



# Comparative Mitgenome Analysis of *Anoplophora horsfieldi* and Other Chrysomeloidea, Cucujiformia Insects Reveals Conserved Mitogenome Organization and Phylogeny

Haifen Qin<sup>1,2</sup>, Yujia Liu<sup>1</sup>, Yujie Zhang<sup>1</sup>, Jingfeng Liu<sup>1</sup>, Zhenkun Zhao<sup>1,2</sup> and Lichun Jiang<sup>1,2\*</sup>

<sup>1</sup>Key Laboratory for Molecular Biology and Biopharmaceutics, School of Life Science and Technology, Mianyang Normal University, Mianyang, Sichuan, 621000, P.R. China.

<sup>2</sup>Ecological Security and Protection Key Laboratory of Sichuan Province, Mianyang Normal University, Mianyang, Sichuan, 621000, P.R. China.

## ABSTRACT

Mitochondrial genomes are important markers using to reconstruct phylogenetic status and reveal insect molecular evolution. In this study, the mitochondrial genome (mitogenome) of *Anoplophora horsfieldi* (Coleoptera: Chrysomelidae) was determined using high-throughput sequencing. The size of circular mitogenome is 15,796 bp and it includes a typical structure of 13 protein-coding genes (PCGs), two ribosomal RNA genes (*rRNAs*), 22 transfer RNA genes (*tRNAs*) and an adenine + thymine (A+T)-rich region. The base composition of the major strand is A: 39.74%, T: 39.57%, G: 8.22%, and C: 12.47%, with a A+T content of 79.31%. The genomic analyses indicated that its gene arrangement and content is similar to that of other Cucujiformia species. The *Anoplophora* species control region sequence with rich A+T content exhibited high genetic variability. All PCGs initiate with ATN and terminate with the TAA or TAG except for COXI-COXIII and ND3-ND5, where they end with an incomplete stop codon (T-). All tRNAs form clover-leaf structure only apart from trnS (AGN) which possess a reduced DHU arm. The motifs 'ATGATAA' between ND4L and ND4, was more conserved than that between trnS (UCN) and ND1 and between ATP8 and ATP6 in the mitogenomes of Cucujiformia. The 1,143 bp A+T-rich area includes a 16 bp poly-T stretch, 14 bp poly-A stretch region, three microsatellite-like repeats of (TA)<sub>n</sub> and three other random repetitive sequences. Based on 13 PCGs of 118 Cucujiformia mitogenomes, the phylogenetic analyses are reconstructed with both maximum likelihood and Bayesian analyses and a consistent topology is formed. The results show that *A. horsfieldi* grouped with *A. glabripennis* and *A. chinensis* with high nodal supports. It also supports such phylogenetic relationships of ((Lymexyloidea + Tenebrionoidea) + (Curculionoidea + (Chrysomeloidea + (Cucujoidea + (Cleroidea + Coccinelloidea)))) within Cucujiformia. Therefore, *A. horsfieldi* mitogenome enriches our understanding of the phylogenetic relationship of Chrysomelidae. In addition, it is used to establish phylogenetic trees and further study the phylogenetic relationship between Cucujiformia and Chrysomeloidea.

## Article Information

Received 24 August 2022

Revised 15 September 2022

Accepted 01 October 2022

Available online 08 May 2023

(early access)

Published 26 August 2024

## Authors' Contribution

HQ investigation; collect samples, analyze the data. YL investigation; collect samples; contribute analysis tools. YZ investigation; collect samples; contribute analysis tools; organize tables and beautify pictures. JL analyze the data; prepare figures and tables. ZZ investigation, collect samples. LJ conceive and designed the experiments; write the paper.

## Key words

*Anoplophora horsfieldi*, Coleoptera, Cucujiformia, Chrysomeloidea, mitochondrial genome, molecular phylogeny

## INTRODUCTION

*Anoplophora* is an important genus in the longhorned beetle family Cerambycidae. They are widely distributed in the Oriental and eastern Palearctic regions,

and contain 36 species (McDougall, 2001; Richard and Gregory, 2002). Several notorious pest insects belong to this genus (McDougall, 2001). *Anoplophora horsfieldi* has body length of about 35 mm. *Camellia sinensis*, *Celtis sinensis*, *Quercus glauca* and *Ulmus pumila* are main host plants. Adults appeared in summer, and lived in low and middle mountainous areas. This species can be found in India, Vietnam and China (Chen et al., 1959). Most of the Cerambycidae species, as a large family in the Coleoptera order, are forestry pests with above 25,000 species worldwide (Sama et al., 2010; Wang et al., 2012). Although Cerambycidae shows large taxonomic diversity, it is still limited knowledge about the Cerambycidae mitogenome.

Mitogenomes are important molecular markers,

\* Corresponding author: [jiang\\_lichun@126.com](mailto:jiang_lichun@126.com)  
0030-9923/2024/0005-2449 \$ 9.00/0



Copyright 2024 by the authors. Licensee Zoological Society of Pakistan.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

which can be used for exploring the phylogenetics, population genetics and phylogeographics of species due to low recombination rate, maternal inheritance, conserved gene organization and orientation and high mutation rate (Dowton *et al.*, 2002; Jin *et al.*, 2004; Wang and Tang, 2018). In general, the metazoan mitogenome forms a closed-circular structure, with a length of 15 to 18 kilobases (kb) (Dowton *et al.*, 2002; Lu *et al.*, 2013; Cameron, 2014), including 13 PCGs, 22 tRNAs, 2 rRNAs and a non-coding region (A+T-rich region) that in other Cucujiformia and in invertebrates is named as A+T-rich region because of its extremely high content in Adenines and Thymines (Wolstenholme, 1992; Cameron, 2014; Sun *et al.*, 2017; Zhang *et al.*, 2019). Recent high throughput sequencing development has enriched insect mitogenome datasets and accelerated the research on insect molecular evolution, too. At present, insect mitochondrial genome data has been used for species identification, comparative genomics, evolutionary genetics, reconsecration of phylogeny and phylogeography (Cameron and Whiting, 2008; Hao *et al.*, 2012; Ma *et al.*, 2012; Timmermans *et al.*, 2014; Li *et al.*, 2016).

Coleoptera is an ancient order and can adapt to a variety of habitats (Li *et al.*, 2016). Coleoptera has approximately 400,000 described species and contains Adephaga, Archostemata, Myxophaga and Polyphaga (Crowson, 1960; Nie and Yang, 2013; Zhang, 2013). Based on morphological features, Crowson (1960) suggested Polyphaga can be divided into five infraorders and 18 superfamilies. However, Bouchard *et al.* (2011) think that Polyphaga can be divided into six infraorders and 16 superfamilies. Because of different phylogenetic positioning of some superfamilies, so these classification systems contain different numbers of infraorders (Nie and Yang, 2013). Cucujiformia, as the most highly diversified infraorders of Polyphaga (Crowson, 1960; ØDegaard, 2000; Bouchard *et al.*, 2009; Wang and Tang, 2018) includes six superfamilies and most of two superfamilies species (Curculionoidea and Chrysomeloidea) are plant-feeding beetles (Grimaldi and Engel, 2005). Recently, Robertson *et al.* (2015) and Yuan *et al.* (2016) suggest that Cucujiformia can be divided into seven superfamilies (Cleroidea, Coccinelloidea, Lymexyloidea, Tenebrionoidea, Cucujoidea, Chrysomeloidea, Curculionoidea). Moreover, long-horned beetles and leaf beetles in Chrysomeloidea and weevils in Curculionoidea are considered as Coleoptera 'Phytophaga'. Chrysomeloidea contains the families Chrysomelidae, Cerambycidae, Megalopodidae, Vesperidae, Oxypeltidae, Disteniidae and Orsodacnidae (Hunt *et al.*, 2007).

The phylogeny and evolution of Chrysomeloidea have received considerable attention of systematists.

Nonetheless, the interrelationships (families and subfamilies) are unclear and many chrysomeloid evolutionary questions still persist. This is especially true for Cerambycidae, for which there is a few molecular phylogenies, however, the numbers and species used in the construction of phylogenetic trees are limited. Several studies have explored the phylogenetic relationships of Chrysomeloidea by using the combined data including morphological data (Farrell and Sequeira, 2004; Gómez-Zurita *et al.*, 2007) and molecular data (18S rRNA Hunt *et al.*, 2007; 28S rDNA Marvaldi *et al.*, 2009, and partial mitochondrial genes Bocak *et al.*, 2014; Li *et al.*, 2016; Wang and Tang, 2018). In addition, there are very few molecular studies about Chrysomeloidea taxa. Currently, some researchers have found some new evidence to understand the Chrysomeloidea phylogenetic relationships, but some conflicting results are often achieved due to the differences of datasets and analytical methods. Therefore, the phylogenetic relationships of Cucujiformia and Chrysomeloidea need to be further evaluated.

In this study, we achieved the complete sequence of the mitogenome of *A. horsfieldi* using next-generation sequencing and compared it with other cucujiform insects. We aimed to analyze the structural characteristics of *A. horsfieldi* mitogenome and compared it with previously sequenced *Anoplophora* mitogenomes. Moreover, based on all 117 mitogenome sequences also used BI and ML methods to reconstruct a tentative phylogeny to evaluate relationships among Cucujiformia and Chrysomeloidea. Our review is structured taxonomies, commencing with the phylogenetic neighborhood of Chrysomeloidea and their relatives, followed by sections on Chrysomelidae, Cerambycidae and then Cucujiformia and its relatives. The sequence characteristics and annotation of *A. horsfieldi* mitogenome will be a significant increase in further research of Cucujiformia and Chrysomeloidea mitogenome structures and phylogenetics.

## MATERIALS AND METHODS

### *Sample collection and genomic DNA extraction*

The specimen of the *A. horsfieldi*, was sampled from an adult vulture captured in the Yongxing Town, Mianyang City, Sichuan Province, China in June 2020 (104°39'5.43"E, 31°27'39.06"N, 475 m.a.s.l). The specimen was preserved in 95% ethanol and stored -76°C until genomic DNA extraction. Total genomic DNA was extracted from the thorax muscle of a single specimen using an E.Z.N.A.® Tissue DNA Kit (Omega, Norcross, GA) abiding by the manufacturer's instructions. The extracted DNA quality was examined by 0.9% agarose gel electrophoresis (w/v) and used to sequence the whole mitogenome of *A.*

*horsfieldi* with next-generation sequencing.

#### Library preparation and sequencing

Genome sequencing libraries with about 400 bp of insertion fragment were constructed with a NEXTflex™ Rapid DNA-Seq Kit (Illumina, San Diego, CA) according to the manufacture's protocols. The library was sequenced on Illumina HiSeq X Ten platform to produce 150 bp paired end reads (300 cycles).

#### Mitochondrial genome assembly and analysis

SOAPdenovo v2.0 (Luo *et al.*, 2012), MITObim v1.8 (Christoph *et al.*, 2013) and NOVOPlasty v2.7.1 (Dierckxsens *et al.*, 2017) were alternate application to assembly the mitochondrial genome. The assembled mitochondrial fragments were identified by BlastX using *A. chinensis* (Li *et al.*, 2015) (NC\_029230) and *A. glabripennis* (Fang *et al.*, 2016) (NC\_008221) mitochondrial genes as queries. Prediction and annotation of 13 PCGs, 22 tRNA and 2 rRNA genes were performed with DOGMA (<http://dogma.cccb.utexas.edu/>) or MITOS (<http://mitos.bioinf.uni-leipzig.de/index.py>) using the support of annotation from reference mitogenome.

The mitogenome sequence is further verified and verified below. 12S and 16S rRNA genes were determined by comparison using homologous sequences of mitochondrial DNA from other *Anoplophora* species with ClustalX version 2.0 (Larkin *et al.*, 2007) 13 PCG sequences were translated into putative proteins according to the Invertebrate Mitochondrial Genetic Code. The base composition of nucleotide sequences was calculated by skewness on the basis of the following formulas: AT skew =  $[A-T]/[A+T]$  and GC skew =  $[G-C]/[G+C]$  (Perna and Kocher, 1995). The content of A+T and relative synonymous codon usage (RSCU) values were described with MEGA 7.0 (Kumar *et al.*, 2016). The tRNA genes were resolved with the tRNAscan-SE 1.21 (<http://lowelab.ucsc.edu/tRNAscan-SE/>) (Lowe and Chan, 2016) and the MITOS Web Server (Perna and Kocher, 1995; Bernt *et al.*, 2013). The tRNAs secondary structure were achieved with RNAviz v2.0 (De Rijk *et al.*, 2003). The tandem repeats of the A+T-rich region were explored by the tandem repeat finder (<http://tandem.bu.edu/trf/trf.html>) (Benson, 1999; Timmermans *et al.*, 2014).

#### Phylogenetic analysis

The newly sequenced mitogenomes of *A. horsfieldi* was aligned with 116 mitogenomes of Cucujiformia and Chrysomeloidea available in GenBank, with *Prosopocoilus gracilis* (Coleoptera: Scarabaeiformia: Lucanidae: KP735805), *Cheirotonus jansonii* (Coleoptera: Scarabaeiformia: Scarabaeidae: NC\_023246) and

*Necrophila americana* (Coleoptera: Staphyliniformia: Silphidae: NC\_018352) as outgroups (Supplementary Table S1). The multiple alignments (13 PCGs concatenated nucleotide sequence datasets) were performed with ClustalX soft (v2.0 Larkin *et al.*, 2007) with the default settings. Accession numbers of all mitogenomes are listed in the phylogenetic trees (Supplementary Table S1). We employed the nucleotide sequences of the 13 PCGs as the dataset to construct the maximum likelihood (ML) and Bayesian inference (BI) phylogenetic trees. The nucleotide sequences of 13 PCGs from all 117 Cucujiformia mitogenome were aligned separately in MEGA 7.0 (Kumar *et al.*, 2016). The final alignments were refined and the conserved sequences were identified with Gblocks v0.91b (Talavera and Castresana, 2007). Geneious 8.1.6 was used to concatenate the resulting alignments (Kearse *et al.*, 2012). The optimal nucleotide substitution model was selected on the basis of the Akaike Information Criterion (Posada and Buckley, 2004) with J Model test v0.1.1 (Posada, 2008). ML and BI phylogenetic trees show different algorithms. ML phylogenetic tree was constructed by PhyML v3.0 (Guindon *et al.*, 2010) based on the best-fit model (as above) with 1,000 replicates. BI tree was conducted using MrBayes 3.2 with a GTR + I + G model, each of four chains (three hot and one cold), with run length of 20 million generations and sampling every 1,000 generations (Ronquist *et al.*, 2012). Convergence was assessed with Tracer 1.5 (Rambaut and Drummond, 2007) and trees from the first 25% of the samples were removed as burn-in. Based the value of Bayesian posterior probabilities (BPP), node support was assessed. Figtree v1.4.3 software was used to view and edit the consensus trees (Rambaut, 2009).

## RESULTS AND DISCUSSION

#### Genome structure, organization and composition

We deposited the complete mitogenome of *A. horsfieldi* in the GenBank database (MN248534) and the mitogenome sequence of *A. horsfieldi* is 15,796 bp in size (Table I and Fig. 1), which is located in the range of whole sequenced Cucujiformia species with the length ranging from 15,064 bp in *Naupactus xanthographus* (Curculionidae) to 20,124 bp in *Ceutorhynchus obstructus* (Curculionidae) (Supplementary Table S1). Alignment results revealed 38 mitogenome regions, including 13 PCGs, two rRNA, 22 tRNA and a non-coding region with high A+T-rich composition, which is common in most animal mtDNAs (Table I). The majority strand (J-strand) of the mitogenome of *A. horsfieldi* encodes nine PCGs and fourteen tRNAs, while the remaining genes are encoded on the minority strand (N-strand) (Table I).

Table 1. The mtDNA gene arrangement of three *Anoplophora* species (*Anoplophora horsfieldi*, *Anoplophora chinensis*, *Anoplophora glabripennis*).

| Gene            | <i>Anoplophora horsfieldi</i> (15,796 bp) |      |                     |               |                 |                          |       |       |                                     |        | <i>Anoplophora chinensis</i> (15,805 bp) |                     |               |                 |                          |       |       |                                     |        |     | <i>Anoplophora glabripennis</i> (15,774 bp) |               |                 |                          |       |       |                                     |      |    |  |
|-----------------|-------------------------------------------|------|---------------------|---------------|-----------------|--------------------------|-------|-------|-------------------------------------|--------|------------------------------------------|---------------------|---------------|-----------------|--------------------------|-------|-------|-------------------------------------|--------|-----|---------------------------------------------|---------------|-----------------|--------------------------|-------|-------|-------------------------------------|------|----|--|
|                 | Strand                                    |      | Position<br>From To | Sizes<br>(bp) | Nucle-<br>otide | Anti<br>codons<br>(tRNA) | Start | Stop* | Inter-<br>genic<br>nucle-<br>otide† | Strand |                                          | Position<br>From To | Sizes<br>(bp) | Nucle-<br>otide | Anti<br>codons<br>(tRNA) | Start | Stop* | Inter-<br>genic<br>nucle-<br>otide† | Strand |     | Position<br>From To                         | Sizes<br>(bp) | Nucle-<br>otide | Anti<br>codons<br>(tRNA) | Start | Stop* | Inter-<br>genic<br>nucleo-<br>tide† |      |    |  |
|                 | From                                      | To   |                     |               |                 |                          |       |       |                                     | From   | To                                       |                     |               |                 |                          |       |       |                                     | From   | To  |                                             |               |                 |                          |       |       |                                     | From | To |  |
| <i>tRNA-Ile</i> | H                                         | 1    | 63                  | 63            | GAT             |                          |       | 0     | 1                                   | 69     | 69                                       |                     | TTG           |                 | 0                        | 1     | 68    | 68                                  |        | TTG |                                             | 0             |                 |                          |       |       |                                     | 0    |    |  |
| <i>tRNA-Gln</i> | L                                         | 70   | 138                 | 69            | TTG             |                          |       | 6     | 77                                  | 145    | 69                                       |                     | TTG           |                 | 7                        | 78    | 146   | 69                                  |        | TTG |                                             | 9             |                 |                          |       |       |                                     | 9    |    |  |
| <i>tRNA-Met</i> | H                                         | 138  | 207                 | 70            | CAT             |                          |       | -1    | 145                                 | 213    | 69                                       |                     | CAT           |                 | -1                       | 146   | 214   | 69                                  |        | CAT |                                             | -1            |                 |                          |       |       |                                     | -1   |    |  |
| <i>ND2</i>      | H                                         | 205  | 1215                | 1011          |                 | ATA                      | TAA   | -3    | 214                                 | 1224   | 1011                                     |                     | ATT           | TAA             | 0                        | 215   | 1225  | 1011                                |        | ATT | TAA                                         | 0             |                 |                          |       |       |                                     | 0    |    |  |
| <i>tRNA-Trp</i> | H                                         | 1214 | 1278                | 65            | TCA             |                          |       | -2    | 1223                                | 1287   | 65                                       |                     | TCA           |                 | -2                       | 1224  | 1288  | 65                                  |        | TCA |                                             | -2            |                 |                          |       |       |                                     | -2   |    |  |
| <i>tRNA-Cys</i> | L                                         | 1271 | 1334                | 64            | GCA             |                          |       | -8    | 1280                                | 1341   | 62                                       |                     | GCA           |                 | -8                       | 1281  | 1343  | 63                                  |        | GCA |                                             | -8            |                 |                          |       |       |                                     | -8   |    |  |
| <i>tRNA-Tyr</i> | L                                         | 1335 | 1400                | 66            | GTA             |                          |       | 0     | 1342                                | 1407   | 66                                       |                     | GTA           |                 | 0                        | 1344  | 1412  | 69                                  |        | GTA |                                             | 0             |                 |                          |       |       |                                     | 0    |    |  |
| <i>COXI</i>     | H                                         | 1393 | 2935                | 1543          |                 | ATT                      | T--   | -8    | 1400                                | 2942   | 1543                                     |                     | ATT           | T--             | -8                       | 1405  | 2947  | 1543                                |        | ATT | T--                                         | -8            |                 |                          |       |       |                                     | -8   |    |  |
| <i>tRNA-Leu</i> | H                                         | 2936 | 3002                | 67            | TAA             |                          |       | 0     | 2943                                | 3007   | 65                                       |                     | TAA           |                 | 0                        | 2953  | 3017  | 65                                  |        | TAA |                                             | 5             |                 |                          |       |       |                                     | 5    |    |  |
| <i>COXII</i>    | H                                         | 3003 | 3690                | 688           |                 | ATT                      | T--   | 0     | 3008                                | 3695   | 688                                      |                     | ATT           | T--             | 0                        | 3018  | 3705  | 688                                 |        | ATT | T--                                         | 0             |                 |                          |       |       |                                     | 0    |    |  |
| <i>tRNA-Lys</i> | H                                         | 3691 | 3761                | 71            | TTT             |                          |       | 0     | 3696                                | 3765   | 70                                       |                     | TTT           |                 | 0                        | 3706  | 3775  | 70                                  |        | TTT |                                             | 0             |                 |                          |       |       |                                     | 0    |    |  |
| <i>tRNA-Asp</i> | H                                         | 3749 | 3824                | 76            | GTC             |                          |       | -13   | 3765                                | 3848   | 84                                       |                     | GTC           |                 | -1                       | 3775  | 3839  | 65                                  |        | GTC |                                             | -1            |                 |                          |       |       |                                     | -1   |    |  |
| <i>ATP8</i>     | H                                         | 3825 | 3980                | 156           |                 | ATT                      | TAA   | 0     | 3849                                | 4004   | 156                                      |                     | ATT           | TAA             | 0                        | 3840  | 3995  | 156                                 |        | ATT | TAA                                         | 0             |                 |                          |       |       |                                     | 0    |    |  |
| <i>ATP6</i>     | H                                         | 3974 | 4648                | 675           |                 | ATG                      | TAA   | -7    | 3998                                | 4672   | 675                                      |                     | ATG           | TAA             | -7                       | 3989  | 4660  | 672                                 |        | ATG | TAA                                         | -7            |                 |                          |       |       |                                     | -7   |    |  |
| <i>COXIII</i>   | H                                         | 4648 | 5434                | 787           |                 | ATG                      | T--   | -1    | 4672                                | 5460   | 789                                      |                     | ATG           | TAA             | -1                       | 4660  | 5448  | 789                                 |        | ATG | T--                                         | -1            |                 |                          |       |       |                                     | -1   |    |  |
| <i>tRNA-Gly</i> | H                                         | 5441 | 5504                | 64            | TCC             |                          |       | 6     | 5464                                | 5525   | 62                                       |                     | TCC           |                 | 3                        | 5452  | 5516  | 65                                  |        |     |                                             | 3             |                 |                          |       |       |                                     | 3    |    |  |
| <i>ND3</i>      | H                                         | 5505 | 5856                | 352           |                 | ATT                      | T--   | 0     | 5528                                | 5879   | 352                                      |                     | ATT           | T--             | 2                        | 5508  | 5856  | 349                                 |        | ATT | T--                                         | -9            |                 |                          |       |       |                                     | -9   |    |  |
| <i>tRNA-Ala</i> | H                                         | 5857 | 5921                | 65            | TGC             |                          |       | 0     | 5880                                | 5944   | 65                                       |                     | TGC           |                 | 0                        | 5857  | 5921  | 65                                  |        | TGC |                                             | 0             |                 |                          |       |       |                                     | 0    |    |  |
| <i>tRNA-Thr</i> | H                                         | 5919 | 5978                | 60            |                 |                          |       | -3    | 5941                                | 6003   | 63                                       |                     |               |                 | -4                       | 5918  | 5980  | 63                                  |        |     |                                             | -4            |                 |                          |       |       |                                     | -4   |    |  |
| <i>tRNA-Asn</i> | H                                         | 5981 | 6044                | 64            | GTT             |                          |       | 2     | 6006                                | 6069   | 64                                       |                     | GTT           |                 | 2                        | 5983  | 6046  | 64                                  |        | GTT |                                             | 2             |                 |                          |       |       |                                     | 2    |    |  |
| <i>tRNA-Ser</i> | H                                         | 6043 | 6111                | 69            | TCT             |                          |       | -2    | 6068                                | 6136   | 69                                       |                     |               |                 | -2                       | 6055  | 6123  | 69                                  |        |     |                                             | 8             |                 |                          |       |       |                                     | 8    |    |  |
| <i>tRNA-Glu</i> | H                                         | 6112 | 6174                | 63            | TTC             |                          |       | 0     | 6138                                | 6200   | 63                                       |                     |               |                 | 1                        | 6115  | 6177  | 63                                  |        |     |                                             | -9            |                 |                          |       |       |                                     | -9   |    |  |
| <i>tRNA-Phe</i> | L                                         | 6174 | 6237                | 64            | GAA             |                          |       | -1    | 6200                                | 6263   | 64                                       |                     | GAA           |                 | -1                       | 6177  | 6240  | 64                                  |        |     |                                             | -1            |                 |                          |       |       |                                     | -1   |    |  |
| <i>ND5</i>      | L                                         | 6241 | 7954                | 1714          |                 | ATA                      | T--   | 3     | 6267                                | 7977   | 1711                                     |                     | GTG           | T--             | 3                        | 6244  | 7957  | 1714                                |        | ATA | T--                                         | 3             |                 |                          |       |       |                                     | 3    |    |  |
| <i>tRNA-His</i> | L                                         | 7952 | 8014                | 63            | GTG             |                          |       | -3    | 7978                                | 8040   | 63                                       |                     |               |                 | 0                        | 7957  | 8018  | 62                                  |        |     |                                             | -1            |                 |                          |       |       |                                     | -1   |    |  |
| <i>ND4</i>      | L                                         | 8015 | 9344                | 1330          |                 | ATA                      | T--   | 0     | 8041                                | 9370   | 1330                                     |                     | ATA           | T--             | 0                        | 8020  | 9349  | 1330                                |        | ATA | T--                                         | 1             |                 |                          |       |       |                                     | 1    |    |  |
| <i>ND4L</i>     | L                                         | 9341 | 9628                | 288           |                 | ATG                      | TAA   | -4    | 9367                                | 9654   | 288                                      |                     | ATG           | TAA             | -4                       | 9346  | 9633  | 288                                 |        | ATG | TAA                                         | -4            |                 |                          |       |       |                                     | -4   |    |  |
| <i>tRNA-Thr</i> | H                                         | 9631 | 9697                | 67            | TGT             |                          |       | 2     | 9657                                | 9720   | 64                                       |                     | TGT           |                 | 2                        | 9636  | 9699  | 64                                  |        | TGT |                                             | 2             |                 |                          |       |       |                                     | 2    |    |  |
| <i>tRNA-Pro</i> | L                                         | 9698 | 9762                | 65            | TGG             |                          |       | 0     | 9721                                | 9785   | 65                                       |                     | TGG           |                 | 0                        | 9699  | 9772  | 74                                  |        | GGG |                                             | -1            |                 |                          |       |       |                                     | -1   |    |  |

Table continued on next page.....

The gene arrangement and orientation of *A. horsfieldi* mitogenome are similar to the mitogenome of typical Cerambycidae species. The differences in size of mitogenomes among Cucujiformia insects can be explained by the numbers and types of repetitive sequences in the A+T-rich regions.

Three species of the genus *Anoplophora* possess the same characteristics in the mitochondrial genome. The total base composition of *A. horsfieldi* was A (39.74%), T (39.57%), C (12.47%) and G (8.22%); A (39.48%), T (38.17%), C (13.57%) and G (8.78%) in *A. chinensis* and A (39.62%), T (38.71%), C (13.07%) and G (8.59%) in *A. glabripennis*, respectively. A heavy AT bias was found in the three species of mitochondrial genomes (79.31% in *A. horsfieldi*, 77.65% in *A. chinensis* and 78.34% in *A. glabripennis*, [Supplementary Table S2-S4](#)), which is similar to other sequenced Cerambycidae species ([Wang et al., 2013](#); [Fang et al., 2016](#); [Li et al., 2016](#)). All three *Anoplophora* species present a positive AT skew and negative GC skew in the whole mitogenome and had a higher A + T content in rRNAs than in tRNAs ([Table II](#)). In fact, the region rich in A + T has been viewed as the main source of the variations in the length of the whole mitogenome, and in the process of polynucleotide operation, insertion/deletion, the number of nucleotides is variable, and the copy numbers and types of tandem repeat elements are also very different between different species ([Zhang and Hewitt, 1997](#)).

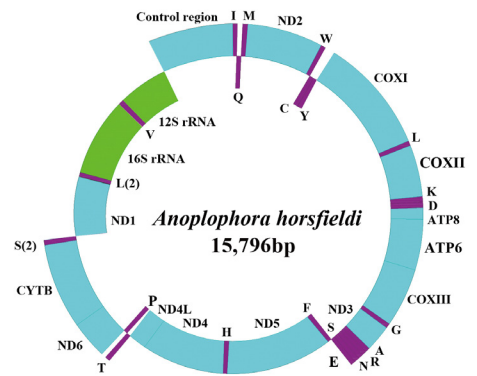


Fig. 1. Complete mitogenome organization and gene arrangement of *Anoplophora horsfieldi* from China. Genes coded on the H strand are directed to the outer ring, while the genes coded on the L strand are indicated in the interior of the ring. Genes are abbreviated as follows: ATP6 and ATP8 (subunits 6 and 8 of ATPase), COXI-COXIII (cytochrome c oxidase subunits 1-3), Cytb (cytochrome b), ND1-ND6 and ND4L (NADH dehydrogenase subunits 1-6 and 4L), 12S rRNA and 16S rRNA (ribosomal RNA of 12S and 16S), CR (control region). One-letter amino acid abbreviations were used to label the corresponding tRNA genes.

| Gene                 | <i>Anoplophora horsfieldi</i> (15,796 bp) |                     |                                  |                          |                |                                     |                     | <i>Anoplophora chinensis</i> (15,805 bp) |                          |                |                                     |                     |                                  |                          | <i>Anoplophora glabripennis</i> (15,774 bp) |                                     |  |  |  |  |  |
|----------------------|-------------------------------------------|---------------------|----------------------------------|--------------------------|----------------|-------------------------------------|---------------------|------------------------------------------|--------------------------|----------------|-------------------------------------|---------------------|----------------------------------|--------------------------|---------------------------------------------|-------------------------------------|--|--|--|--|--|
|                      | Strand                                    | Position<br>From To | Sizes<br>Nucle-<br>otide<br>(bp) | Anti<br>codons<br>(tRNA) | Start<br>Stop* | Inter-<br>genic<br>nucle-<br>otide† | Position<br>From To | Sizes<br>Nucle-<br>otide<br>(bp)         | Anti<br>codons<br>(tRNA) | Start<br>Stop* | Inter-<br>genic<br>nucle-<br>otide† | Position<br>From To | Sizes<br>Nucle-<br>otide<br>(bp) | Anti<br>codons<br>(tRNA) | Start<br>Stop*                              | Inter-<br>genic<br>nucle-<br>otide† |  |  |  |  |  |
| ND6                  | H                                         | 9765 10268          | 504                              |                          | 2              |                                     | 9788 10291          | 504                                      |                          | ATT TAA        | 2                                   | 9772 10269          | 498                              |                          | ATA TAA                                     | -1                                  |  |  |  |  |  |
| Cytb                 | H                                         | 10268 11407         | 1140                             |                          | ATT TAA        | -1                                  | 10291 11430         | 1140                                     |                          | ATG TAG        | -1                                  | 10269 11408         | 1140                             |                          | ATG TAG                                     | -1                                  |  |  |  |  |  |
| tRNA-Ser             | H                                         | 11406 11472         | 67                               | TGA                      | -2             |                                     | 11429 11495         | 67                                       | TGA                      | -2             |                                     | 11407 11473         | 67                               | TGA                      |                                             | -2                                  |  |  |  |  |  |
| ND1                  | L                                         | 11490 12437         | 948                              |                          | ATA TAG        | 17                                  | 11513 12460         | 948                                      |                          | ATA TAG        | 17                                  | 11491 12421         | 931                              |                          | ATT T-                                      | 17                                  |  |  |  |  |  |
| tRNA-Leu             | L                                         | 12442 12506         | 65                               | TAG                      | 4              |                                     | 12465 12529         | 65                                       | TAG                      | 4              |                                     | 12447 12511         | 65                               | TAG                      |                                             | 25                                  |  |  |  |  |  |
| 16S ribosomal<br>RNA | L                                         | 12507 13779         | 1273                             |                          |                | 0                                   | 12530 13802         | 1273                                     |                          |                | 0                                   | 12547 13782         | 1236                             |                          |                                             | 35                                  |  |  |  |  |  |
| tRNA-Ile             | L                                         | 13780 13849         | 70                               | TAC                      | 0              |                                     | 13804 13873         | 70                                       | TAC                      | 1              |                                     | 13786 13856         | 71                               | TAC                      |                                             | 3                                   |  |  |  |  |  |
| 12S ribosomal<br>RNA | L                                         | 13850 14653         | 804                              |                          | 0              |                                     | 13874 14677         | 804                                      |                          | 0              |                                     | 13857 14662         | 806                              |                          |                                             | 0                                   |  |  |  |  |  |
| control region       | L                                         | 14654 15796         | 1143                             |                          | 0              |                                     | 14675 15805         | 1131                                     |                          |                |                                     | 14660 15774         | 1115                             |                          |                                             | -3                                  |  |  |  |  |  |

\* T<sub>1</sub> represents incomplete stop codons. † Intergenic bp indicates gap nucleotides (positive value) or overlapped nucleotides (negative value) between two adjacent genes. ‡ H and L indicate genes transcribed on the heavy and light strands, respectively.

\* T, represents incomplete stop codons. † Intergenic bp indicates gap nucleotides (positive value) or overlapped nucleotides (negative value) between two adjacent genes. ‡ H and L indicate genes transcribed on the heavy and light strands, respectively.

### Protein-coding genes and codon usage

The length of the total PCGs is 11,133 bp for *A. horsfieldi*, 11,130 bp for *A. chinensis* and 11,092 bp for *A. glabripennis*, accounting for 70.48%, 70.42% and 70.25% of their total mitogenomes, respectively (Table II). Four PCGs (ND1, ND4, ND4L and ND5) are from N strand, and the remaining genes are found on J strand (Fig. 1 and Table I). Thirteen PCGs showed variable range from 156 bp (ATP8) to 1711/1714 bp (ND5) in *A. horsfieldi*, *A. chinensis* and *A. glabripennis* (Table I). All three mitogenomes showed similar characteristics with the smallest size of ATP8 and the largest that of ND5 among PCGs (Fang *et al.*, 2016; Li *et al.*, 2016). In addition, their AT skewness and GC skewness are negative and positive, respectively which are not consistent with those of complete mitogenome (Table II).

Almost all PCGs of three species start with the ATN codon, only ND5 in *A. chinensis* starts with the GTG codon (Table II). In *A. horsfieldi*, an incomplete stop codon “T–” is used for COXI, COXII, COXIII, ND3, ND4 and ND5, while ND1 used TAG as stop codon, and the remaining five PCGs used TAA as stop codon (Table I). In *A. chinensis*, an incomplete stop codon “T” is used for COXI, COXII, ND3, ND4 and ND5, while ND1 and Cytb used TAG as a complete stop codon, and the remaining six PCGs used TAA as complete stop codon (Table I). Moreover, in *A. glabripennis* an incomplete stop codon “T–” is used for COXI, COXII, COXII, ND1, ND3, ND4 and ND5, while TAG is used as a complete stop codon for Cytb, and TAA is used as a complete stop codon for the remaining five PCGs (Table I). In all three mitochondrial genomes, the termination codon TAA appeared higher than TAG and at least five incomplete stop codons “T–” were present (Table I) (Fang *et al.*, 2016; Li *et al.*, 2016). The incomplete termination codons could be added as TAA by post-transcriptional polyadenylation during the mRNA process maturation (Ojala *et al.*, 1981; Schuster and Stern, 2009; Guindon *et al.*, 2010; Ronquist *et al.*, 2012). And this phenomenon has also been discovered in other insects as well (James and Andrew, 2006; Larkin *et al.*, 2007; Wang *et al.*, 2013; Wu *et al.*, 2014; Huang *et al.*, 2015; Liu *et al.*, 2018).

The similar amino acid base composition and the relative synonymous codon usage (RSCU) of the three *Anoplophora* species are found (Fig. 2 and Table III). The genome-wide bias towards AT was viewed to reflect in the codon usage by the PCGs. The total number of codons of the PCGs ranges from 3695 to 3709 (Table II). UUA (Leu2), AUU (Ile), UUU (Phe) and AUA (Met) are the most frequently utilized amino acids (Fig. 3), all high frequency codons are composed of A or U. The third codon frequency of A/T is significantly

higher than that of G/C, reflecting nucleotide A + T bias in the mitochondrial PCGs among Cucujiformia. The composition of most commonly used amino acids, like Leu, Ile, and Phe, varied from 42.47% (*A. glabripennis*) to 43.12% (*A. horsfieldi*) (Fig. 3). This pattern with rich with A or T nucleotides in all PCGs was also similar to other Cucujiformia insects (Kim *et al.*, 2009; Du *et al.*, 2016; Wang and Tang, 2018).

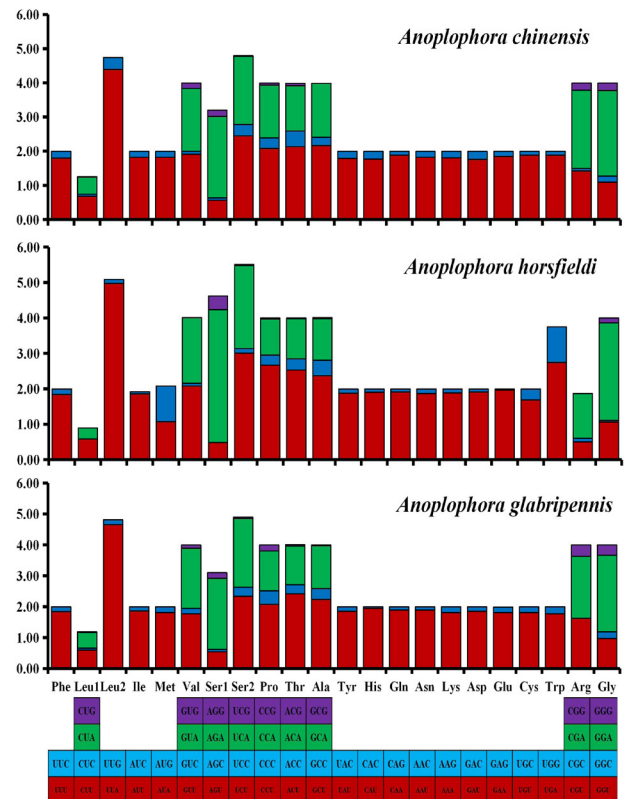


Fig. 2. The relative synonymous codon usage (RSCU) in the mitogenomes of three *Anoplophora* species. Codon families are provided on the x-axis along with the different combinations of synonymous codons that code for that amino acid. RSCU are provided on the y-axis.

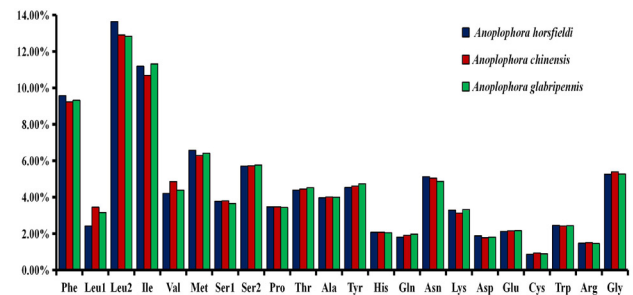


Fig. 3. Total codons in three *Anoplophora* mitogenomes.

**Table II. Nucleotide composition and skew values in *Anoplophora horsfieldi*, *Anoplophora chinensis* and *Anoplophora glabripennis*.**

|                                 | Size (bp) | A (bp) | T (bp) | G(bp) | C (bp) | T %   | C %   | A %   | G %   | A+T % | AT skew | GC skew |
|---------------------------------|-----------|--------|--------|-------|--------|-------|-------|-------|-------|-------|---------|---------|
| <i>Anoplophora horsfieldi</i>   |           |        |        |       |        |       |       |       |       |       |         |         |
| Whole genome                    | 15796     | 6277   | 6251   | 1298  | 1970   | 39.57 | 12.47 | 39.74 | 8.22  | 79.31 | 0.002   | -0.206  |
| 13 Protein-coding genes         | 11133     | 3730   | 4987   | 1220  | 1196   | 44.79 | 10.74 | 33.50 | 10.96 | 78.30 | -0.144  | 0.010   |
| 1 <sup>st</sup>                 | 3711      | 1274   | 1372   | 647   | 418    | 36.97 | 11.26 | 34.33 | 17.43 | 71.30 | -0.037  | 0.215   |
| 2 <sup>nd</sup>                 | 3711      | 781    | 1766   | 513   | 651    | 47.59 | 17.54 | 21.05 | 13.82 | 68.63 | -0.387  | -0.119  |
| 3 <sup>th</sup>                 | 3711      | 1675   | 1849   | 60    | 127    | 49.82 | 3.42  | 45.14 | 1.62  | 94.96 | -0.049  | -0.358  |
| Protein-coding genes-H strand   | 7773      | 2668   | 3288   | 816   | 1001   | 42.30 | 12.88 | 34.32 | 10.50 | 76.62 | -0.104  | -0.102  |
| 1 <sup>st</sup>                 | 2591      | 925    | 864    | 442   | 360    | 33.35 | 13.89 | 35.70 | 17.06 | 69.05 | 0.034   | 0.102   |
| 2 <sup>nd</sup>                 | 2591      | 611    | 1157   | 331   | 492    | 44.65 | 18.99 | 23.58 | 12.77 | 68.24 | -0.309  | -0.196  |
| 3 <sup>th</sup>                 | 2591      | 1132   | 1267   | 43    | 149    | 48.90 | 5.75  | 43.69 | 1.66  | 92.59 | -0.056  | -0.552  |
| Protein-coding genes-L strand   | 8016      | 2998   | 3563   | 690   | 765    | 44.45 | 9.54  | 37.40 | 8.61  | 81.85 | -0.086  | -0.052  |
| 1 <sup>st</sup>                 | 2672      | 982    | 1142   | 301   | 247    | 42.74 | 9.24  | 36.75 | 11.26 | 79.49 | -0.075  | 0.099   |
| 2 <sup>nd</sup>                 | 2672      | 829    | 1210   | 276   | 357    | 45.28 | 13.36 | 31.03 | 10.33 | 76.31 | -0.187  | -0.128  |
| 3 <sup>th</sup>                 | 2672      | 1187   | 1211   | 113   | 161    | 45.32 | 6.03  | 44.42 | 4.23  | 89.75 | -0.010  | -0.175  |
| tRNA genes                      | 1437      | 586    | 547    | 129   | 175    | 38.07 | 12.18 | 40.78 | 8.98  | 78.84 | 0.034   | -0.151  |
| rRNA genes                      | 2076      | 848    | 831    | 122   | 275    | 40.03 | 13.25 | 40.85 | 5.88  | 80.88 | 0.010   | -0.385  |
| Control region                  | 1143      | 502    | 486    | 35    | 120    | 42.52 | 10.50 | 43.92 | 3.06  | 86.44 | 0.016   | -0.548  |
| <i>Anoplophora chinensis</i>    |           |        |        |       |        |       |       |       |       |       |         |         |
| Whole genome                    | 15805     | 6240   | 6032   | 1388  | 2145   | 38.17 | 13.57 | 39.48 | 8.78  | 77.65 | 0.017   | -0.214  |
| 13 Protein-coding genes         | 11130     | 3655   | 4843   | 1316  | 1316   | 43.51 | 11.82 | 32.84 | 11.82 | 76.35 | -0.140  | 0.000   |
| 1 <sup>st</sup>                 | 3710      | 1238   | 1336   | 675   | 461    | 36.01 | 12.43 | 33.37 | 18.19 | 69.38 | -0.038  | 0.188   |
| 2 <sup>nd</sup>                 | 3710      | 776    | 1757   | 522   | 655    | 47.36 | 17.65 | 20.92 | 14.07 | 68.27 | -0.387  | -0.113  |
| 3 <sup>th</sup>                 | 3710      | 1641   | 1750   | 119   | 200    | 47.17 | 5.39  | 44.23 | 3.21  | 91.40 | -0.032  | -0.254  |
| Protein-coding genes-H strand   | 7785      | 2654   | 3157   | 857   | 1117   | 40.55 | 14.35 | 34.09 | 11.01 | 74.64 | -0.087  | -0.132  |
| 1 <sup>st</sup>                 | 2595      | 900    | 834    | 463   | 398    | 32.14 | 15.34 | 34.68 | 17.84 | 66.82 | 0.038   | 0.075   |
| 2 <sup>nd</sup>                 | 2595      | 611    | 1163   | 331   | 490    | 44.82 | 18.88 | 23.55 | 12.76 | 68.36 | -0.311  | -0.194  |
| 3 <sup>th</sup>                 | 2595      | 1143   | 1160   | 63    | 229    | 44.70 | 8.82  | 44.05 | 2.43  | 88.75 | -0.007  | -0.568  |
| Protein-coding genes-L strand   | 7998      | 2863   | 3571   | 1031  | 533    | 44.65 | 6.66  | 35.80 | 12.89 | 80.45 | -0.110  | 0.318   |
| 1 <sup>st</sup>                 | 2666      | 966    | 1115   | 413   | 172    | 44.65 | 6.66  | 35.80 | 12.89 | 80.45 | -0.110  | 0.318   |
| 2 <sup>nd</sup>                 | 2666      | 785    | 1213   | 386   | 282    | 45.50 | 10.58 | 29.44 | 14.48 | 74.94 | -0.214  | 0.156   |
| 3 <sup>th</sup>                 | 2666      | 1112   | 1243   | 232   | 79     | 46.62 | 2.96  | 41.71 | 8.70  | 88.33 | -0.056  | 0.492   |
| tRNA genes                      | 1446      | 579    | 539    | 181   | 147    | 37.28 | 10.17 | 40.04 | 12.52 | 77.32 | 0.036   | 0.104   |
| rRNA genes                      | 2076      | 810    | 851    | 287   | 128    | 40.99 | 6.17  | 39.02 | 13.82 | 80.01 | -0.025  | 0.383   |
| Control region                  | 1131      | 473    | 495    | 104   | 59     | 43.77 | 5.22  | 41.82 | 9.20  | 85.59 | -0.023  | 0.276   |
| <i>Anoplophora glabripennis</i> |           |        |        |       |        |       |       |       |       |       |         |         |
| Whole genome                    | 15768     | 6248   | 6104   | 1355  | 2061   | 38.71 | 13.07 | 39.62 | 8.59  | 78.34 | 0.012   | -0.207  |
| 13 Protein-coding genes         | 11092     | 3676   | 4867   | 1277  | 1272   | 43.88 | 11.47 | 33.14 | 11.51 | 77.02 | -0.139  | 0.002   |
| 1 <sup>st</sup>                 | 3698      | 1262   | 1336   | 653   | 447    | 36.13 | 12.09 | 34.13 | 17.66 | 70.25 | -0.028  | 0.187   |
| 2 <sup>nd</sup>                 | 3698      | 781    | 1752   | 509   | 656    | 47.38 | 17.74 | 21.12 | 13.76 | 68.50 | -0.383  | -0.126  |
| 3 <sup>th</sup>                 | 3696      | 1633   | 1779   | 115   | 169    | 48.13 | 4.57  | 44.18 | 3.11  | 92.32 | -0.043  | -0.190  |
| Protein-coding genes-H strand   | 8863      | 3153   | 3678   | 876   | 1156   | 41.50 | 13.04 | 35.57 | 9.88  | 77.07 | -0.077  | -0.138  |
| 1 <sup>st</sup>                 | 2954      | 1076   | 1004   | 451   | 423    | 33.99 | 14.32 | 36.43 | 15.27 | 70.41 | 0.035   | 0.032   |
| 2 <sup>nd</sup>                 | 2955      | 776    | 1328   | 341   | 510    | 44.94 | 17.26 | 26.26 | 11.54 | 71.20 | -0.262  | -0.199  |
| 3 <sup>th</sup>                 | 2954      | 1301   | 1346   | 84    | 223    | 45.57 | 7.55  | 44.04 | 2.84  | 89.61 | -0.017  | -0.453  |
| Protein-coding genes-L strand   | 6812      | 2391   | 3044   | 901   | 476    | 44.69 | 6.99  | 35.10 | 13.23 | 79.79 | -0.120  | 0.309   |
| 1 <sup>st</sup>                 | 2271      | 793    | 967    | 364   | 147    | 42.58 | 6.47  | 34.92 | 16.03 | 77.50 | -0.099  | 0.425   |
| 2 <sup>nd</sup>                 | 2271      | 635    | 1049   | 325   | 262    | 46.19 | 11.54 | 27.96 | 14.31 | 74.15 | -0.246  | 0.107   |
| 3 <sup>th</sup>                 | 2270      | 963    | 1028   | 212   | 67     | 45.29 | 2.95  | 42.42 | 9.34  | 87.71 | -0.033  | 0.520   |
| tRNA genes                      | 1431      | 572    | 545    | 173   | 141    | 38.09 | 9.85  | 39.97 | 12.09 | 78.06 | 0.024   | 0.102   |
| rRNA genes                      | 2039      | 794    | 833    | 285   | 127    | 40.85 | 6.23  | 38.94 | 13.98 | 79.79 | -0.024  | 0.383   |
| Control region                  | 1115      | 503    | 478    | 42    | 92     | 42.87 | 8.25  | 45.11 | 3.77  | 87.98 | 0.025   | -0.373  |

**Table III.** The codon number and relative synonymous codon usage in mitochondrial protein coding genes.

| Ami-<br>no<br>acid | Codon  | Count |     |     | RSCU |      |      | Frequency (%) |        |        | Ami-<br>no<br>acid | Codon  | Count |     |     | RSCU |      |      | Frequency (%) |       |       |
|--------------------|--------|-------|-----|-----|------|------|------|---------------|--------|--------|--------------------|--------|-------|-----|-----|------|------|------|---------------|-------|-------|
|                    |        | AH    | AC  | AG  | AH   | AC   | AG   | AH            | AC     | AG     |                    |        | AH    | AC  | AG  | AH   | AC   | AG   | AH            | AC    | AG    |
| Phe                | UUU(F) | 326   | 308 | 316 | 1.84 | 1.80 | 1.84 | 8.79%         | 8.32%  | 8.56%  | Tyr                | UAU(Y) | 158   | 153 | 162 | 1.88 | 1.79 | 1.85 | 4.27%         | 4.13% | 4.39% |
| Phe                | UUC(F) | 29    | 34  | 28  | 0.16 | 0.20 | 0.16 | 0.78%         | 0.92%  | 0.76%  | Tyr                | UAC(Y) | 10    | 18  | 13  | 0.12 | 0.21 | 0.15 | 0.27%         | 0.49% | 0.35% |
| Leu                | UUA(L) | 495   | 444 | 459 | 4.98 | 4.40 | 4.66 | 13.35%        | 11.99% | 12.44% | Stop               | UAA(*) | 6     | 6   | 5   | 0.19 | 1.50 | 1.67 | 0.16%         | 0.16% | 0.14% |
| Leu                | UUG(L) | 11    | 34  | 15  | 0.11 | 0.34 | 0.15 | 0.30%         | 0.92%  | 0.41%  | Stop               | UAG(*) | 2     | 2   | 1   | 0.06 | 0.50 | 0.33 | 0.05%         | 0.05% | 0.03% |
| Leu                | CUU(L) | 58    | 69  | 59  | 0.58 | 0.68 | 0.60 | 1.56%         | 1.86%  | 1.60%  | His                | CAU(H) | 73    | 68  | 74  | 1.90 | 1.77 | 1.95 | 1.97%         | 1.84% | 2.00% |
| Leu                | CUC(L) | 1     | 7   | 6   | 0.01 | 0.07 | 0.06 | 0.03%         | 0.19%  | 0.16%  | His                | CAC(H) | 4     | 9   | 2   | 0.10 | 0.23 | 0.05 | 0.11%         | 0.24% | 0.05% |
| Leu                | CUA(L) | 31    | 51  | 50  | 0.31 | 0.50 | 0.51 | 0.84%         | 1.38%  | 1.35%  | Gln                | CAA(Q) | 64    | 67  | 69  | 1.91 | 1.89 | 1.89 | 1.73%         | 1.81% | 1.87% |
| Leu                | CUG(L) | 0     | 1   | 2   | 0.00 | 0.01 | 0.02 | 0.00%         | 0.03%  | 0.05%  | Gln                | CAG(Q) | 3     | 4   | 4   | 0.09 | 0.11 | 0.11 | 0.08%         | 0.11% | 0.11% |
| Ile                | AUU(I) | 403   | 361 | 388 | 1.86 | 1.82 | 1.86 | 10.87%        | 9.75%  | 10.51% | Asn                | AAU(N) | 178   | 170 | 170 | 1.87 | 1.82 | 1.89 | 4.81%         | 4.59% | 4.61% |
| Ile                | AUC(I) | 12    | 35  | 30  | 0.06 | 0.18 | 0.14 | 0.32%         | 0.95%  | 0.81%  | Asn                | AAC(N) | 12    | 17  | 10  | 0.13 | 0.18 | 0.11 | 0.32%         | 0.46% | 0.27% |
| Ile                | AUA(M) | 234   | 212 | 216 | 1.08 | 1.82 | 1.82 | 6.31%         | 5.73%  | 5.85%  | Lys                | AAA(K) | 115   | 105 | 112 | 1.89 | 1.81 | 1.82 | 3.11%         | 2.84% | 3.03% |
| Met                | AUG(M) | 10    | 21  | 21  | 1.00 | 0.18 | 0.18 | 0.27%         | 0.57%  | 0.57%  | Lys                | AAG(K) | 7     | 11  | 11  | 0.11 | 0.19 | 0.18 | 0.19%         | 0.30% | 0.30% |
| Val                | GUU(V) | 81    | 86  | 72  | 2.08 | 1.91 | 1.78 | 2.19%         | 2.32%  | 1.95%  | Asp                | GAU(D) | 67    | 58  | 62  | 1.91 | 1.76 | 1.85 | 1.81%         | 1.57% | 1.68% |
| Val                | GUC(V) | 3     | 4   | 7   | 0.08 | 0.09 | 0.17 | 0.08%         | 0.11%  | 0.19%  | Asp                | GAC(D) | 3     | 8   | 5   | 0.09 | 0.24 | 0.15 | 0.08%         | 0.22% | 0.14% |
| Val                | GUA(V) | 72    | 83  | 79  | 1.85 | 1.84 | 1.95 | 1.94%         | 2.24%  | 2.14%  | Glu                | GAA(E) | 78    | 74  | 73  | 1.97 | 1.85 | 1.82 | 2.11%         | 2.00% | 1.98% |
| Val                | GUG(V) | 0     | 7   | 4   | 0.00 | 0.16 | 0.10 | 0.00%         | 0.19%  | 0.11%  | Glu                | GAG(E) | 1     | 6   | 7   | 0.03 | 0.15 | 0.17 | 0.03%         | 0.16% | 0.19% |
| Ser                | UCU(S) | 116   | 108 | 102 | 3.01 | 2.45 | 2.34 | 3.13%         | 2.92%  | 2.76%  | Cys                | UGU(C) | 27    | 33  | 30  | 1.69 | 1.89 | 1.82 | 0.73%         | 0.89% | 0.81% |
| Ser                | UCC(S) | 5     | 15  | 13  | 0.13 | 0.34 | 0.30 | 0.13%         | 0.41%  | 0.35%  | Cys                | UGC(C) | 5     | 2   | 3   | 0.31 | 0.11 | 0.18 | 0.14%         | 0.05% | 0.08% |
| Ser                | UCA(S) | 90    | 88  | 96  | 2.34 | 1.99 | 2.21 | 2.43%         | 2.38%  | 2.60%  | Stop               | UGA(W) | 87    | 85  | 80  | 2.75 | 1.89 | 1.78 | 2.35%         | 2.30% | 2.17% |
| Ser                | UCG(S) | 1     | 1   | 2   | 0.03 | 0.02 | 0.05 | 0.03%         | 0.03%  | 0.05%  | Trp                | UGG(W) | 4     | 5   | 10  | 1.00 | 0.11 | 0.22 | 0.11%         | 0.14% | 0.27% |
| Pro                | CCU(P) | 86    | 67  | 66  | 2.67 | 2.08 | 2.08 | 2.32%         | 1.81%  | 1.79%  | Arg                | CGU(R) | 15    | 20  | 22  | 0.51 | 1.43 | 1.63 | 0.41%         | 0.54% | 0.60% |
| Pro                | CCC(P) | 9     | 10  | 14  | 0.28 | 0.31 | 0.44 | 0.24%         | 0.27%  | 0.38%  | Arg                | CGC(R) | 3     | 1   | 0   | 0.10 | 0.07 | 0.00 | 0.08%         | 0.03% | 0.00% |
| Pro                | CCA(P) | 33    | 50  | 41  | 1.02 | 1.55 | 1.29 | 0.89%         | 1.35%  | 1.11%  | Arg                | CGA(R) | 37    | 32  | 27  | 1.26 | 2.29 | 2.00 | 1.00%         | 0.86% | 0.73% |
| Pro                | CCG(P) | 1     | 2   | 6   | 0.03 | 0.06 | 0.19 | 0.03%         | 0.05%  | 0.16%  | Arg                | CGG(R) | 0     | 3   | 5   | 0.00 | 0.21 | 0.37 | 0.00%         | 0.08% | 0.14% |
| Thr                | ACU(T) | 103   | 88  | 101 | 2.53 | 2.13 | 2.42 | 2.78%         | 2.38%  | 2.74%  | Ser                | AGU(S) | 19    | 25  | 24  | 0.49 | 0.57 | 0.55 | 0.51%         | 0.68% | 0.65% |
| Thr                | ACC(T) | 13    | 19  | 12  | 0.32 | 0.46 | 0.29 | 0.35%         | 0.51%  | 0.33%  | Ser                | AGC(S) | 0     | 3   | 3   | 0.00 | 0.07 | 0.07 | 0.00%         | 0.08% | 0.08% |
| Thr                | ACA(T) | 46    | 55  | 52  | 1.13 | 1.33 | 1.25 | 1.24%         | 1.49%  | 1.41%  | Arg                | AGA(S) | 110   | 105 | 100 | 3.75 | 2.38 | 2.30 | 2.97%         | 2.84% | 2.71% |
| Thr                | ACG(T) | 1     | 3   | 2   | 0.02 | 0.07 | 0.05 | 0.03%         | 0.08%  | 0.05%  | Arg                | AGG(S) | 11    | 8   | 8   | 0.38 | 0.18 | 0.18 | 0.30%         | 0.22% | 0.22% |
| Ala                | GCU(A) | 87    | 81  | 83  | 2.37 | 2.17 | 2.24 | 2.35%         | 2.19%  | 2.25%  | Gly                | GGU(G) | 52    | 55  | 48  | 1.07 | 1.10 | 0.98 | 1.40%         | 1.49% | 1.30% |
| Ala                | GCC(A) | 16    | 9   | 13  | 0.44 | 0.24 | 0.35 | 0.43%         | 0.24%  | 0.35%  | Gly                | GGC(G) | 2     | 9   | 10  | 0.04 | 0.18 | 0.21 | 0.05%         | 0.24% | 0.27% |
| Ala                | GCA(A) | 43    | 59  | 51  | 1.17 | 1.58 | 1.38 | 1.16%         | 1.59%  | 1.38%  | Gly                | GGA(G) | 134   | 125 | 121 | 2.75 | 2.50 | 2.48 | 3.62%         | 3.38% | 3.28% |
| Ala                | GCG(A) | 1     | 0   | 1   | 0.03 | 0.00 | 0.03 | 0.03%         | 0.00%  | 0.03%  | Gly                | GGG(G) | 7     | 11  | 16  | 0.14 | 0.22 | 0.33 | 0.19%         | 0.30% | 0.43% |

RSCU, relative synonymous codon usage; AH, *Anoplophora horsfieldi*; AC, *Anoplophora chinensis* (NC\_029230); AG, *Anoplophora glabripennis* (NC\_008221).

#### Ribosomal and transfer RNA genes

Twenty-two tRNAs of *A. horsfieldi*, *A. chinensis* and *A. glabripennis* mitogenomes scattered discontinuously across the complete mitogenome (Table I). The tRNAs length of these three species were 1,437, 1,446 and 1,431 bp, respectively, and accounted for 9.10%, 9.15% and 9.08% of the total mitogenomes, respectively (Table I). The three mitogenomes possess 22 tRNA genes, eight

transcribed from N strand and 14 from J strand (Fig. 1, Table I). The length of these 22 tRNAs range from 60 (trnR) to 76 bp (trnD) in *A. horsfieldi*, from 62 (trnC) to 84 bp (trnD) in *A. chinensis* and from 62 (trnH) to 74 bp (trnP) in *A. glabripennis* (Table I). 21 of 22 tRNAs can form typical cloverleaf secondary structure, while trnS1 (AGN) formed a simple loop due to lacking the DHU arm (Fig. 4, Supplementary Figs. S1-S2), as discovered in other

coleopteran species (Stewart and Beckenbach, 2003; Song *et al.*, 2010; Cabrera-Brandt and Gaitan-Espitia, 2015). The lack of DHU stem in trnS1 is a common phenomenon in insect mitogenomes (Beckenbach, 2011; Cameron, 2014; Li *et al.*, 2015), and has been proven as a typical characteristic of metazoan mitogenomes (Wang *et al.*, 2019). Many nucleotide substitutions are spotted in four different stems and the anticodon loop is highly conserved (Fig. 4). Besides the orthodox AU and CG pairs, some mismatched base pairs are also discovered in different stems. A total of 15 GU mismatches, 3 UU mismatches, and 1 AG mismatch are found in *A. horsfieldi* (Fig. 4), 17 GU mismatches, 2 UU mismatches, 2 AG mismatches, and 1 AA mismatch are found in *A. glabripennis* while 17 GU mismatches, 2 UU mismatches, 1 AG mismatch, and 1 AC mismatches are found in *A. chinensis* (Fig. 4, Supplementary Figs. S1-S2, Table I).

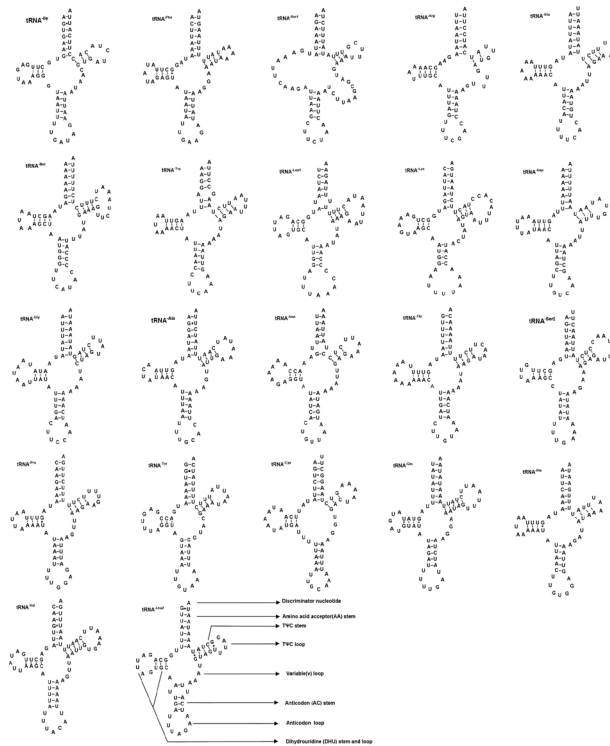


Fig. 4. Putative tRNA secondary structures predicted from the 22 tRNA gene sequences found in the *Anoplophora horsfieldi* mitogenome.

Two rRNA genes (rrnL and rrnS) are encoded by the N-strand in *A. horsfieldi*, *A. chinensis* and *A. glabripennis* and rrnL is distributed between trnL1 and trnV, while rrnS lies between trnV and CR. The genes rrnL and rrnS are 1273 and 804 bp long in *A. horsfieldi* and *A. chinensis*, and 1236 and 806 bp long in *A. glabripennis*, respectively

(Table I). The lengths range from 1236 to 1273 bp in rrnL, and from 804 to 806 bp in rrnS in these three mitogenomes of genus *Anoplophora* (Fang *et al.*, 2016; Li *et al.*, 2016). The A+T content of the two rRNAs (80.88%) in *A. horsfieldi* was higher than 80.01% in *A. chinensis* and 79.79% in *A. glabripennis*, with positive AT skew (0.010) and negative GC skew (-0.385), however, in the other two species, the AT and GC skews are negative and positive, respectively (Table III).

#### Overlapping and intergenic spacer regions

A total of 15 gene overlap in the *A. horsfieldi* mitogenome from 1 to 13 bp, adding up to 59 bp, 14 gene overlaps exist in the *A. chinensis* mitogenome with the length from 1 to 8 bp, amounting to 45 bp, while 18 gene overlaps occur in the *A. glabripennis* mitogenome with the length from 1 to 9 bp, adding up to 64 bp (Table I). The longest overlap region (13 bp) of the three mitogenomes is found between trnK and trnD. All three *Anoplophora* species have eight identical overlap regions, including trnQ-trnM (1 bp), ATP6-COXIII (1 bp), trnE-trnF (1 bp), ND6-Cytb (1 bp), Cytb-trnS (2 bp), ND4-ND4L (4 bp), ATP8-ATP6 (7 bp), trnY-COXI (8 bp) and trnW-trnC (8 bp) (Coates, 2014; Fang *et al.*, 2016; Li *et al.*, 2016; Yao *et al.*, 2017). Eight intergenic spacers appear in *A. horsfieldi*, ranging in length from 2 to 17 bp and amounting to 42 bp. Nine intergenic spacers occur in *A. chinensis*, ranging in size from 1 to 17 bp and adding up to 41 bp. A total of 115 bp in *A. glabripennis* distributed in 12 intergenic spacers, ranging in size from 1 to 35 bp. All three mitochondrial genomes share four identical intergenic spacers (trnR-trnN (2 bp), ND4L-trnT (2 bp), ND5-trnS (3 bp), trnS-ND1 (17 bp)) (Table I). The size of the intergenic spacers is more variable than overlaps.

Two 7-bp long overlaps (ATGATAA) were discovered in the *Anoplophora* species, which were also observed in other Polyphaga insects (Fig. 5), however, 7-bp long overlaps (ATGATTA) was discovered in *Epicauta chinensis* (KP692789) and 7-bp long overlaps (ATGATAG) was found in *Tribolium castaneum* (AJ312413) (Friedrich and Muqim, 2003) and *Cryptolestes pusillus* (NC\_028204) (Li *et al.*, 2016) between ATP8 and ATP6. The overlaps lie between ATP8 and ATP6 on the J-strand and between ND4L and ND4 on the N-strand, respectively. The overlapped sequences were translated as a bicstron (Stewart and Beckenbach, 2005; Wang and Tang, 2018). The other two 5 bp (TTAAT) and 7 bp (TTTAGT) long motif were detected between trnSer (UCN) and ND1 in mitogenomes of the three *Anoplophora* species, which was also present in other cucujiform beetle (Fig. 6). However, another 5 bp long motif (TAGTA) was found at this location (Wang and Tang, 2018). This has been explained as the possible

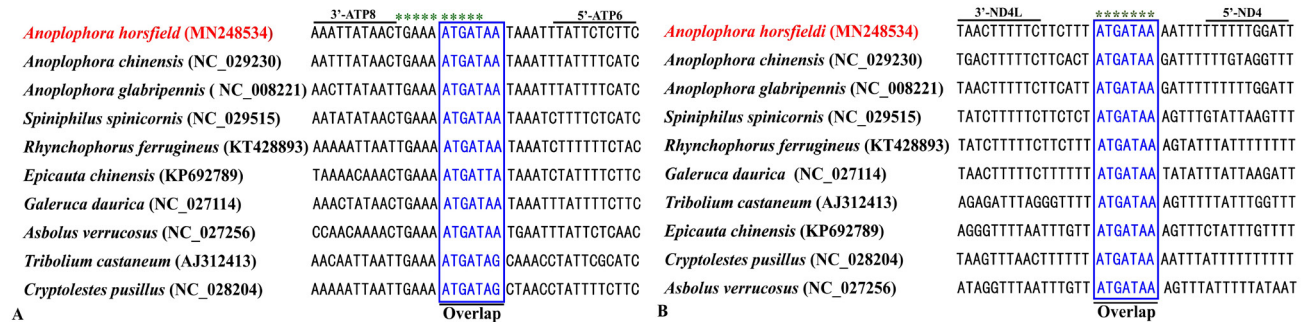


Fig. 5. Alignment of overlapping region across Cucujiformia and other insects. (A) Atp8 and Atp6, the motifs “ATGATA/TA/G” in the box are conservative; (B) ND4L and ND4, the motifs “ATGATAA” in the box are conservative.

binding site of mtTERM because it is located at the end of the H-strand region in the circular mitochondrial genome (Taanman, 1999). Based on mitogenomic comparisons, the motifs between ATP8 and ATP6, and between trnSer and ND1 were relatively conserved in four suborders (Polyphaga, Adephaga, Archostemata and Myxophaga). However, the motif “ATGATAA” was only found between ND4 and ND4L of the mitogenomes of Polyphaga insects, while “ATGTAA” was found in the other three suborders’ mitogenomes (Adephaga, Archostemata and Myxophaga) (Wang and Tang, 2018).

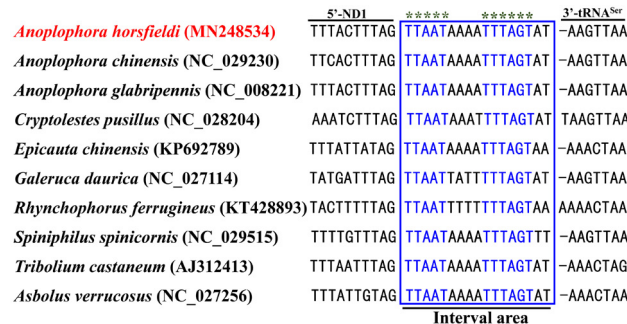


Fig. 6. Alignment of the intergenic spacer region between trnS (UCN) and ND1 of several Cucujiformia insects. The shaded “TTAAT” and “TTTAGT” motif is conserved across the Cucujiformia species.

#### A+T-rich region

The A+T-rich region in insect mitochondrial genome, similar to the control region of vertebrate mitochondrial genome, possess the transcription and replication origin sites (Andrews *et al.*, 1999; Yukuhiro *et al.*, 2002; Zhu *et al.*, 2013). The largest A+T-rich region lied between the genes *rrnS* and *trnI*. This region plays important roles in regulating transcription and replication (Zhang and Hewitt, 1997). The lengths of A+T-rich region are 1,143 bp in *A. horsfieldi*. In all three *Anoplophora* species, *A. glabripennis* has the smallest

control region of 1,115 bp, while *A. horsfieldi* has the largest that of 1,143 bp (Fig. 7, Table I). The A + T contents are 86.44% in *A. horsfieldi*, 85.59% in *A. chinensis* and 87.98% in *A. glabripennis*. The motif of tandem repeats in A+T-rich region has been recorded in many insect orders (Zhang and Hewitt, 1997; Du *et al.*, 2017; Wang *et al.*, 2018; Huang *et al.*, 2019; Li *et al.*, 2019), but is found in *A. horsfieldi* for the first time. *A. horsfieldi* and *A. glabripennis* have three types of long repeat tandem units and one types of microsatellite-like repeat, while there are only two different long repeats in *A. chinensis* (Fig. 7B). The repeat units presented obvious differences in size and copy number among Coleoptera species, thus resulting in various lengths of the A+T-rich region.

Tandem repeats (due to difference in type and number) can be used as microsatellites for phylogeography and population genetics researches and also as a molecular marker in evolutionary biology (Wang *et al.*, 2018; Zhang *et al.*, 2019). In addition, in some insect species (Kim *et al.*, 2006; Yin *et al.*, 2012), there was no long tandem repetitive sequence in this region, but it did have some microsatellite-like repeats (such as, (A)<sub>9</sub>, (T)<sub>15</sub>, (C)<sub>8</sub>, (TA)<sub>6</sub>, (TA)<sub>8</sub>). The A+T-rich region in *A. horsfieldi* possess a 16 bp long poly-T and 14 bp long poly-A stretches which has been considered as a possible recognition site for the initiation of replication of the mtDNA minor strand (Fig. 7A, Andrews *et al.*, 1999; Kim *et al.*, 2009; Du *et al.*, 2016). When compared with the A+T-rich region of three *Anoplophora* species, three conserved sequence blocks (CSBs) were recognized in those species. These CSBs ranged in length from 23 to 111 bp, and their sequence similarity among species was generally more than 50%. These CSBs were also found in other stoneflies mitogenomes (Qian *et al.*, 2014; Chen and Du, 2015, 2017; Cao *et al.*, 2019), but the functions and characteristics of these conserved blocks need to be further studied. In *A. horsfieldi* and *A. glabripennis*, the AT and GC skewness of the A+T-rich region are negative and positive, respectively, while they are opposite in *A. chinensis* (Table II).

*rrnS*-14,654-AAAAAAAAAAGAACTACTCCCCCTATTAAAAATATTCAAACCTCATTTTTTTAGCACTAA  
 AACTTTCTATATAAAATTATTTCAATATCCCCTAAATACTGATTTTCATAAACTTCCATAAAAC  
 TCATTTATAAAATTATTTCAATATCCCCTAAATACTGATTTTCATAAACTTCCATAAAACTCAT  
 TTATAAAATTATTTCAATATCCCCTAAATACTGATTTTCATAAACTTCCATAAAACTCATTAT  
 AAAATTATTTCAATATCCCCTAAATACTGATTTTCATAAACTTCCATAAAACTCATTATATAAA  
 TTATTTCAATATCCCCTAAATACTGATTTTCATAAACTTCCATAAAACTCATTATATAAAATTAT  
 TTCAATATCCCCTAAATACTGATTTTCATAAACTTCCATTTAAATATAAT<sup>CSB1</sup>AATAATTTTATAAA  
CATATTATTATATAAAATTTAATATTTTATATAATTAAAAATAATTATTTAATAAAATTATAATTTT  
 ATATTATATCATTTTTTTATACTAATATAAAAAAAGGGTAGTTTTTTTTTTTTTTTTAAAAATTGTT  
 ATTTTCAATAACAAATTTATATATTAATAATATATAAATTATTTATATATATAACAATTATATTT  
 ATATATATTACTAAATATATTATATATATAAGTAT<sup>CSB2</sup>ATATTATTTTTATATTTATATTAAATTGTT  
ATATATATATAAGATATTTATATATCTATAATAATATAATATAAATGTATATATATAAATAATAAA  
ATAAATAATATTGATTCATTATAGAAATAATTAATACTAATATAAATTTCTATTATCAATTTTATTT  
 TTTTCATTATTGATTTATTTAAAAATAGTATTATTCCTAAAATTTTATTAATATCTGATAAAAAATA  
 GGAGATCCTTATATTAATATAAATAGCTTATATAAATATATAAACTGATTAGGCATATATTAATAT  
 AATTTTTATTTTAAACAAAAAAT<sup>CSB3</sup>AAAAAAAAAAAAAATGTAAAAAATAACAATCATTGCATTTTCT  
 TTTAGTTAATTTTT<sup>CSB3</sup>TTTCGCTAAAACCCAATTTTACAAAAATTTACATTAAATTCATGTTTATTTA  
 ATTTAATAACTAAAAAAGTTAT-15,796-*trnI*

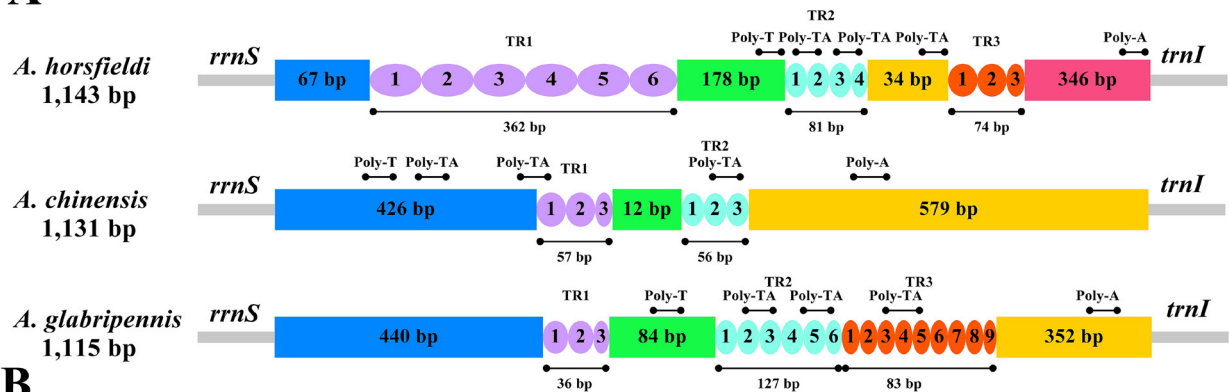
**A****B**

Fig. 7. (A) Organization of the A+T-rich regions in the *Anoplophora horsfieldi* mitogenomes. The square box represents the feature sequence CSB1-3. The poly-T stretch is underlined while the poly-A stretch is double underlined. The single microsatellite T/A repeats sequence are indicated by single underlining. (B) The colored ovals indicate the tandem repeats; the remaining regions are shown with different color boxes. TR1-3 represents different types of repeat sequences. Horizontal lines with solid dots at both ends represent the structures of poly (A), poly (T) and poly (TA), respectively.

#### Phylogenetic analyses

The phylogenetic analyses were carried out using the concatenated nucleotide sequences of 13 PCGs from 117 Cucujiformia mitogenomes (Supplementary Table S1). Similar topologies were achieved by using Bayesian (BI) and maximum likelihood (ML) analyses (Fig. 8). Phylogenetic analysis showed that *A. horsfieldi* grouped with *A. glabripennis* and *A. chinensis* with high nodal

supports. In *Anoplophora*, *A. horsfieldi* is at the base of the genus. In our phylogenetic analyses, *Anoplophora* was clustered with *Monochamus*, rather than genus *Psacotheta* and it was consistent with the results of previous phylogenetic analyses based on mitochondrial data (Coates, 2014; Yuan *et al.*, 2016; Liu *et al.*, 2018; Wang *et al.*, 2019). In addition to *Orsodacne lineola* (Vesperidae), Prioninae, Cerambycinae, Disteniinae, Cassidinae and

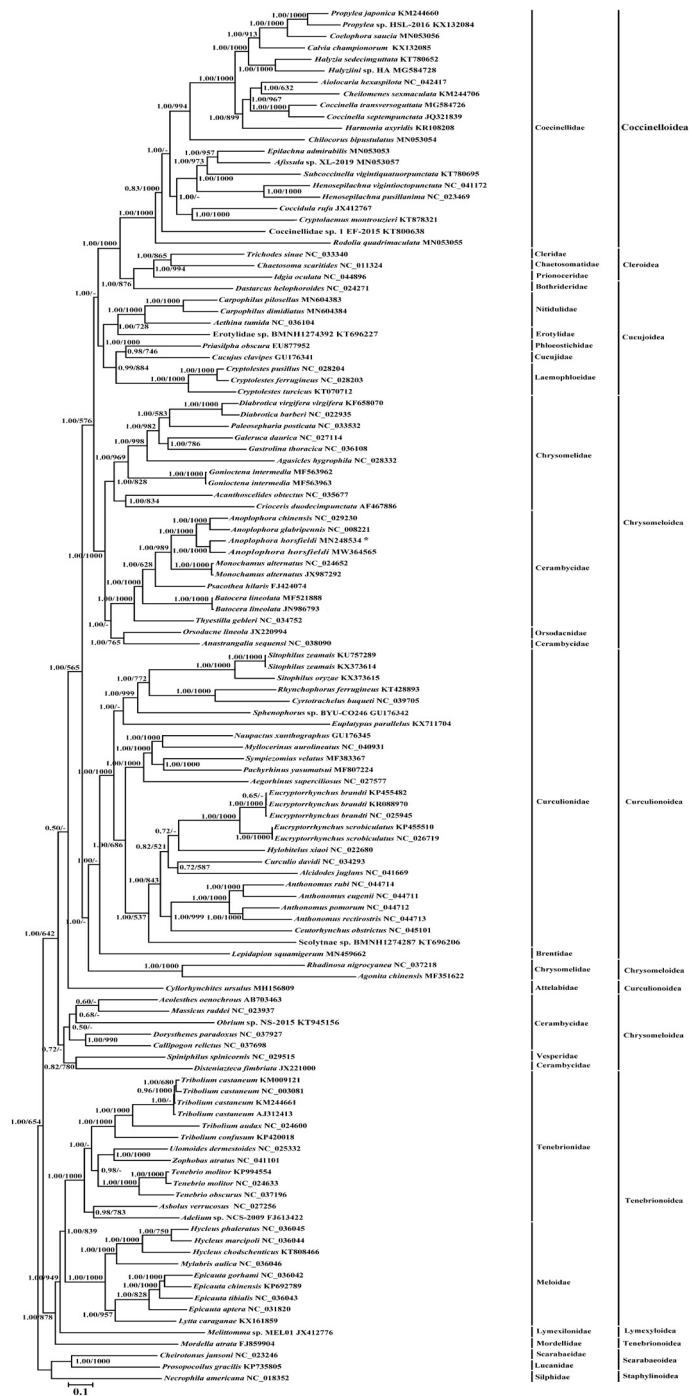


Fig. 8. Phylogenetic tree of the relationships among 117 mitogenome sequences of Cucujiformia based on the nucleotide dataset of the 13 mitochondrial protein-coding genes of 11,154 nucleotides with *Prosopocoilus gracilis* (Coleoptera: Scarabaeiformia: Lucanidae: KP735805), *Cheirotonus jansoni* (Coleoptera: Scarabaeiformia: Scarabaeidae: NC\_023246) and *Necrophila americana* (Coleoptera: Staphyliniformia: Silphidae: NC\_018352) as outgroups. Branch lengths and topology are from the BI analysis. Numbers above branches specify posterior probabilities from Bayesian inference (BI) and bootstrap percentages from maximum likelihood (ML, 1000 replications) analyses. The GenBank numbers and scientific names of all species are shown in [Supplementary Table S1](#). Tree topologies produced by Bayesian inferences (BI) and maximum likelihood (ML) analyses were equivalent. Bayesian posterior probability and bootstrap support values for ML analyses are shown orderly on the nodes. The asterisks indicate new sequences generated in this study.

Philinae species, Cerambycidae is closely related with family Chrysomelidae, and *Anastrangalia sequensi* was positioned as the basal lineage of Cerambycidae (Fig. 8). At the same time, the phylogenetic topological structure supports the two families as monophyletic. However, comparing our results to previous morphological and molecular analyses of Cerambycidae and Chrysomelidae relationships was harder to determine (Marvaldi et al., 2009; Yuan et al., 2016; Liu et al., 2018). Previous phylogenetic analyses of molecular data resulted in a monophyletic Cerambycidae and Chrysomelidae (Zhou et al., 2016; Liu et al., 2017; Yao et al., 2017). This may be because of the limited mitogenomic data, especially for Chrysomelidae species. Based on phylogenetic tree, the monophyly of Chrysomelidae and Cerambycidae cannot be supported in Chrysomeloidea, and their evolutionary relationships are relatively complex (Fig. 8) and need to be confirmed by additional samples and species and phylogenetic information, including nuclear genes.

Previously, the series Cucujiformia insects contain six superfamilies (Bouchard et al., 2011; Lawrence et al., 2011). However, our results support that Cucujiformia is divided into seven superfamilies, which is in agreement with the views of Robertson et al. (2015) and Yuan et al. (2016). Although the topological structures of phylogenetic trees are different among different data strings and analysis methods, these studies do not support the previous classification of Cucujoidea as monophyletic (Hunt et al., 2007; Marvaldi et al., 2009; Lawrence et al., 2011; Bocak et al., 2014; Robertson et al., 2015; Timmermans et al., 2016), which is consistent with our phylogenetic results (Fig. 8), and found that Elateroidea is embedded between Coccinelloidea and Cleroidea classification. Recently, Robertson et al. (2015) constructed a high-level phylogenetic analysis for Cucujoidea according to four mitochondrial and four nuclear genes, where two superfamilies (Cleroidea and Cucujoidea) were redefined and a new superfamily Coccinelloidea (cerylonid series) was proposed, and our reconstructed phylogenetic results also support the Coccinelloidea as a new superfamily (Fig. 8). In these seven superfamilies, except for *Dastarcus helophoroides* and *Cyllorhynchites ursulus*, five superfamilies (Lymexyloidea, Cucujoidea, Coccinelloidea, Cleroidea and Curculionoidea) are consistently supported as monophyletic with ML and BI methods, which is in accord with the results of Timmermans et al. (2016) and Yuan et al. (2016) (Fig. 8). Although the monophyly of other two superfamilies (Chrysomeloidea and Coccinelloidea) are recovered in PhyloBayes analyses with the P123R dataset, our results support that Chrysomeloidea and Tenebrionoidea belong to paraphyletic relationships based on 13 PCGs (Fig. 8). Yuan et al. (2016) carried

out the PhyloBayes analyses with the P123R dataset, and formed such a cucujiform relationship of (Cleroidea + (Coccinelloidea + ((Lymexyloidea + Tenebrionoidea) + (Cucujoidea + (Chrysomeloidea + Curculionoidea)))) (Fig. 8), however, our topology tree supports such phylogenetic relationships of ((Lymexyloidea + Tenebrionoidea) + (Curculionoidea + (Chrysomeloidea + (Cucujoidea + (Cleroidea + Coccinelloidea)))) (Fig. 8). This is different from the studies of (((Chrysomeloidea + Cucujoidea) + Tenebrionoidea) + Cleroidea) obtained by Kim et al. (2009), (((Curculionoidea + Chrysomeloidea) + Cucujoidea) + (Cucujoidea + Cleroidea)) Tenebrionoidea in the PCG-BI tree obtained by Li et al. (2016), (((Curculionoidea + Chrysomeloidea) + Cucujoidea) + (Cleroidea + (cerylonid lineages + (Lymexyloidea + Tenebrionoidea))) constructed by Timmermans et al. (2016) and (Cleroidea + (Erotylid series + (Tenebrionoidea + (Cucujoidea + (Chrysomeloidea + Curculionoidea)))) of Crampton-Platt et al. (2015) where these superfamily relationships within Cucujiformia cannot be well resolved. Lymexyloidea + Tenebrionoidea were positioned as the basal lineage of the series Cucujiformia insects in the PCG-BI/ML tree (Fig. 8), which was inconsistent with previous reports (Coates, 2014; Li et al., 2016), whereas superfamily Tenebrionoidea is located at the basal position in the ML and BI trees, and is also different from the basal position of Cleroidea in Cucujiformia (Kim et al., 2009; Marvaldi et al., 2009). These differences may be due to the use of different numbers of samples, molecular markers and analytical methods.

From the view of topology tree, the layout of Chrysomeloidea species is complicated and paraphyletic, which is consistent with Bocak et al. (2014). Relationships of some families in Chrysomeloidea remain to be poorly supported, especially for the small families Orsodacnidae and Vesperidae. Besides Prioninae, Cerambycinae, Disteniinae, Cassidinae and Philinae, the phylogenetic relationship of the family Cerambycidae is monophyletic. Unsatisfactory results were also found for internal relationships of Cerambycidae for which the sampling number and density of species were low and especially the representativeness of conservative loci was poor (Bocak et al., 2014). The Prioninae, Cerambycinae, Disteniinae, Cassidinae and Philinae species are not clustered in the family Chrysomelidae, which leads to the formation of paraphyly of the family. Besides the five subfamilies, Chrysomelidae were well supported and this is in accord with previous study results from Gomez-Zurita et al. (2007) and Bocak et al. (2014). The major clades were resolved (Fig. 8), containing the basal splits of a group of Donaciinae, Criocerinae and Bruchinae, the monophyly of Cassidinae. However, our results support the paraphyly of

Galerucinae, Chrysomelinae and Cerambycinae (Fig. 8). The phylogenetic relationships at subfamily Prioninae, Cerambycinae, Disteniinae, Cassidinae and Philinae, especially in family Vesperidae and Orsodacnidae, still have not been resolved due to the lack of sufficient phylogenetic information and many more new species.

In word, as far as Cucujiformia, Chrysomeloidea and Chrysomelidae is concerned, the conflict results about phylogeny might be due to the different types of molecular markers and analytical methods (Kim *et al.*, 2009; Marvaldi *et al.*, 2009; Coates, 2014; Crampton-Platt *et al.*, 2015; Meiklejohn *et al.*, 2014). Therefore, the phylogenetic relationship and taxonomic status within the Cucujiformia insects require a large number of additional samples and more molecular markers to elucidate their evolutionary relationship. Undoubtedly, this study will contribute to explore the phylogeny, systematics and taxonomy of the Cucujiformia insects.

## CONCLUSION

In this study, the entire mitogenome of *A. horsfieldi* (16,759 bp) was sequenced and annotated. The mitochondrial gene order and orientation of *A. horsfieldi* were similar to Chrysomeloidea species. The motifs, 'ATGATAA' between ND4L and ND4, was more conserved than that between trnS (UCN) and ND1 and between ATP8 and ATP6 in Cucujiformia. Phylogenetic tree showed that *A. horsfieldi* grouped with *A. glabripennis* and *A. chinensis* with high nodal supports. Within Cucujiformia, such phylogenetic relationships of (Lymexyloidea + (Tenebrionoidea + (Curculionoidea + (Chrysomeloidea + (Cucujoidea + (Clerioidea + Coccinelloidea)))))) were rebuilt, and the paraphyletic relationships that Chrysomeloidea, Coccinelloidea and Curculionoidea were supported. We believe that the mitogenomes of *A. horsfieldi* will help people to further the studies of molecular evolution, and Cucujiformia mitogenome architecture and phylogenetics.

## ACKNOWLEDGEMENT

We would like to express our gratitude to all those who helped us during the writing of this manuscript.

### Funding

This research was supported by the Natural Science Foundation Project of Science and Technology, Department of Sichuan Province (No. 2023NSFSC0206), the National Natural Science Foundation of China (No. 32270457), the Research Project of Ecological Security and Protection Key Laboratory of Sichuan Province (No. ESP2003).

### IRB approval

The animal handling procedures conformed with the China Animal Welfare guidelines and have been approved by the Ethic and Animal Welfare Committee and the Animal Protection and Use Commission of Mianyang Normal University (MYSY2017JC02).

### Ethical statement

This study was carried out in accordance with the Animal Care and Use Committee at the Mianyang Normal University. Efforts were taken to minimize suffering and included administering anesthesia. The study did not involve endangered or protected species.

### Supplementary material

There is supplementary material associated with this article. Access the material online at: <https://dx.doi.org/10.17582/journal.pjz/20220824100800>

### Statement of conflict of interest

The authors have declared no conflict of interest.

## REFERENCES

- Andrews, R., Kubacka, M.I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M. and Howell, N., 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.*, **23**: 147. <https://doi.org/10.1038/13779>
- Beckenbach, A.T., 2011. Mitochondrial genome sequences of representatives of three families of scorpionflies (Order Mecoptera) and evolution in a major duplication of coding sequence. *Genome*, **54**: 368-376. <https://doi.org/10.1139/g11-006>
- Benson, G., 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucl. Acids Res.*, **27**: 573-580. <https://doi.org/10.1093/nar/27.2.573>
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsche, G., Pütz, J., Middendorf, M. and Stadler, P.F., 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.*, **69**: 313-319. <https://doi.org/10.1016/j.ympev.2012.08.023>
- Bocak, C., Barton, L., Crampton-Platt, A., Chesters, D., Ahrens, D. and Vogler, A.P., 2014. Building the coleoptera tree-of-life for > 8000 species: Composition of public dna data and fit with linnaean classification. *Syst. Ent.*, **39**: 97-110. <https://doi.org/10.1111/syen.12037>
- Bouchard, P., Bousquet, Y., Davies, A.E., Alonso-Zarazaga, M.A., Lawrence, J.F., Lyal, C.H.C.,

- Newton, A., Reid, C., Schmitt, M. and Slipinski, A., 2011. Family-group names in Coleoptera (Insecta). *Zookeys*, **88**: 1-972. <https://doi.org/10.3897/zookeys.88.807>
- Bouchard, P., Grebennikov, V.V., Smith, A.B.T. and Douglas, H., 2009. Chapter 11. Biodiversity of coleoptera. In: *Insect biodiversity: Science and society* (eds. R.G. Foottit and P.H. Adler). Blackwell. Publishing. Ltd., Oxford
- Cabrera-Brandt, M.A. and Gaitan-Espitia, J.D., 2015. Phylogenetic analysis of the complete mitogenome sequence of the raspberry weevil, *Aegorhinus superciliosus* (Coleoptera: Curculionidae), supports monophyly of the tribe Aterpini. *Gene*, **571**: 205-211. <https://doi.org/10.1016/j.gene.2015.06.059>
- Cameron, S.L. and Whiting, M.F., 2008. 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*, **40**: 112-123. <https://doi.org/10.1016/j.gene.2007.10.023>
- Cameron, S.L., 2014. Insect mitochondrial genomics: Implications for evolution and phylogeny. *Annu. Rev. Ent.*, **59**: 95-117. <https://doi.org/10.1146/annurev-ento-011613-162007>
- Cao, J., Wang, Y., and Li, W., 2019. Comparative mitogenomic analysis of species in the subfamily Amphinemurinae (Plecoptera: Nemouridae) reveal conserved mitochondrial genome organization. *Int. J. Biol. Macromol.*, **138**: 292-301. <https://doi.org/10.1016/j.ijbiomac.2019.07.087>
- Chen, S., Xie, Y. and Deng, G., 1959. *Economic entomology of China: Volume 1. Coleoptera: Cerambycidae*. Science Press. Beijing, China.
- Chen, Z.T. and Du, Y.Z., 2015. Comparison of the complete mitochondrial genome of the stonefly *Sweltsa longistyla* (Plecoptera: Chloroperlidae) with mitogenomes of three other stoneflies. *Gene*, **558**: 82-87. <https://doi.org/10.1016/j.gene.2014.12.049>
- Chen, Z.T. and Du, Y.Z., 2017. First mitochondrial genome from Nemouridae (Plecoptera) reveals novel features of the elongated control region and phylogenetic implications. *Int. J. mol. Sci.*, **18**: 996. <https://doi.org/10.3390/ijms18050996>
- Christoph, H., Lutz, B. and Bastien, C., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucl. Acids Res.*, **41**: e129. <https://doi.org/10.1093/nar/gkt371>
- Coates, B.S., 2014. Assembly and annotation of full mitochondrial genomes for the corn rootworm species, *Diabrotica virgifera virgifera* and *Diabrotica barberi* (Insecta: Coleoptera: Chrysomelidae), using Next Generation Sequence data. *Gene*, **542**: 190-197. <https://doi.org/10.1016/j.gene.2014.03.035>
- Crampton-Platt, A., Timmermans, M.J.T.N., Gimmel, M.L., Kuty, S.N., Cockerill, T.D., Vun Khen C. and Vogler, A.P., 2015. Soup to tree: the phylogeny of beetles inferred by mitochondrial metagenomics of a Bornean rainforest sample. *Mol. Biol. Evol.*, **32**: 2302-2316. <https://doi.org/10.1093/molbev/msv111>
- Crowson, R.A., 1960. The phylogeny of coleopteran. *Annu. Rev. Ent.*, **5**: 111-134. <https://doi.org/10.1146/annurev.en.05.010160.000551>
- De Rijk, P., Wuyts, J. and De Wachter, R., 2003. RnaViz 2: An improved representation of RNA secondary structure. *Bioinformatics*, **19**: 99-300. <https://doi.org/10.1093/bioinformatics/19.2.299>
- Dierckxsens, N., and Mardulyn, P. and Smits, G., 2017. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucl. Acids Res.*, **45**: e18. <https://doi.org/10.1093/nar/gkw955>
- Dowton, M., Castro, L.R. and Austin, A.D., 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: The examination of genome morphology. *Invertebr. Syst.*, **16**: 345-356. <https://doi.org/10.1071/IS02003>
- Du, C., He, S.L., Song, X.H., Liao, Q., Zhang, X.Y. and Yue, B.S., 2016. The complete mitochondrial genome of *Epicauta chinensis* (Coleoptera: Meloidae) and phylogenetic analysis among coleopteran insects. *Gene*, **578**: 274-280. <https://doi.org/10.1016/j.gene.2015.12.036>
- Du, Y.M., Zhang, C.N., Dietrich, C.H., Zhang, Y.L. and Dai, W., 2017. Characterization of the complete mitochondrial genomes of *Maiestas dorsalis* and *Japananus hyalinus* (Hemiptera: Cicadellidae) and comparison with other Membracoidea. *Sci. Rep. U.K.*, **7**: 14197. <https://doi.org/10.1038/s41598-017-14703-3>
- Fang, J., Qian, L., Xu, M., Yang, X., Wang, B. and An, Y., 2016. The complete nucleotide sequence of the mitochondrial genome of the Asian longhorn beetle, *Anoplophora glabripennis* (Coleoptera: Cerambycidae). *Mitochondrial DNA A*, **27**: 3299-3300. <https://doi.org/10.3109/19401736.2015.1015012>
- Farrell, B.D. and Sequeira, A.S., 2004. Evolutionary rates in the adaptive radiation of beetles on plants. *Evolution*, **58**: 1984-2001. <https://doi.org/10.1111/j.0014-3820.2004.tb00484.x>

- Friedrich, M. and Muqim, N., 2003. Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castaneum*. *Mol. Phylogenet. Evol.*, **26**: 502-512. [https://doi.org/10.1016/S1055-7903\(02\)00335-4](https://doi.org/10.1016/S1055-7903(02)00335-4)
- Gómez-Zurita, J., Hunt, T., Kopliku, F. and Vogler, A.P., 2007. Recalibrated tree of leaf beetles (Chrysomelidae) indicates independent diversification of angiosperms and their insect herbivores. *PLoS One*, **2**: e360. <https://doi.org/10.1371/journal.pone.0000360>
- Grimaldi, D. and Engel, M.S., 2003. Evolution of the insects, Mitochondrial genome of the flour beetle *Tribolium castaneum*. *Mol. Phylogenet. Evol.*, **26**: 502-512. [https://doi.org/10.1016/S1055-7903\(02\)00335-4](https://doi.org/10.1016/S1055-7903(02)00335-4)
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.*, **59**: 307-321. <https://doi.org/10.1093/sysbio/syq010>
- Hao, J.S., Sun, Q.Q., Zhao, H.B., Sun, X.Y., Gai, Y.H. and Yang, Q., 2012. The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperidae: Pyrginae) and its phylogenetic implication. *Int. J. Genom.*, **2012**: 328049. <https://doi.org/10.1155/2012/328049>
- Huang, M.C., Wang, Y.Y., Liu, X.Y., Li, W., Kang, Z., Wang, K., Li, X. and Yang, D., 2015. The complete mitochondrial genome and its remarkable secondary structure for a stonefly *Acroneuria hainana* Wu (Insecta: Plecoptera, Perlidae). *Gene*, **557**: 52-60. <https://doi.org/10.1016/j.gene.2014.12.009>
- Huang, Y., Liu, Y., Zhu, X.Y., Xin, Z.Z., Zhang, H.B., Zhang, D.Z., Wang, J.L., Tang, B.P., Zhou, C.L., Liu, Q.N. and Dai, L.S., 2019. Comparative mitochondrial genome analysis of *Grammodes geometrica* and other noctuid insects reveals conserved mitochondrial genome organization and phylogeny. *Int. J. Biol. Macromol.*, **125**: 1257-1265. <https://doi.org/10.1016/j.ijbiomac.2018.09.104>
- Hunt, T., Bergsten, J., Levkanicova, Z., Papadopoulou, A., John, O.S., Wild, R., Hammond, P.M., Ahrens, D., Balke, M. and Caterino, M.S., 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, **318**: 1913-1916. <https://doi.org/10.1126/science.1146954>
- James, B.S. and Andrew, T.B., 2006. Insect mitochondrial genomics 2: The complete mitochondrial genome sequence of a giant stonefly, *Pteronarcys princeps*, asymmetric directional mutation bias, and conserved plecopteran A + T-region elements. *Genome*, **49**: 815-824. <https://doi.org/10.1139/g06-037>
- Jin, S.B., Kim, I., Sohn, H.D. and Jin, B.R., 2004. The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization and phylogenetic analysis with other insects. *Mol. Biol. Evol.*, **32**: 978-985. <https://doi.org/10.1016/j.ympev.2004.03.009>
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. and Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**: 1647-1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Kim, I., Lee, E.M., Seol, K.Y., Yun, E.Y., Lee, Y.B., Hwang, J.S. and Jin, B.R., 2006. The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). *Insect mol. Biol.*, **15**: 217-225. <https://doi.org/10.1111/j.1365-2583.2006.00630.x>
- Kim, K.G., Hong, M.Y., Kim, M.J., Im, H.H., Kim, M.I., Bae, C.H., Seo, S.J., Lee, S.H. and Kim, I., 2009. Complete mitochondrial genome sequence of the yellow-spotted long-horned beetle *Psacotha hilaris* (Coleoptera: Cerambycidae) and phylogenetic analysis among coleopteran insects. *Mol. Cells*, **27**: 429-441. <https://doi.org/10.1007/s10059-009-0064-5>
- Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.*, **33**: 870-874. <https://doi.org/10.1093/molbev/msw054>
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**: 2947-2948. <https://doi.org/10.1093/bioinformatics/btm404>
- Lawrence, J.F., Slipinski, A., Seago, A.E., Thayer, M.K., Newton, A.F. and Marvaldi, A.E., 2011. Phylogeny of the Coleoptera based on morphological characters of adults and larvae. *Annls Zool.*, **61**: 1-217. <https://doi.org/10.3161/000345411X576725>
- Li, H., Liu, H., Shi, A., Stys, P., Zhou, X. and Cai, W., 2012. The complete mitochondrial genome and novel gene arrangement of the unique-headed bug *Stenopirates* sp. (Hemiptera: Enicocephalidae). *PLoS One*, **7**: e29419. <https://doi.org/10.1371/>

- [journal.pone.0029419](https://doi.org/10.1371/journal.pone.0029419)
- Li, L., Liu, G., Sun, T., Xin, T., Li, M., Zou, Z. and Xia, B., 2016. Complete mitochondrial genome of *Cryptolestes pusillus* (Coleoptera: Laemophloeidae). *Mitochondrial DNA A*, **27**: 3703-3704. <https://doi.org/10.3109/19401736.2015.1079865>
- Li, N., Hu, G. and Hua, B., 2019. Complete mitochondrial genomes of *Bittacus strigosus* and *Panorpa debilis* and genomic comparisons of Mecoptera. *Int. J. Biol. Macromol.*, **140**: 672-681. <https://doi.org/10.1016/j.ijbiomac.2019.08.152>
- Li, Q., Wei, S.J., Shi, M. and Chen, X.X., 2015. The complete mitochondrial genome of *Neopanorpa pulchra* (Mecoptera: Panorpididae). *Mitochondrial DNA A*, **26**: 305-306. <https://doi.org/10.3109/19401736.2013.825787>
- Li, W., Yang, X., Qian, L., An, Y. and Fang, J., 2016. The complete mitochondrial genome of the citrus longhorn beetle, *Anoplophora chinensis* (Coleoptera: Cerambycidae). *Mitochondrial DNA A*, **27**: 4665-4667. <https://doi.org/10.3109/19401736.2015.1106493>
- Li, X., Ou, J., Wei, Z., Li, Y. and Tian, Y., 2016. The mitogenomes of three beetles (Coleoptera: Polyphaga: Cucujiformia): New gene rearrangement and phylogeny. *Biochem. Syst. Ecol.*, **69**: 101-107. <https://doi.org/10.1016/j.bse.2016.08.012>
- Liu, J., Jia, P., Luo, T. and Wang, Q., 2017. Complete mitochondrial genome of white-striped longhorn beetle, *Batocera lineolata* (Coleoptera: Cerambycidae) by next-generation sequencing and its phylogenetic relationship within superfamily Chrysomeloidea. *Mitochondrial DNA B*, **2**: 520-521. <https://doi.org/10.1080/23802359.2017.1361797>
- Liu, Y.Q., Chen, D.B., Liu, H.H., Hu, H.L., Bian, H.X., Zhang, R.S., Yang, R.S., Jiang, X.F. and Shi, S.L., 2018. The complete mitochondrial genome of the longhorn beetle *Dorysthenes paradoxus* (Coleoptera: Cerambycidae: Prionini) and the implication for the phylogenetic relationships of the Cerambycidae species. *J. Insect Sci.*, **18**: 21. <https://doi.org/10.1093/jisesa/iey012>
- Lowe, T.M. and Chan, P.P., 2016. tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes. *Nucl. Acids Res.*, **44**: W54-W57. <https://doi.org/10.1093/nar/gkw413>
- Lu, H.F., Su, T.J., Luo, A.R., Zhu, C.D. and Wu, C.S., 2013. Characterization of the complete mitochondrion genome of diurnal moth *Amata emma* (Butler) (Lepidoptera: Erebiidae) and its phylogenetic implications. *PLoS One*, **8**: e72410. <https://doi.org/10.1371/journal.pone.0072410>
- Luo, R., Liu, B., Xie, Y., Li, Z.Y., Huang, W.H., Yuan, J.Y., He, G.Z., Chen, Y.X., Pan, Q., Liu, Y.J., Tang, J.B., Wu, G.X., Zhang, H., Shi, Y.J., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C.L., Cheung, D.W., Yiu, S.M., Peng, S.L., Zhu, X.Q., Liu, G.M., Liao, X.K., Li, Y.R., Yang, H.M., Wang, J., Lam T.W. and Wang, J., 2012. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Giga Sci.*, **1**: 18. <https://doi.org/10.1186/2047-217X-1-18>
- Ma, C., Yang, P.C., Jiang, F., Chapuis, M.P., Shali, Y., Sword, G.A. and Kang, L., 2012. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Mol. Ecol.*, **21**: 4344-4358. <https://doi.org/10.1111/j.1365-294X.2012.05684.x>
- Marvaldi, A.E., Duckett, C.N., Kjer, K.M. and Gillespie, J.J., 2009. Structural alignment of 18s and 28s rDNA sequences provides insights into phylogeny of Phytophaga (Coleoptera: Curculionoidea and Chrysomeloidea). *Zool. Scr.*, **38**: 63-77. <https://doi.org/10.1111/j.1463-6409.2008.00360.x>
- McDougall, D.N., 2001. *Anoplophora chinensis*. USDA Forest Service, State and Private Forestry. North American Forest Commission Exotic Forest Pest Information System (NAFC-ExFor).
- Meiklejohn, K.A., Danielson, M.J., Faircloth, B.C., Glenn, T.C., Braun, E.L. and Kimball, R.T., 2014. Incongruence among different mitochondrial regions: A case study using complete mitogenomes. *Mol. Phylogenet. Evol.*, **78**: 314-323. <https://doi.org/10.1016/j.ympev.2014.06.003>
- Nie, R.E. and Yang, X.K., 2013. Higher level phylogeny of Coleoptera based on molecular methods: current status and problems. *Acta entomol. Sin.*, **56**: 1055-1062.
- Ødegaard, F., 2000. How many species of arthropods? Erwin's estimate revised. *Biol. J. Linn. Soc.*, **71**: 583-597. <https://doi.org/10.1111/j.1095-8312.2000.tb01279.x>
- Ojala, D., Montoya, J. and Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*, **290**: 470-474. <https://doi.org/10.1038/290470a0>
- Perna, N.T. and Kocher, T.D., 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. mol. Evol.*, **41**: 353-358. <https://doi.org/10.1007/BF01215182>
- Posada, D. and Buckley, T.R., 2004. Model selection and model averaging in phylogenetics:

- Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst. Biol.*, **53**: 793-808. <https://doi.org/10.1080/10635150490522304>
- Posada, D., 2008. J model test: Phylogenetic model averaging. *Mol. Biol. Evol.*, **25**: 1253-1256. <https://doi.org/10.1093/molbev/msn083>
- Qian, Y.H., Wu, H.Y., Ji, X.Y., Yu, W.W. and Du, Y.Z., 2014. Mitochondrial genome of the stonefly *Kamimuria wangi* (Plecoptera: Perlidae) and phylogenetic position of Plecoptera based on mitogenomes. *PLoS One*, **9**: e86328. <https://doi.org/10.1371/journal.pone.0086328>
- Rambaut, A. and Drummond, A.J., 2007. Tracer v. 1.5. Available at <http://www.beast.bio.ed.ac.uk/Tracer>.
- Rambaut, A., 2009. FigTree v1.3.1. Available at <http://tree.bio.ed.ac.uk/software/> (accessed on 31 September 2017).
- Richard, H.E. and Gregory, P., 2002. Longhorned beetles of the genus *Anoplophora* and Lithography: Alien invaders in the eye of the beholder. *Am. Entomol.*, **48**: 200-206. <https://doi.org/10.1093/ae/48.4.200>
- Robertson, J.A., Ślipiński, A., Moulton, M., Shockley, F.W., Giorgi, A., Lord, N.P., McKenna, D.D., Tomaszewska, W., Forrester, J., Miller, K.B., Whiting, M.F. and Mchugh, J.V., 2015. Phylogeny and classification of Cucujoidea and the recognition of a new superfamily Coccinelloidea (Coleoptera: Cucujiformia). *Syst. Ent.*, **40**: 745-778. <https://doi.org/10.1111/syen.12138>
- Ronquist, F., Teslenko, M., Mark, P.V.D., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. and Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.*, **61**: 539-542. <https://doi.org/10.1093/sysbio/sys029>
- Sama, G., Buse, J., Orbach, E., Friedman, A., Rittner, O. and Chikatunov, V., 2010. A new catalogue of the Cerambycidae (Coleoptera) of Israel with notes on their distribution and host plants. *Munis Ent. Zool.*, **5**: 1-55.
- Schuster, G. and Stern, D., 2009. RNA polyadenylation and decay in mitochondria and chloroplasts. *Prog. mol. Biol. Transl.*, **85**: 393-422. [https://doi.org/10.1016/S0079-6603\(08\)00810-6](https://doi.org/10.1016/S0079-6603(08)00810-6)
- Song, H., Sheffield, N.C., Cameron, S.L., Miller, K.B. and Whiting, M.F., 2010. When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics. *Syst. Ent.*, **35**: 429-448. <https://doi.org/10.1111/j.1365-3113.2009.00517.x>
- Stewart, J.B. and Beckenbach, A.T., 2003. Phylogenetic and genomic analysis of the complete mitochondrial DNA sequence of the spotted asparagus beetle *Crioceris duodecimpunctata*. *Mol. Phylogenet. Evol.*, **26**: 513-526. [https://doi.org/10.1016/S1055-7903\(02\)00421-9](https://doi.org/10.1016/S1055-7903(02)00421-9)
- Stewart, J.B. and Beckenbach, A.T., 2005. Insect mitochondrial genomics: The complete mitochondrial genome sequence of the meadow spittlebug *Philaenus spumarius* (Hemiptera: Auchenorrhyncha: Cercopoidae). *Genome*, **48**: 46-54. <https://doi.org/10.1139/g04-090>
- Sun, Y., Chen, C., Gao, J., Abbas, M.N., Kausar, S., Qian, C., Wang, L., Wei, G.Q., Zhu, B.J. and Liu, C.L., 2017. Comparative mitochondrial genome analysis of *Daphnis nerii* and other lepidopteran insects reveals conserved mitochondrial genome organization and phylogenetic relationships. *PLoS One*, **12**: e0178773. <https://doi.org/10.1371/journal.pone.0178773>
- Taanman, J.W., 1999. The human mitochondrial genome: structure, transcription, translation and replication. *BBA-Bioenerg.*, **14**: 103-123. [https://doi.org/10.1016/S0005-2728\(98\)00161-3](https://doi.org/10.1016/S0005-2728(98)00161-3)
- Talavera, G. and Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.*, **56**: 564-577. <https://doi.org/10.1080/10635150701472164>
- Timmermans, M.J., Lees, D.C. and Simonsen, T.J., 2014. Towards a mitogenomic phylogeny of Lepidoptera. *Mol. Biol. Evol.*, **7**: 169-178. <https://doi.org/10.1016/j.ympev.2014.05.031>
- Timmermans, M.J.T.N., Barton, C., Haran, J., Ahrens, D., Culverwell, C.L., Ollikainen, A., Dodsworth, S., Foster, P.G., Bocak, L., Vogler, A.P., 2016. Family-level sampling of mitochondrial genomes in coleoptera: Compositional heterogeneity and phylogenetics. *Genome Biol. Evol.*, **8**: 161-175. <https://doi.org/10.1093/gbe/evv241>
- Wang, C., Feng, Y. and Chen, X., 2012. Complete sequence and gene organization of the mitochondrial genome of *Batocera lineolata* Chevrolat (Coleoptera: Cerambycidae). *Chin. Sci. B-Chin.*, **57**: 78-85. <https://doi.org/10.1007/s11434-012-5271-3>
- Wang, C.Y., Feng, Y. and Chen, X.M., 2013. Complete coding region of the mitochondrial genome of *Monochamus alternatus* Hope (Coleoptera: Cerambycidae). *Zool. Sci.*, **30**: 570-576. <https://doi.org/10.2108/zsj.30.570>

- Wang, J., Dai, X., Xu, X., Zhang, Z., Yu, D., Storey, K.B. and Zhang, J., 2019. The complete mitochondrial genomes of five longicorn beetles (Coleoptera: Cerambycidae) and phylogenetic relationships within Cerambycidae. *PeerJ*, **7**: e7633. <https://doi.org/10.7717/peerj.7633>
- Wang, J.J., Yang, M.F., Dai, R.H., Li, H. and Wang, X.Y., 2018. Characterization and phylogenetic implications of the complete mitochondrial genome of Idiocerinae (Hemiptera: Cicadellidae). *Int. J. Biol. Macromol.*, **120**: 2366-2372. <https://doi.org/10.1016/j.ijbiomac.2018.08.191>
- Wang, Q. and Tang, G., 2018. The mitochondrial genomes of two walnut pests, *Gastrolina depressa* and *G. depressa thoracica* (Coleoptera: Chrysomelidae) and phylogenetic analyses. *PeerJ*, **6**: e4919. <https://doi.org/10.7717/peerj.4919>
- Wang, Y., Cao, J.J., Li, N., Ma, G.Y. and Li, W.H., 2019. The first mitochondrial genome from Scopuridae (Insecta: Plecoptera) reveals structural features and phylogenetic implications. *Int. J. Biol. Macromol.*, **122**: 893-902. <https://doi.org/10.1016/j.ijbiomac.2018.11.019>
- Wolstenholme, D.R., 1992. Animal mitochondrial DNA: Structure and evolution. *Int. Rev. Cytol.*, **141**: 173-216. [https://doi.org/10.1016/S0074-7696\(08\)62066-5](https://doi.org/10.1016/S0074-7696(08)62066-5)
- Wu, H.Y., Ji, X.Y., Yu, W.W. and Du, Y.Z., 2014. Complete mitochondrial genome of the stonefly *Cryptoperla stilifera* Sivec (Plecoptera: Peltoperlidae) and the phylogeny of Polyneopteran insects. *Gene*, **537**: 177-183. <https://doi.org/10.1016/j.gene.2013.12.044>
- Yao, J., Yang, H. and Dai, R., 2017. Characterization of the complete mitochondrial genome of *Acanthoscelides obtectus* (Coleoptera: Chrysomelidae: Bruchinae) with phylogenetic analysis. *Genetica*, **145**: 397-408. <https://doi.org/10.1007/s10709-017-9975-9>
- Yin, H., Zhi, Y.C., Jiang, H.D., Wang, P.X., Yin, X.C. and Zhang, D.C., 2012. The complete mitochondrial genome of *Gomphocerus tibetanus* Uvarov, (Orthoptera: Acrididae: Gomphocerinae). *Gene*, **492**: 214-218. <https://doi.org/10.1016/j.gene.2011.12.020>
- Yuan, M., Zhang, Q., Zhang, L., Guo, Z., Liu, Y., Shen, Y. and Shao, R.F., 2016. High-level phylogeny of the coleoptera inferred with mitochondrial genome sequences. *Mol. Phylogenet. Evol.*, **104**: 99-111. <https://doi.org/10.1016/j.ympev.2016.08.002>
- Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K. and Banno, Y., 2002. Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkmoth, *Bombyx mandarina*, and its close relative, the domesticated silkmoth, *Bombyx mori*. *Mol. Biol. Evol.*, **19**: 1385-1389. <https://doi.org/10.1093/oxfordjournals.molbev.a004200>
- Zhang, D.X. and Hewitt, G.M., 1997. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. *Biochem. Syst. Ecol.*, **25**: 99-120. [https://doi.org/10.1016/S0305-1978\(96\)00042-7](https://doi.org/10.1016/S0305-1978(96)00042-7)
- Zhang, S.Q., Che, L.H., Li, Y., Liang, D., Pang, H., Ślipiński, A. and Zhang, P., 2018. Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nat. Commun.*, **9**: 205. <https://doi.org/10.1038/s41467-017-02644-4>
- Zhang, Y., Feng, S., Fekrat, L., Jiang, F., Khathutshelo, M. and Li, Z., 2019. The first two complete mitochondrial genome of *Dacus bivittatus* and *Dacus ciliatus* (Diptera: Tephritidae) by next-generation sequencing and implications for the higher phylogeny of Tephritidae. *Int. Rev. Cytol.*, **140**: 469-476. <https://doi.org/10.1016/j.ijbiomac.2019.08.076>
- Zhang, Z.Q., 2013. Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa*, **37**: 1-82.
- Zhou, X., Han, H., Pang, B. and Zhang, P., 2016. The complete mitochondrial genome of *Galeruca daurica* (Joannis) (Coleoptera: Chrysomelidae). *Mitochondrial DNA A*, **27**: 2891-2892. <https://doi.org/10.3109/19401736.2015.1060424>
- Zhu, J.Y., Yang, P., Zhang, Z., Wu, G.X. and Yang, B., 2013. Transcriptomic immune response of *Tenebrio molitor* pupae to parasitization by *Scleroderma guani*. *PLoS One*, **8**: 173-185. <https://doi.org/10.1371/journal.pone.0054411>