigmata). *Mol. Biol. Evol.* 19: 2313–2317.
- Nie, R., T. Breeschoten, J. T. N. Timmermans, K. Nadein, H. J. Xue, M. Bai, Y. Huang, X. Yang, A. P. Vogler. 2018. The phylogeny of Galerucinae (Coleoptera: Chrysomelidae) and the performance of mitochondrial genomes in phylogenetic inference compared to nuclear rRNA genes. *Cladistics* 34(2): 113–130.
- Rider, D. A. 2006. Family pentatomidae leach, 1815, pp. 233–402. *In* B. Aukema and C. H. Rieger (eds.), *Catalogue of the heteroptera of the palaearctic region. Vol. 5. pentatomomorpha II. The Netherlands Entomological Society, Amsterdam, the Netherlands*, pp. 364–374.
- Rider, D. A. 2015. Pentatomoidea home page. North Dakota State University. Available from <http://www.ndsu.nodak.edu/ndsu/rider/Pentatomoidea/>
- Rider, D. A., and L. Y. Zheng. 2002. Checklist and nomenclatural notes on the Chinese Pentatomidae (Heteroptera) I. Asopinae. *Entomotaxonomia.* 24(2): 107–115.
- Rider, D. A., and L. Y. Zheng. 2005. Checklist and nomenclatural notes on the Chinese Pentatomidae (Heteroptera). III. Phyllocephalinae, Podopinae. *Proc. Entomol. Soc. Washington.* 107(1): 90–98.
- Rider, D. A., L. Y. Zheng, and I. M. Kerzhner. 2002. Checklist and nomenclatural notes on the Chinese Pentatomidae (Heteroptera). II. Pentatominae. *Zoosyst. Ross.* 11(1): 135–153.
- Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61: 539–542.
- Vaidya, G., D. J. Lohman, and R. Meier. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27: 171–180.
- Wang, I. J. 2010. Recognizing the temporal distinctions between landscape genetics and phylogeography. *Mol. Ecol.* 19: 2605–8.
- Wang, Y., J. Chen, L. Y. Jiang, and G. X. Qiao. 2015. Hemipteran mitochondrial genomes: features, structures and implications for phylogeny. *Int. J. Mol. Sci.* 16(6): 12382–12404.
- Wang, Y. H., C. Ying, D. Rédei, P. Baňář, X. Qiang, P. Štys, J. Damgaard, P. P. Chen, W. B. Yi, and W. Ying. 2016. Phylogenetic divergences of the true bugs (Insecta: Hemiptera: Heteroptera), with emphasis on the aquatic lineages: the last piece of the aquatic insect jigsaw originated in the Late Permian/Early Triassic. *Cladistics* 32: 404–407.
- Weirauch, C., and R. T. Schuh. 2010. Systematics and evolution of Heteroptera: 25 years of progress. *Annu. Rev. Entomol.* 56: 487–510.
- Wolstenholme, D. R. 1992. Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.* 141: 173.
- Wu, L. W., H. Chiba, D. C. Lees, Y. Ohshima, and M. L. Jeng. 2019. Unravelling relationships among the shared stripes of sailors: mitogenomic phylogeny of Limenitidini butterflies (Lepidoptera, Nymphalidae, Limenitidinae), focusing on the genera *Athyma* and *Limenitis*. *Mol. Phylogenet. Evol.* 130: 60–66.
- Xia, X. 2018. DAMBE7: new and improved tools for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* 35: 1550–1552.
- Xing, X. 2013. A study on taxonomy of subfamily podopinae from China (Hemiptera: Heteroptera: Pentatomidae). Nankai University, Tianjin, China.
- Yang, W. I. 1962. Hemiptera, Pentatomidae. *In* Zoology Editorial Committee of Chinese Academy of Sciences (eds.), *Economic insect fauna of China, vol. 2. Academia Sinica, Science Press, Beijing, China*, pp. 104–111.
- Yang, Q. Q., S. W. Liu, F. Song, G. F. Liu, and X. P. Yu. 2018. Comparative mitogenome analysis on species of four apple snails (Ampullariidae: Pomacea). *Int. J. Biol. Macromol.* 118: 525–533.

- Yuan, M. L., Q. L. Zhang, Z. L. Guo, J. Wang, and Y. Y. Shen. 2015. Comparative mitogenomic analysis of the superfamily Pentatomoidea (Insecta: Hemiptera: Heteroptera) and phylogenetic implications. *BMC Genomics*. 16: 460.
- Zhao, W., Q. Zhao, M. Li, J. Wei, X. Zhang, and H. Zhang. 2017a. Characterization of the complete mitochondrial genome and phylogenetic implications for *Eurydema maracandica* (Hemiptera: Pentatomidae). *Mitochondrial DNA, Part B*. 2(2): 550–551.
- Zhao, W., Q. Zhao, M. Li, J. Wei, X. Zhang, and H. Zhang. 2017b. DNA Barcoding of Chinese species of the genus *Eurydema* Laporte, 1833 (Hemiptera: Pentatomidae). *Zootaxa* 4286: 151–175.
- Zhao, Q., J. Wang, M. Q. Wang, B. Cai, H. F. Zhang, and J. F. Wei. 2018. Complete mitochondrial genome of *Dinorhynchus dybowskyi* (Hemiptera: Pentatomidae: Asopinae) and phylogenetic analysis of Pentatomomorpha species. *J. Insect Sci.* 18(2): 44.