



Comparative Mitochondrial Genomic Analysis of Longhorn Beetles (Coleoptera: Chrysomeloidea) with Phylogenetic Implications

Yiming Niu^{1,*}, Fengming Shi^{1,*}, Xinyu Li¹, Sainan Zhang¹, Yabei Xu¹, Jing Tao¹, Meng Li¹, Yuxuan Zhao¹, Shixiang Zong¹

¹ Beijing Key Laboratory for Forest Pest Control, Beijing Forestry University, Beijing 100083, China

* These authors contributed equally to the work.

<https://zoobank.org/F9A35225-FBA1-4483-80FE-8DF91414F41E>

Corresponding author: Shixiang Zong (zongshixiang@bjfu.edu.cn)

Received 17 October 2023
Accepted 18 December 2023
Published 22 March 2024

Academic Editors André Nel, Marianna Simões

Citation: Niu Y, Shi F, Li X, Zhang S, Xu Y, Tao J, Li M, Zhao Y, Zong S (2024) Comparative Mitochondrial Genomic Analysis of Longhorn Beetles (Coleoptera: Chrysomeloidea) with Phylogenetic Implications. *Arthropod Systematics & Phylogeny* 82: 133–150. <https://doi.org/10.3897/asp.82.e114299>

Abstract

Longhorn beetles (Cerambycidae) play a vital role in global ecosystems. Some of them contribute to nutrient cycling and pollination, while others, pose a threat to forestry production. Despite their ecological importance, there has been a lack of comprehensive analyses on the mitochondrial genomes of Cerambycidae beetles. In this study, we have conducted mitochondrial genome sequencing and annotation for four Cerambycidae beetles: *Monochamus sutor*, *Monochamus guerryi*, *Monochamus galloprovincialis*, and *Monochamus latifasciatus*. Our analysis revealed a high degree of conservation in these mitochondrial genomes, with rare gene rearrangements observed across the Cerambycidae family. Additionally, a notable bias towards AT content was identified, with most genes using ATN as the start codon and TAA as the stop codon. Except for trnS1, all tRNA genes showed typical cloverleaf secondary structures. Phylogenetic analysis using IQ-TREE and Phylobayes consistently produced congruent topologies. At the gene level analyses, our results highlighted the remarkable conservation of the *COX1* gene. Furthermore, at the species level, we observed strong adaptability in the Spondylidinae and Lepturinae subfamilies. We also offer our insights into contentious aspects of the phylogeny. Overall, our research contributes to a deeper understanding of the phylogeny and evolution of Cerambycidae, laying the groundwork for future population genetic investigations.

Key words

Cerambycidae, Comparison, *Monochamus*, Mitochondrial genomic, Phylogeny, Systematics

1. Introduction

The order Coleoptera (beetles) is the largest and most widely distributed order in insect with a staggering 380,000 documented species worldwide. Beetles make up approximately a quarter of all known animal species

on Earth and many more species yet to be documented (Zhang et al. 2018; Tavakilian et al. 2019). Among the whole Coleoptera, longhorn beetles in the Cerambycidae emerges as one of the most diverse and ecologically sig-

nificant families, comprising approximately 40,000 species (Footitt et al. 2018). The evolutionary trajectory of the Cerambycidae classification has spanned a protracted course, and to this day, it remains a subject of divergent perspectives among scholars. The documentation of longhorn beetle species dates back to 1758 when Linné described and classified 76 such species in the tenth edition of “Systema Naturae,” initially dividing them into three genera: *Cerambyx*, *Leptura* and *Necydalis* (Linnaeus 1758). In 1762, Geoffroy added two more genera: *Prionus* and *Stenocorus* (Geoffroy 1762). In 1801, the Danish entomologist Fabricius documented 594 longhorn beetle species (Fabricius 1801). Despite the continuous expansion of genera and species, it was only in 1815 that Leach formally recognized the familial rank of Cerambycidae and coined the term “Cerambycidea” in the book of the “Edinburgh Encyclopedia” (Brewster 1832). In 1829, Latreille classified the family into Prioninae, Cerambycinae, Lamiinae, and Lepturinae (Latreille 1892). After that, many taxonomists tried to further refine the classification of the Cerambycidae family. Taxonomists systematically summarized and categorized higher-level classifications within the Cerambycidae family through various aspects such as comparative morphology, reproductive systems, and biogeography. Up to now, there is still controversy surrounding the taxonomic status of the Cerambycidae family.

Longhorn beetles play a crucial role in global ecosystems. They aid in the decomposition of dead branches and woody plants, promoting ecosystem cycling. They serve as pollinators, providing essential assistance in the reproduction of plants (Nie et al. 2020). However, longhorn beetles can also pose a threat to forests, such as *Anoplophora glabripennis* (Motschulsky), *Jebusaea hamerschmidtii* Reiche, and *Aromia bungii* (Faldermann) (MacLeod et al. 2002; El-Shafie et al. 2021; Yamamoto et al. 2002). Notably, many *Monochamus* species in the subfamily of Lamiinae serve as vectors for the devastating pine wood nematodes (PWD), responsible for plenty of pine trees dead (Chen et al. 2023). Some *Monochamus* genus longhorn beetles, such as *Monochamus alternatus* Hope and *Monochamus saltuarius* Gebler, share remarkable similarities, making their identification a challenging task (Shi et al. 2022). Traditional identification methods relied on body size, elytra sculpture, and the presence of scutellum hairs (Patel et al. 2012; Monné et al. 2017), often subject to human subjectivity. Consequently, a comprehensive understanding of systematic phylogenetics can greatly assist in the classification and identification of these beetles.

With the continuous advancement of molecular biology technology, mitochondrial genome (mitogenome) has been widely used in species identification, kinship identification, population genetics, phylogeny, Geographic distribution, migration, evolution, and so on (Ovenden et al. 2018; Shi et al. 2023). Insects typically possess complete mitochondrial genomes, which are double-stranded closed circular DNA molecules, ranging in size from 14 to 20 kilobases (kb). These mitogenomes consist of 37 genes, including 13 protein-coding genes (PCGs), 2

ribosomal RNA genes (rRNAs), 22 transfer RNA genes (tRNAs), and 1 control region, also known as the A+T enrichment region (Boore 1999). The unique attributes of insect mitochondrial genomes, such as intron deletion, maternal inheritance, relatively low levels of recombination, and rapid mutation rates, render them invaluable as molecular markers for diverse applications (Behura 2006; Freeland 2020). Moreover, the mitochondrial genome provides a wealth of phylogenetic information, including gene order, RNA secondary structure, codon use frequency, and control region characteristics (Johnson 2019). Additionally, the independence of mitochondrial genetic material from the nuclear genome makes it well-suited for elucidating relationships among different populations and species (Silva-Pinheiro et al. 2022). So integrating and conducting comparative studies of the mitochondrial genome data with other valuable information, such as nDNA, morphology, behavior, etc., will ultimately lead to an enhanced, more comprehensive understanding of systematic biology and the historical relationships among living organisms (Rubinoff et al. 2005).

However, the number of complete mitochondrial genome sequences in longhorn beetles is remarkably limited and there is still a lack of comprehensive comparative analysis of the mitochondrial genomes among them. Therefore, we sequenced the mitochondrial genomes of four species within the *Monochamus* genus of longhorn beetles and conducted comparative mitochondrial genomic analyses in conjunction with other longhorn beetle species available in GenBank (<https://www.ncbi.nlm.nih.gov>). Our study enriches the mitochondrial genome and provides robust data support for an in-depth investigation of the phylogeny of the Cerambycidae. It may shed new light on the complex phylogenetic relationships in Cerambycidae.

2. Materials and methods

2.1. Sample collection and DNA extraction

Adult *Monochamus sutor* (Linnaeus), *Monochamus guerryi* Pic, *Monochamus galloprovincialis* (Olivier), and *Monochamus latefasciatus* (Breuning) species collected information was shown in Table 1, and preserved in anhydrous ethanol (purity $\geq 99.7\%$) at -20°C . Specimens were first identified according to external morphological characters before the DNA extraction (Sun 2019; Silva-Pinheiro et al. 2022). A small amount of muscle tissue was removed from the chest of each four longhorn beetles for total DNA extraction. Total genomic DNA was extracted using TIANamp Genomic DNA Kit (TIANGEN, Beijing, China) following the manufacturer’s instructions. The *cox1* sequences were amplified with the primers LCO149/HCO2198 and then retrieved from the National Center for Biotechnology Information database to identify from the molecular level (Folmer 1994).

Table 1. Collection information of species in this study.

Species Name	Host Plant	Date	Locality
<i>Monochamus sutor</i>	<i>Larix gmelinii</i> (Ruprecht) Kuzeneva	Jul., 2020	CHINA – Beijing – 39°54'N, 116°25'E
<i>Monochamus guerryi</i>	<i>Quercus glauca</i> Thunb.	Jun., 2020	CHINA – Guangxi – 22°02'N, 106°47'E
<i>Monochamus galloprovincialis</i>	<i>Pinus sylvestris</i> var. <i>mongolica</i>	Jul., 2020	CHINA – Heilongjiang – 51°02'N, 122°11'E
<i>Monochamus latefasciatus</i>	unkonwn	Jul., 2019	CHINA – Guangxi – 22°02'N, 106°47'E

2.2. Mitochondrial genome sequencing, assembly, and annotation

DNA samples of four longhorn beetles were sequenced by next-generation sequencing at Biomarker Technologies Co., Ltd. (Beijing, China) to obtain complete mitochondrial genes. An Illumina TruSeq library was generated, with average lengths of 350 bp and 250 bp paired-end reads sequenced on Illumina HiSeq 2500 platform (San Diego, CA). After filtering using the NGS QC Toolkit v2.3.3 (Patel et al. 2012), clean reads with high quality were used for subsequent genome assembly and identified through MitoZ 3.6 (Meng et al. 2019), with the mitochondrial genomes of *M. saltuarius* and *Monochamus urussovii* (Fischer-Waldheim) (OP169419 and OP169420) being used as references.

Protein-coding genes (PCGs) and ribosomal RNA (rRNA) genes were annotated by aligning with known homologous Cerambycidae species mitogenome sequences from GenBank. Additionally, the secondary structures of transfer RNA (tRNA) genes were identified using tRNA-scan-SE Search Server v2.0 (<http://lowelab.ucsc.edu/tRNA-scan-SE/>) (Lowe et al. 2016). The limits of PCGs and rRNA genes were determined by comparing them with analogous genes found in other Cerambycidae species and by referencing the positions of transfer RNA (tRNA) genes. To affirm the precision of the deduced locations, the nucleotide sequences of 13 PCGs were translated into amino acids through the utilization of MEGA X (Kumar et al. 2018). The mitogenomes were visualized using the proksee website (<https://proksee.ca/>).

2.3. Comparative mitogenome analyses

The mitogenomes of 89 species obtained from the NCBI (<https://www.ncbi.nlm.nih.gov>) and 4 newly sequenced species were used for analysis (Table S1). Subsequently, an online software tool called Interactive Tree Of Life (iTOL) v 6.8 (Letunic et al. 2019) was employed to visualize the gene order of all species with different gene represented in distinct colors for enhanced clarity and comparison. Mitogenome rearrangement events were completed through the manual comparison, with *Drosophila yakuba* Burla as a reference (Clary et al. 1985). Pairwise comparisons of DNAs between *M. sutor* and other longhorn beetles mitochondrial genomes were performed using BLASTN searches in the CGView Comparison Tool (CCT) (https://github.com/paulstothard/cgview_comparison_tool). The base composition, ami-

no acid usage, codon usage, and relative synonymous codon usage (RSCU) were analyzed through MEGA11/PhyloSuite v1.2.3 (Tamura et al. 2021; Xiang et al. 2023). The formula: AT-skew = $(A-T) / (A+T)$ and GC-skew = $(G-C) / (G+C)$ was calculated manually to evaluate nucleotide composition bias (Perna et al. 1995). The MEGA 11 (Tamura et al. 2021) software was utilized to compute the average genetic distances employing the Kimura-2-parameter model across 13 PCGs. Synonymous (Ks) and non-synonymous (Ka) substitution rates among PCGs were calculated by DNASP v5.10 (Librado et al. 2009). The effective number of codons (ENC) and codon bias index (CBI) for PCGs were also calculated by the software DNASP 5.10 (Librado et al. 2009).

2.4. Phylogenetic analysis

The mitochondrial genome data were obtained from the NCBI nucleotide database. A total of 86 species in 10 groups were selected from the longhorn beetles. Additionally, the dataset incorporated the mitochondrial genomes of the four newly sequenced species.

After filtering the redundant sequence, 13 PCGs and 2 rRNAs were extracted by PhyloSuite v1.2.3 (Xiang et al. 2023) for downstream analysis. Multiple sequence alignments were conducted for amino acids, PCGs, and rRNA gene using MAFFT v7.313 (Katoh et al. 2013). For the alignment of PCGs, the L-INS-i algorithm was employed, while the G-INS-i algorithm was utilized for the alignment of rRNA genes (Katoh et al. 2013). The alignments of PCGs were imported into the MACSE v2.03 (Ranwez et al. 2018) for optimization. Sequences were concatenated using PhyloSuite version 1.2.3 to create four distinct datasets: (i) PCG123: PCGs with all three codon positions; (ii) PCG123R: PCG123 dataset plus two rRNA genes; (iii) PCG12R: PCG12 dataset plus two rRNA genes; (iv) PCGAA: amino acid sequences translated from PCGs. ModelFinder (Kalyaanamoorthy et al. 2017) and PartitionFinder2 (Lanfear et al. 2017) were used to select optimal partitioning strategies and evolutionary models for PCGsRNA and AA datasets. The heterogeneity and substitution saturation of each dataset were assessed using the AliGROOVE v1.08 (<https://github.com/PatrickKueck/AliGROOVE>) (Kück et al. 2014) and DAMBE v5.5.29 (Xia 2013), respectively.

Maximum likelihood (ML) phylogenies and Phylobayes MPI 1.5a were used to construct phylogenetic trees mainly about Cerambycidae on PhyloSuite v1.2.3 and CIPRES Science Gateway (<https://www.phylo.org>). Seven species with complete sequences from other fam-

ilies that similar to Cerambycidae were selected as outgroups (*Eucryptorrhynchus scrobiculatus* Motschulsky, *Eucryptorrhynchus brandti* (Harold), *Curculio davidi* Fairmaire, *Sympiezomias velatus* (Chevrolat), *Naupactus xanthographus* Germaer, *Aegorhinus superciliosus* (Guérin), *Sitophilus zeamais* Motschulsky). For ML analysis, the IQ-TREE v. 1.6.8 (Minh et al. 2020) was used to reconstruct the ML phylogenetic tree under the Partition Mode (Table S2). A standard bootstrap approximation approach was selected, and 1000 replicates were performed to assess the tree's robustness. Phylobayes MPI on XSEDE (version 1.8c) was utilized through the CIPRES Science Gateway (<https://www.phylo.org>) to construct Bayesian inference (BI) phylogenies. The site-heterogeneous mixture model CAT+GTR was employed for both nucleotide and amino acid datasets (Lartillot et al. 2013). Two independent chains, a Markov chain, and a Monte Carlo chain, were run after removing a fixed point from a sequence. The chains were stopped after two runs of 30000 generations converged satisfactorily, indicated by a maxdiff value of less than 0.1. The initial 25% of trees generated were discarded as burn-in.

3. Results and Discussion

3.1. The general features of mitogenomes

Complete mitogenomes of four Lamiinae species were newly sequenced: *M. sutor*, *M. guerryi*, *M. galloprovincialis* and *M. latefasciatus*, which ranged in size from 15759 (*M. guerryi*) to 15929 (*M. latefasciatus*) base pairs and each contained 37 genes (13PCGs, 22 tRNA and 2 rRNA) (Fig. 1). These mitogenomes exhibit a circular and double-stranded structure. The majority strand (J-strand) encodes 23 genes, including 9 PCGs (*atp6*, *atp8*, *cytb*, *cox1*, *cox2*, *cox3*, *nad2*, *nad3* and *nad6*), and 14 tRNA genes. The minority strand (N-strand) encodes 14 genes, comprising 4 PCGs (*nad1*, *nad4*, *nad4L* and *nad5*), 8 tRNA genes and 2 rRNA genes (*rrnL* and *rrnS*). The gene arrangement in these four new species is consistent with that of other published Cerambycidae mitogenomes and follows the ancestral insect mitogenome arrangement. However, *Vesperus conicollis* Fairmaire

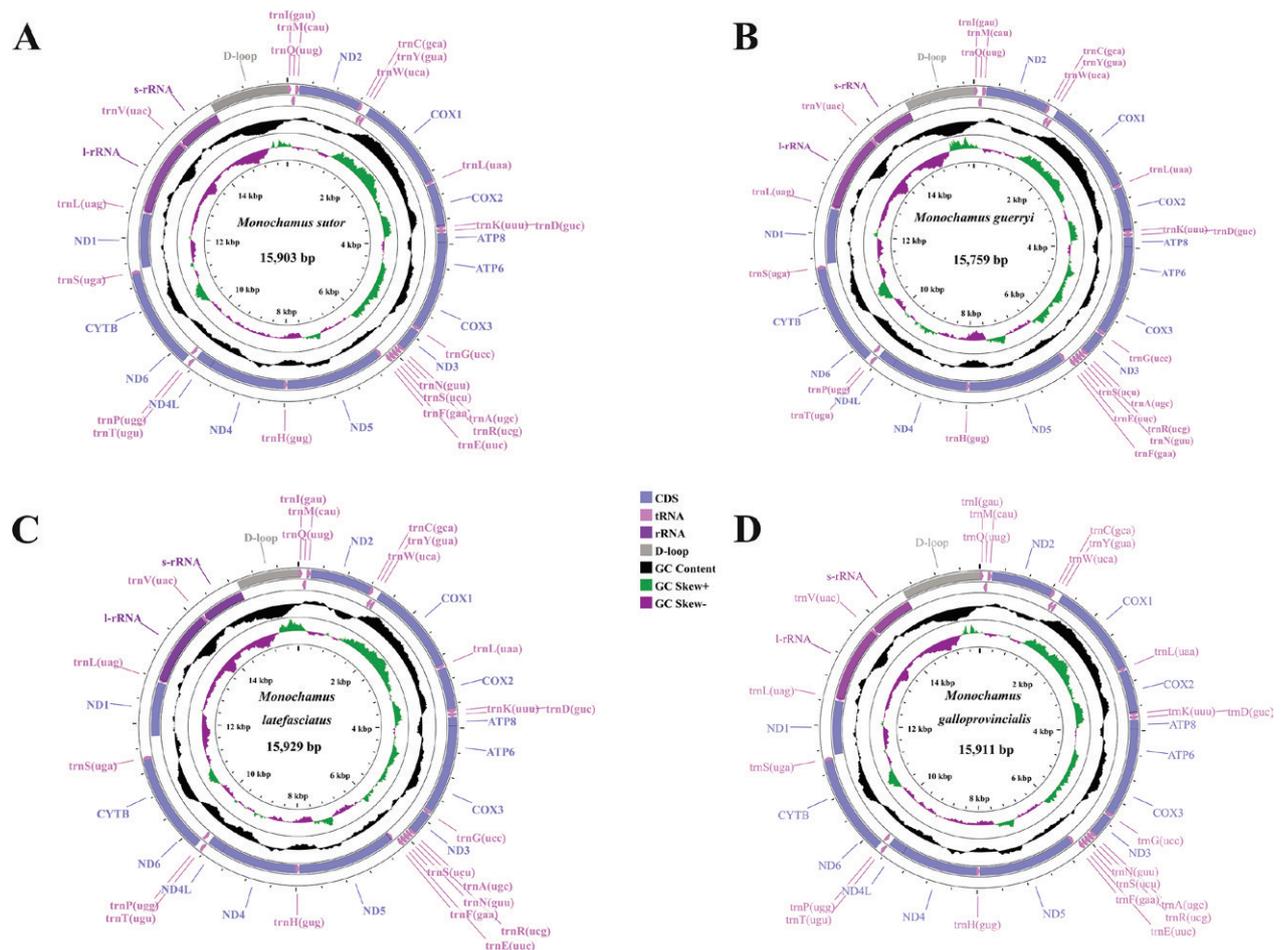


Figure 1. Complete mitochondrial genomes of *M. sutor* (A), *M. guerryi* (B), *M. latefasciatus* (C), and *M. galloprovincialis* (D). The transcriptional direction is indicated by arrows.

& Coquerel exhibited gene rearrangement, with *nad6* positioned ahead of *trnP*. And the translocation of *trnL1* and *trnL2* and a local inversion of *rrnS* were observed in *Apomecyna saltato* (Fabricius) and *Arhopalus unicolor* (Gahan), respectively (Fig. 10). These gene rearrangements mainly occur in the replication origin, contributing to species evolution acceleration. Overall, Cerambycidae mitochondrial genomes display a high degree of conservation, indicating relative stability in mitochondrial function (Liu et al. 2023).

3.2. Nucleotide composition of mitochondrial genomes

The mitochondrial genomes of the longhorn beetles display a noticeable AT bias, with an A+T content ranging from 67.9% (Prioninae, *Agrianome spinicollis* (MacLeay)) to 81.2% (Lepturinae, *Sachalinobia koltzei* (Heyden)). Interestingly, only four species exhibit an A+T content exceeding 80%. Lepturinae stands out with the highest A+T content, averaging 79.5%, while Prioninae has the lowest, averaging 70.8%. The AT-skew values are predominantly positive, ranging from -0.016 (Lamiinae, *Pterolophia* sp. ZJY-2019) to 0.125 (Prioninae, *Priotyranus closteroides* Thomson), while GC-skew values are mainly negative, spanning from -0.286 (Prioninae, *P. closteroides*) to -0.136 (Cerambycinae, *Xystrocera globosa* (Oliver)). Notably, the newly sequenced Lamiinae species all have A+T contents higher than the average, with positive AT skew. Although the nucleotide compo-

sitions across the 10 groups are generally similar, there are differences in A+T content, AT-skew, and GC-skew (Table S3; Fig. 2).

To explore the base content characteristic in the coding region, we specifically analyzed the PCGs of mitochondrial genomes. The results revealed a significant AT bias in the entire PCGs, with A+T content ranging from 65.42% (Prioninae, *A. spinicollis*) to 79.73% (Lepturinae, *S. koltzei*). The Lepturinae again exhibits the highest A+T content, averaging 77.9%, while the Prioninae has the lowest, averaging 67.9%, the same trends observed in the entire mitochondrial genome. The AT-skew values are negative across all PCGs, while most GC-skew values are also negative. This AT bias is particularly prominent at the 3rd codon position, consistent with a high mutational pressure toward AT nucleotides at this position (Meng et al. 2019). Notably, the 1st codon position shows a typical T, G bias, while the 2nd and 3rd codon positions show a typical T, C bias. Among the four newly sequenced species, both the entire PCGs and the second positions of these genes exhibit higher A+T content compared to the corresponding average values. Lepturinae maintains the highest A+T content, while Prioninae retains the lowest (Table S3). Sequences with high A+T content are generally less stable and more mutation-prone, potentially indicating adaptation to specific environments, such as extreme conditions, high altitudes, or high latitude areas (Song et al. 2016). This suggests that species within the Lepturinae may have a competitive advantage in adapting to particular habitats compared to those within the Prioninae. The AT content in the genus *Monochamus* is

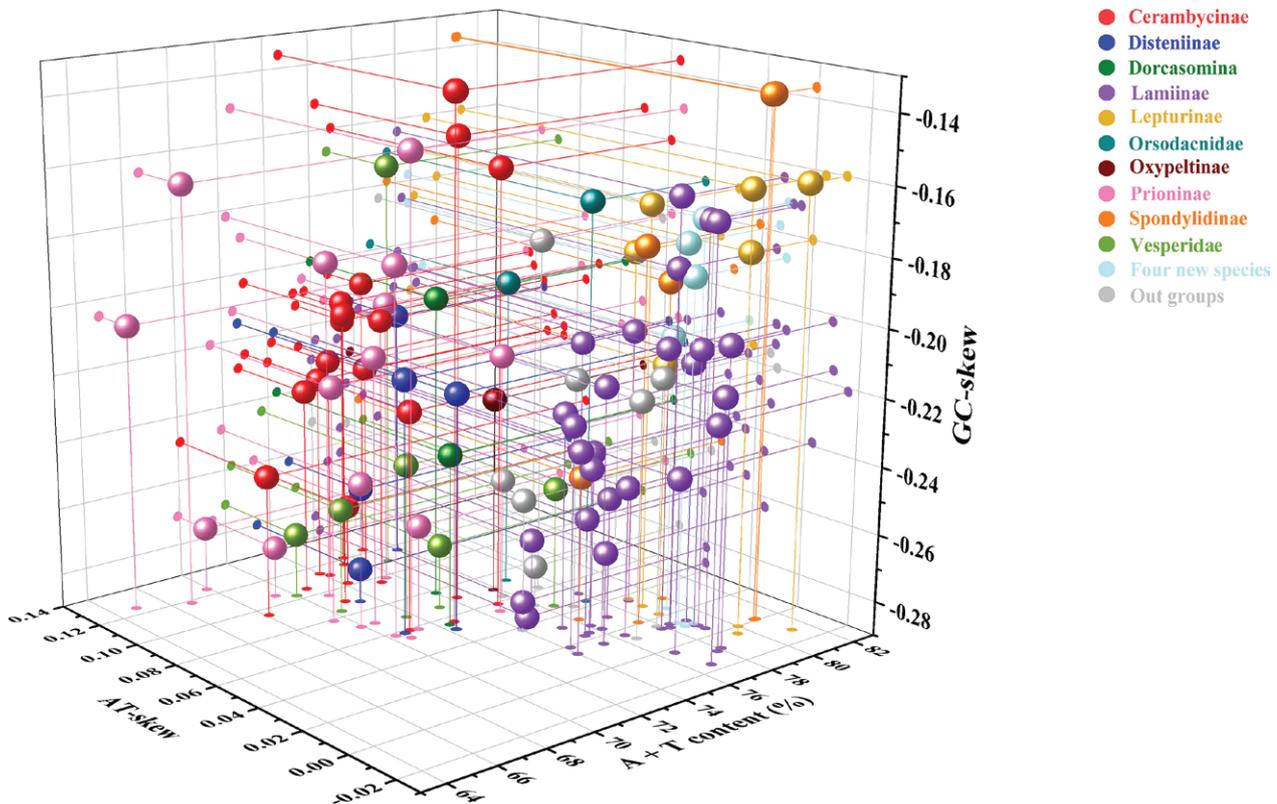


Figure 2. Three-dimensional scatter plots of the AT-skew, GC-skew, and A + T content of 86 mitochondrial genomes. Balls of different colors correspond to different groups.

generally similar, and similar AT bias may indicate shared evolutionary trajectories, especially in the case of the *M. alternatus*, *M. galloprovincialis*, and *M. urussovii*, where A+T content has reached a consistent level. This information enhances our understanding of mitochondrial genome functionality, evolution, and adaptive relationships within different environments.

Furthermore, there is a clear AT bias in RNA genes, with AT-skew and GC-skew values mainly positive in tRNA but predominantly negative in rRNA. Among the four newly sequenced beetles, the AT content in RNA genes is higher than the average, while the GC content is lower than the average (Table S3).

3.3. CGView comparison analysis

As depicted in Figure 3, the species obtained from the BLASTN analysis using the CCT and the mitochondrial genome of *M. sutor* were used as the reference species. The sequence similarity between *M. sutor* and other mitochondrial genomes ranged from 75.80% to 97.6%. Notably, *M. sutor* exhibited its closest relationship with *M. galloprovincialis*, followed by *M. urussovii*, *M. saltuarius*, and *M. alternatus*. Furthermore, the control region displayed the highest degree of variation, and there were varying levels of variation observed in genes such as *ND2*, *ND4*, *ND4L*, *ND5* and *ND6*. These differences pointed to

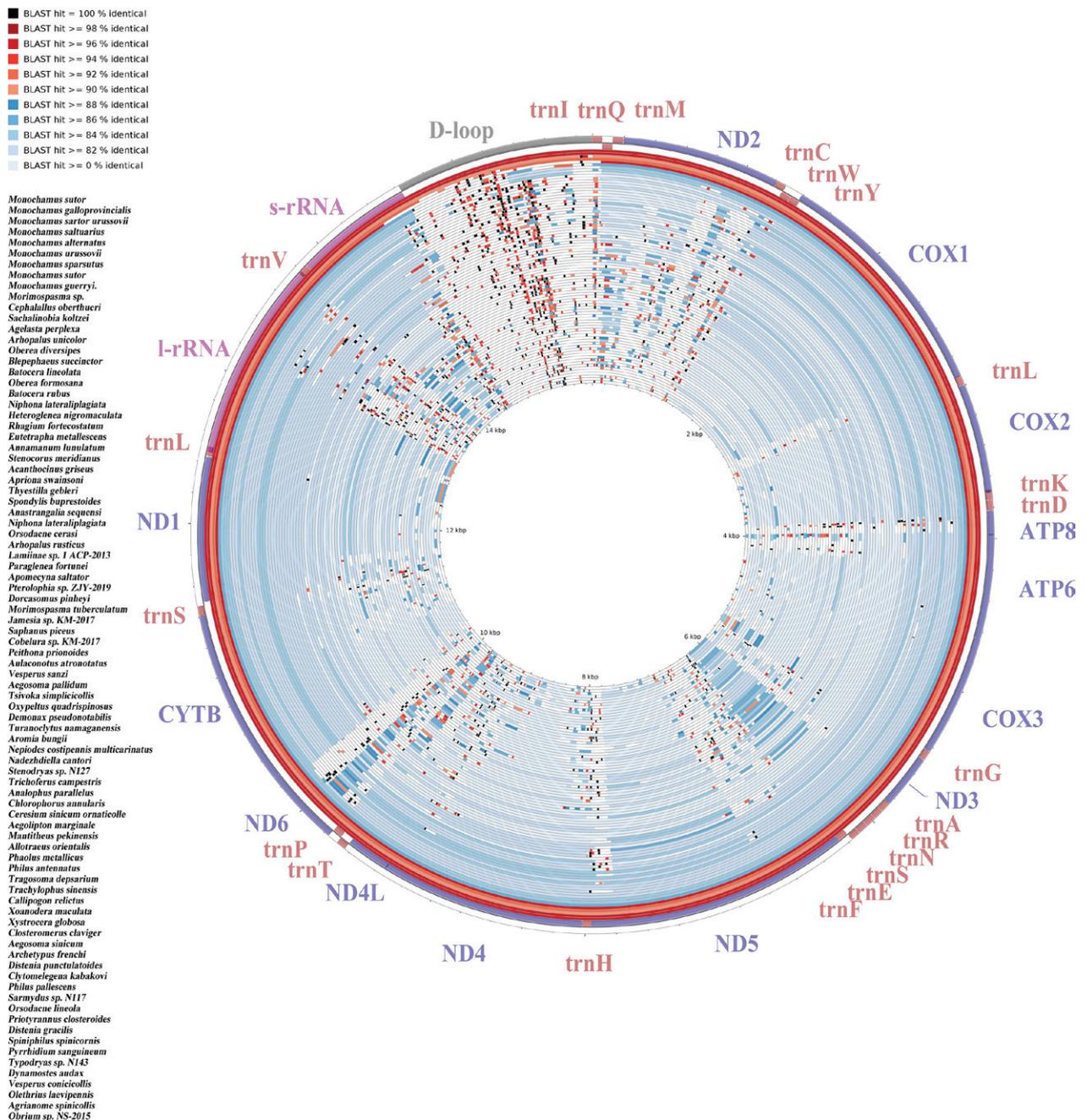


Figure 3. Graphical map of the BLASTN results showing the nucleotide identity between the *M. sutor* mitochondrial genome and that of 85 other species listed in Table S1. CCT arranges the BLASTN results such that the sequence most similar to the reference (*M. sutor*) is placed closer to the outer edge of the map.

the diversity among species and variations in evolutionary rates. In contrast, the *cox1* gene exhibited a lower degree of variation and was relatively more conservative.

3.4. Condon and amino acid usage

After analyzing the PCGs of 86 mitogenomes, we observed that the starting codon usage bias was quite similar at the species level when ATN codons were used as the start codons. Most species predominantly employed ATT or ATG as the start codon, while some species used ATA, ATC, and a few even utilized TTG and GGA. As for stop codons, TAA, TAG, and TTA were used in these mitogenomes, with TAA being the most frequently employed. At the gene level, most PCGs preferred ATT and ATG as start codons and still commonly used TAA as a stop codon. However, there was diversity in the types of start codons for NADH dehydrogenase genes, with some less common ones like TTG. Additionally, 8 genes exhibited incomplete stop codons in PCGs, with only *atp6*, *atp8*, *nad6*, and *nad1* being exceptions. In most cases, the incomplete stop codon was “T”, and very rarely “TA” (Table S5). These incomplete stop codons are relatively common in insect mitochondrial genomes (Wu et al. 2023). Overall, the start codon

patterns align with the frequent use of NNU and NNA in each codon.

Relative synonymous codon usage (RSCU) values of all 10 groups and the four new sequenced species were plotted in Fig. 4. The codon usage of PCGs exhibited a strong bias towards NNU and NNA codons. UUA (Leu2) had the highest RSCU in all samples, indicating that it had undergone strong selection during evolution. Conversely, the RSCU values for CGC (Arg), ACG (Thr), and CUG (Leu1), which end with G and C, were relatively low. Notably, Leu2 and Ser2 consistently possessed the highest RSCU values in all groups. These results demonstrated a connection between RSCU values and codon base selection bias (Table S6). Commonly used codons consisted of A and T, contributing to a higher A+T content in the mitochondrial genome. Furthermore, almost all frequently used codons ended with A or T, reinforcing the bias toward AT content in the third codon position (Song et al. 2016). In Dorcasominae, Prioninae, Cerambycinae and Disteniinae, the RSCU value of Leu2, particularly UUA in Leu2, distinctly decreased, while that of Leu1 increased. This could be a contributing factor to the lower AT content in these four groups (Zheng 2022). Overall, the differences in codon usage among different groups were minimal, suggesting that they might share a similar genetic evolutionary background.

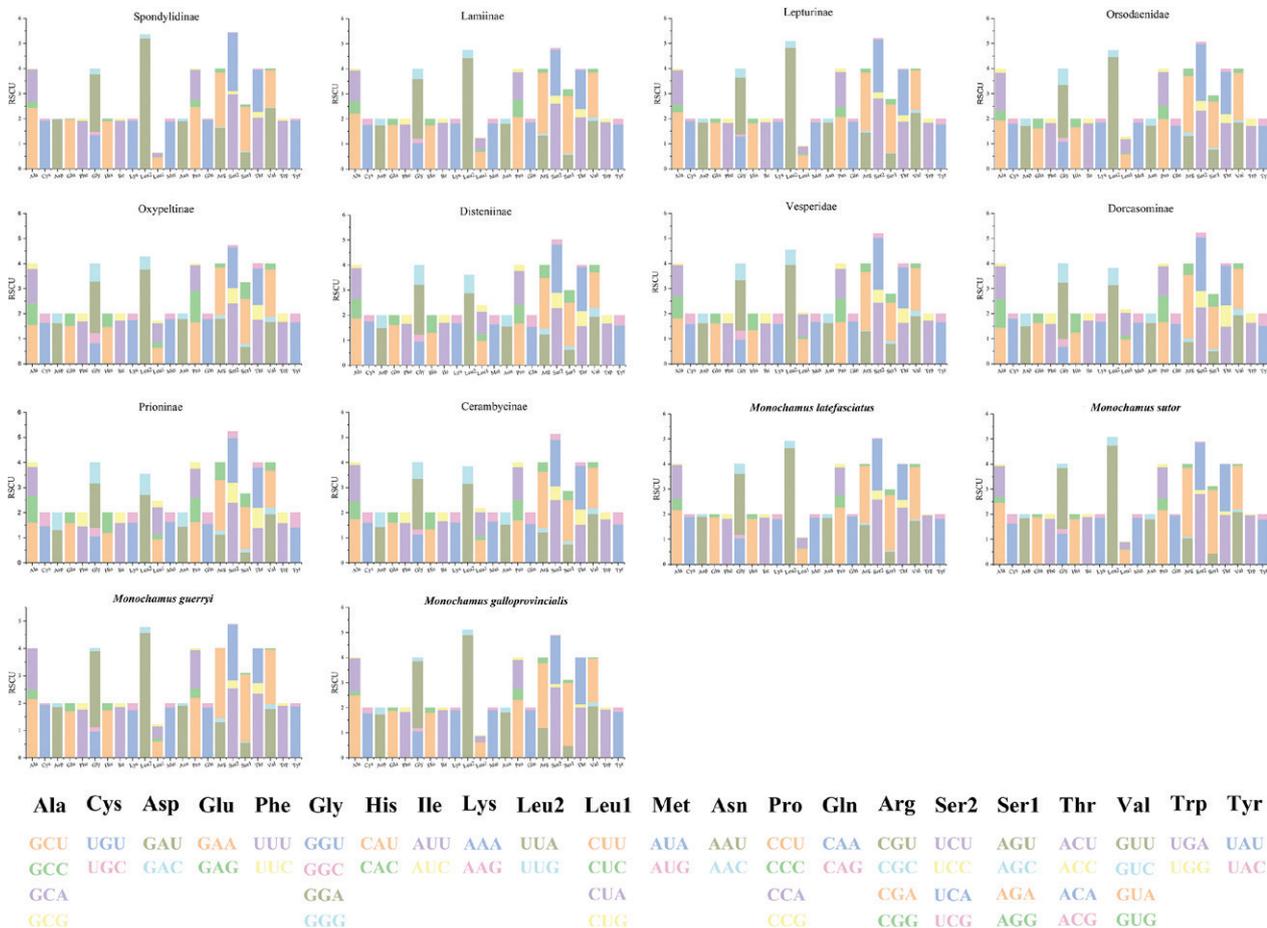


Figure 4. Relative synonymous codon usage (RSCU) in protein-coding genes of 10 groups and four new sequenced species. Codon families are indicated on the x-axis, along with different combinations of synonymous codons that code for specific amino acids. RSCU is shown on the y-axis.

In the 10 groups, the frequency of amino acid usage was relatively consistent. Overall, Leu was the most frequently used amino acid, followed by Ile, Phe, Met, and Ser in that order. However, some variations were observed in Spondylidinae, Prioninae, Cerambycinae and

Disteniinae. In these groups, there was a general trend of Ser being more frequently used than Met (Table S6; Fig 5). To sum it up, the frequency of amino acid usage in PCGs of Cerambycidae exhibited stability, with only minor differences in specific subfamilies.

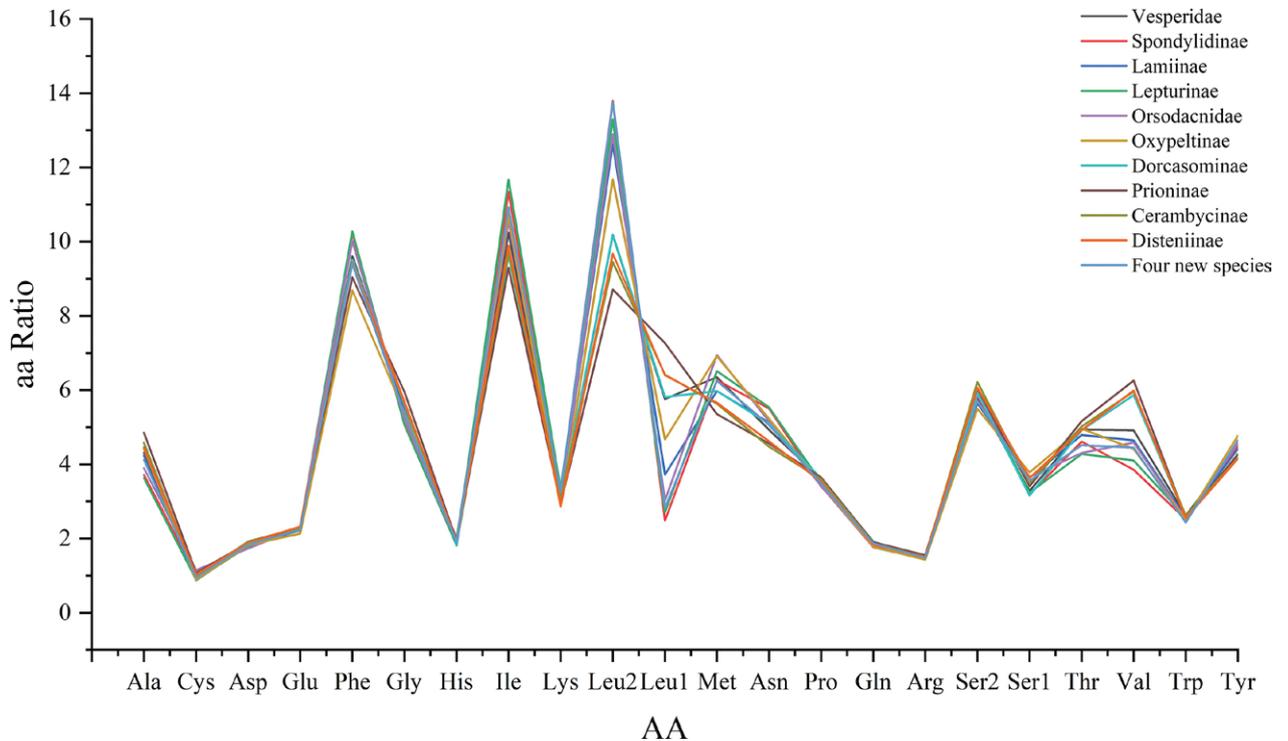


Figure 5. Amino acid usage in the mitochondrial genome. The different colored dashed lines represent different groups, the x-axis represents different amino acid types, and the y-axis represents the usage of amino acids.

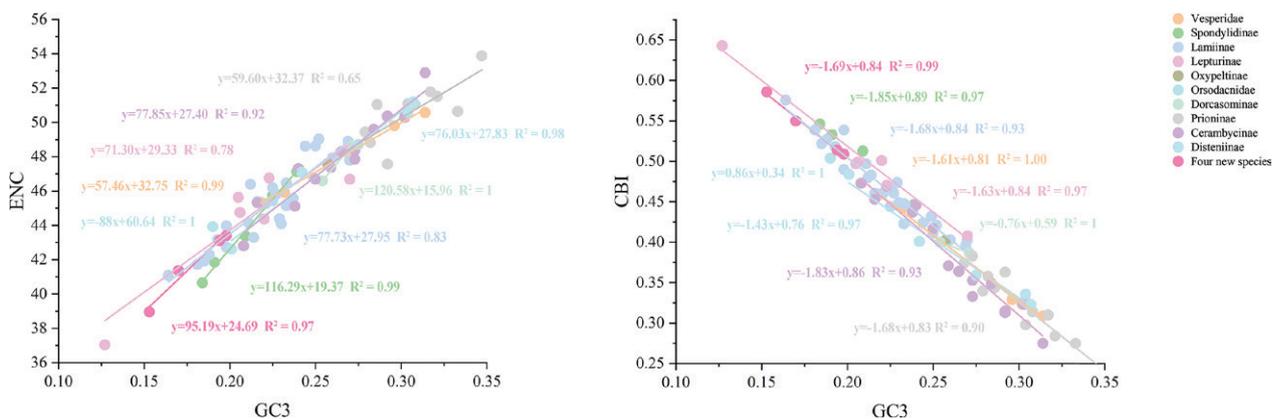


Figure 6. Nucleotide compositions of the mitogenomes. **A** Scatter plot of the GC content of 3rd codon sites versus ENC; **B** Scatter plot of the GC content of 3rd codon sites versus CBI.

The effective number of codons (ENC) and the codon bias index (CBI) in these species ranged from 37.03 (Lepturinae, *S. koltzei*) – 51.2 (Disteniinae, *Typodryas* sp. N143) for ENC, and from 0.275 (Prioninae, *P. closteroides*) – 0.643 (Lepturinae, *S. koltzei*) for CBI. Notably, Spondylidinae and Orsodacnadae had the lower average ENC value of 43.79 and 43.61, while Prioninae had the highest average ENC value of 50.32. As for CBI, Prion-

inae had the lowest average value of 0.33, while Spondylidinae, Lepturinae and Orsodacnadae had the higher average CBI value of 0.50 (Table S4). The ENC and CBI results showed that Spondylidinae and Orsodacnadae species exhibited a higher codon bias, suggesting that synonymous codons were more frequently used in their genomes with a lower codon selection, implying relatively lower genomic complexity. These Spondylidinae and

Table 2. The interval region of mitogenomes

Groups	Intergenic region amount	Gene overlap amount	<i>atp8-atp6</i> Normal intergenic region (-7bp) amount	<i>atp8-atp6</i> Special intergenic region	<i>nad4-nad4L</i> Normal intergenic region (-7bp) amount	<i>nad4-nad4L</i> Special intergenic region
Lamiinae	4-14	9-25	26	- 4, <i>H. nigromaculata</i> - 4, <i>Batocera rubus</i> (Linné) - 4, <i>Annamanum lunulatum</i> (Pic) - 4, <i>Pterolophia</i> sp. ZJY-2019 - 1, <i>Blepephaeus succinctor</i> (Chevrolat) - 4, <i>Morimospasma tuberculatum</i> Breuning	17	- 4, <i>Eutetrappa metallescens</i> (Motschulsky) - 4, <i>Paraglenea fortunei</i> (Saunders) - 4, <i>Thyestilla gebleri</i> (Faldermann) - 4, <i>H. nigromaculata</i> - 4, <i>B. rubus</i> - 4, <i>A. lunulatum</i> - 4, <i>M. latefasciatus</i> - 4, <i>M. guerryi</i> - 4, <i>M. sutor</i> - 4, <i>M. urussovii</i> - 4, <i>Agelasta perplexa</i> Pascoe - 4, <i>A. saltator</i> - 4, <i>Morimospasma</i> sp. - 4, <i>Niphona lateraliplagiata</i> Breuning - 4, <i>Jamesia</i> sp. KM-2017
Vesperidae	8-13	6-8	6	/	6	/
Spondylidinae	6-13	10-12	2	- 4, <i>C. oberthueri</i>	4	- 4, <i>C. oberthueri</i>
Lepturinae	6-9	9-13	6	/	5	- 4, <i>Anastrangalia sequensi</i> (Reitter)
Oxypeltinae	7	13	1	/	0	- 4, <i>Oxypeltus quadrispinosus</i> Blanchard
Orsodacnidae	7-11	6-10	2	/	2	/
Dorcasominae	6-9	10-11	2	/	1	- 4, <i>Tsvoka simplicicollis</i> (Gahan)
Prioninae	4-14	9-14	7	- 4, <i>P. closteroides</i> - 4, <i>Phaolus metallicus</i> (Newman) - 4, <i>Agrianome spinicollis</i> (MacLeay) - 4, <i>Analophus parallelus</i> Waterhouse - 4, <i>Archetypus frenchi</i> (Blackburn) - 4, <i>Olethrius laevipennis</i> Vitali	10	- 4, <i>P. closteroides</i> - 4, <i>Callipogon relictus</i> Semenov - 4, <i>Aegosoma sinicum</i> White
Cerambycinae	2-12	8-15	12	- 4, <i>Closteromerus claviger</i> Fairmaire - 4, <i>Demonax pseudonotabilis</i> Gressitt & Rondon - 4, <i>Ceresium sinicum</i> ornaticolle	12	- 0, <i>C. claviger</i> - 4, <i>D. pseudonotabilis</i> - 4, <i>Xystrocera globosa</i> (Olivier)
Disteniinae	5-7	5-11	5	/	5	/

Orsodacnidae species appear to possess strong adaptability to specific ecological environments (Li et al. 2023; Wu et al. 2023). Among the four newly sequenced species, *M. sutor* showed a relatively low ENC value of 38.95 and a relatively high CBI value of 0.59. These values indicate that *M. sutor* has strong environmental adaptability and more uniform codon usage (Li et al. 2023; Wu et al. 2023). Furthermore, the correlation analysis depicted in Fig. 6 revealed that ENC and CBI were positively and negatively correlated with GC3s, respectively. This suggests that the genomic G+C content plays a crucial role in determining codon preference across different species (Xu et al. 2020; Yan et al. 2021). The correlations between GC3 and both ENC and CBI are high across different groups except Prioninae. Higher GC3 levels in the genome may exert stronger selective pressure, leading to

more uniform codon usage (Wei et al. 2014). Moreover, elevated GC3 levels could result in specific functional genes displaying pronounced codon usage preferences with the principles of neutral mutational theories (Andersson et al. 1996).

3.5. Characteristics of intergenic regions

The number of intergenic regions in the mitochondrial genomes of the entire 10 groups ranged from 2 (Cerambycinae, *Ceresium sinicum ornaticolle* Pic) to 14 (Lamiinae, *Lamiinae* sp. 1 ACP-2013 and Prioninae, *P. closteroides*). On the other hand, gene overlap regions varied from 5 (Disteniinae, *Distenia punctulatoidea* Hubweber) to 25

(Lamiinae, *Obera diversipes* Pic). These gene overlaps were generally small, often spanning from 1 to 7 bp. The largest overlap regions were frequently observed between *atp6* and *atp8* (Wei 2009; Zheng et al. 2022), as well as between *nad4* and *nad4L*. The lengths of the intergenic regions displayed significant variation, with the longest intergenic region reaching up to 1813 bp (Lamiinae, *Jamesia* sp. KM-2017). Among the mitochondrial genomes analyzed in the beetle family, a total of 69 instances of a 7 bp intergenic region between *atp6* and *atp8* were observed, and 62 instances of a 7 bp gene intergenic region between *nad4* and *nad4L* were identified (Table 2; Table S5).

The genes *atp6*, *atp8*, *nad4*, and *nad4L* are likely to play crucial roles in essential cellular processes like energy production and oxidative phosphorylation (Pogoda et al. 2018; Yang et al. 2019). The presence of special intergenic regions between these genes may reflect their functional relationships and coordination (Zhang et al. 2013). Additionally, these distinctive intergenic regions could have implications for the adaptation of insect species to specific environments or ecological conditions (Yévenes et al. 2022).

3.6. Rate of mitogenomes evolution

As depicted in Fig. 7, within these groups, certain mitochondrial genes displayed distinct rates of evolution. Specifically, *nad6* (0.354), *atp8* (0.335), *nad3* (0.312), and *nad2* (0.300) were identified as fast-evolving genes, whereas *cox1* (0.191), *cox2* (0.202), *nad1* (0.207) and *atp6* (0.207) exhibited a slower rate of evolution. Genes characterized by smaller genetic distances are indicative of relatively stable functions and fewer accumulated mutations within their sequences (Nei 1978). In groups, it was noted that *nad6* consistently exhibited the higher genetic distances, while *cox1* consistently displayed the lower genetic distances, except for Spondylidinae. In Cerambycidae, all calculated values of ka/ks were less

than 1, ranging from 0.067 (*cox1*) to 0.441 (*atp8*). This indicates a higher likelihood of retaining synonymous mutations within these genes. Notably, genes such as *atp8* (0.441), *nad4L* (0.318), and *nad6* (0.300) exhibited a high rate of evolution, whereas *cox1* (0.067), *cox2* (0.094) and *cytb* (0.110) showed a low rate of evolution. Further analysis of ka/ks revealed that *atp8* and *nad6* had higher ka/ks ratios, suggesting that these genes have experienced relaxed selection pressures. In contrast, *cox1* and *cox2* displayed lower ka/ks ratios, indicating a history of strong purifying selection acting upon them (Sun et al. 2020). Combining the findings from both Ka/Ks and genetic distance analyses, it became evident that *cox1* and *cox2* are exceptionally conserved genes, relatively unaffected by natural selection (Lynch et al. 2000; Wertheim et al. 2015). The *cox1* gene, in particular, is employed as a DNA barcode for species identification, leveraging the high level of conservatism in its sequence (Yévenes et al. 2022; Fonseca et al. 2022). This utility is particularly valuable in biodiversity surveys and species identification, aiding in distinguishing species that are otherwise challenging to discern. Conversely, genes such as *atp8*, *nad4L*, and *nad6* play crucial roles in cellular respiration and energy metabolism pathways (Hui et al. 2018; Yang et al. 2021). They necessitate ongoing variation and evolution to adapt to diverse ecological environments, resulting in greater sequence variability.

3.7. tRNA and rRNA

All four newly sequenced species exhibited a complete set of 22 tRNA genes, with entire tRNA lengths 1441 (*M. latefasciatus*), 1444 (*M. guerryi*), 1452 (*M. sutor*), 1452 (*M. galloprovincialis*). When individually examining these tRNAs, their lengths ranged from 61 to 70 in *M. latefasciatus*, from 63 to 70 in *M. guerryi*, from 63 to 69 in *M. sutor*, and from 63 to 69 in *M. galloprovincialis*, respectively. Except for trnS1, all tRNAs exhibited the

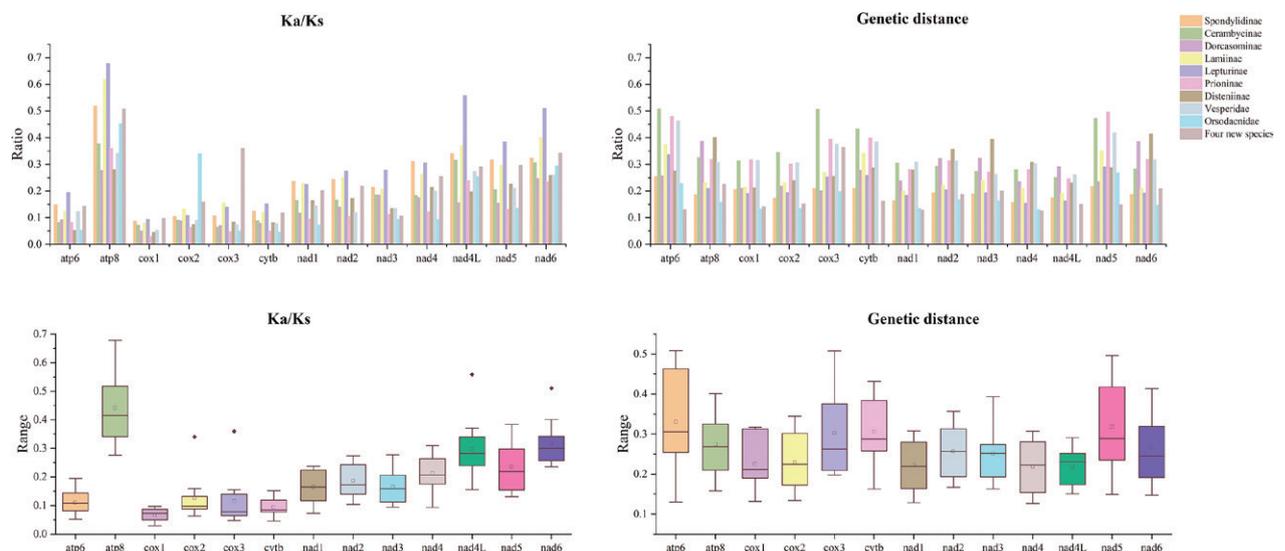


Figure 7. Mean Ka/Ks (A, B) and Genetic distances (C, D) (Kimura-2-parameter) of mitochondrial gene sequences in groups. \pm SD, hollow dots; Median, middle line inside each box; IQR (interquartile range), the box containing 50% of the data;

characteristic cloverleaf structure, where the dihydrouridine arm of trnS1 formed as a single loop. This peculiar structure of trnS1 may arise from abnormalities during transcription or post-transcriptional processes, which could subsequently affect the development of its typical cloverleaf structure (Feng et al. 2020) (Table S5; Fig 8). Such an occurrence has been frequently observed within the longhorn beetle genus *Monochamus*. The absence of the standard cloverleaf configuration implies functional versatility for mitochondrial trnS1, potentially enabling diverse roles in various biological processes (Garey et al. 1989; Wolstenholme et al. 1992). Nevertheless, this also indicates a deviation from the traditional functionality associated with a typical tRNA.

The tRNA structures in all four newly sequenced beetles exhibited a cloverleaf-like secondary structure overall; the phenomenon of tRNA-Ser1 lacking the DHU arm was present in all four newly sequenced beetles. Additionally, some degree of base mismatch occurred, with the most common type being the G-U mismatch.

Specifically, there were 25 (21 G-U, 3 U-U, 1 A-G) mismatched base pairs in *M. latefasciatus*, 24 (20 G-U, 3 U-U, 1 A-G) mismatched base pairs in *M. guerryi*, 20 (15 G-U, 4 U-U, 1 A-G) mismatched base pairs in *M. sutor*, 21 (16 G-U, 4 U-U, 1 A-G) mismatched base pairs in *M. galloprovincialis* (Fig. 8). The absence of the DHU arm and the occurrence of base mismatches in mitochondrial tRNA may result from various factors, including adaptive changes during evolution, genetic mutations, and selective pressures (Ji et al. 2023). The DHU arm played a role in the stability and function of the tRNA molecule, and its absence could potentially impact the stability and proper folding of tRNA-Ser1, thereby affecting its role in protein synthesis. Base mismatches may lead to mutations in genes encoding proteins, accumulating across generations and influencing the genetic traits of the population (Acuna-Hidalgo et al. 2016). Overall, the structural variations observed in these four newly sequenced beetles were consistent and largely consistent with the tRNA structures found in most species.

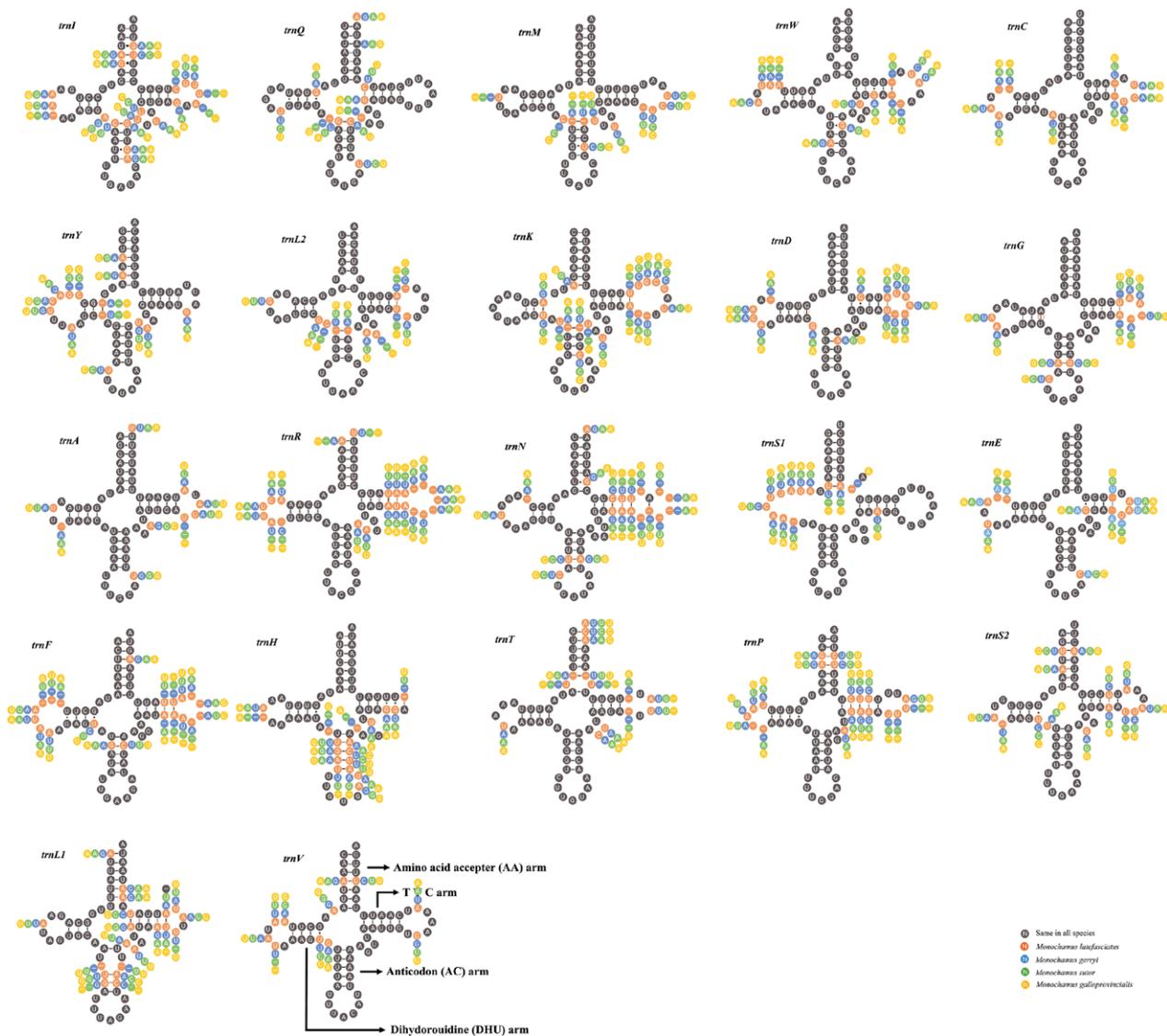


Figure 8. Putative secondary structures of tRNAs from the *M. sutor*, *M. guerryi*, *M. galloprovincialis*, and *M. latefasciatus* mitochondrial genome. Dashes indicate the Watson–Crick base pairs; dots indicate wobble GU pairs; and the other non-canonical pairs are not marked.

The rRNA components in the mitochondrial genomes of the four newly sequenced species encompass both lrRNA and srRNA. Specifically, lrRNA is situated between tRNA-Leu and tRNA-Val, while srRNA is positioned between tRNA-Val and the control region. The length of lrRNA genes ranged from 1329 bp in *M. guerryi* to 1334 bp in *M. latefasciatus*, while the sizes of srRNA genes varied from 775 bp in *M. latefasciatus* to 781 bp in *M. guerryi*, *M. galloprovincialis*, and *M. sutor* (Table S5).

3.8. Substitution saturation tests and nucleotide heterogeneity

The substitution saturation index (Iss) for the six nucleotide sequence datasets was notably lower than the critical Iss values (Iss.cSym and Iss.cAym) for both symmetrical

and asymmetrical trees. This observation strongly indicated the absence of substitution saturation, providing clear evidence of a robust phylogenetic signal (all Iss < Iss.cSym or Iss.cAym, $P < 0.05$) (Table 3, Fig. S1) (Yan et al. 2021).

To assess nucleotide divergence heterogeneity, pairwise comparisons were conducted within a multiple sequence alignment. The results revealed that datasets involving the third codon position (PCG123, PCG123R) exhibited lower heterogeneity in sequence composition. In comparison, datasets with the third codon position removed (PCG12, PCG12R) displayed intermediate heterogeneity levels. The AA dataset exhibited the highest heterogeneity. It's worth noting that some species, such as *C. claviger*, *Dynamostes audax* Pascoe, *H. nigromaculata*, *Oberea formosana* Pic, and *V. conicicollis*, exhibited greater heterogeneity due to incomplete sequences caused by missing data (Fig. 9) (Li et al. 2020).

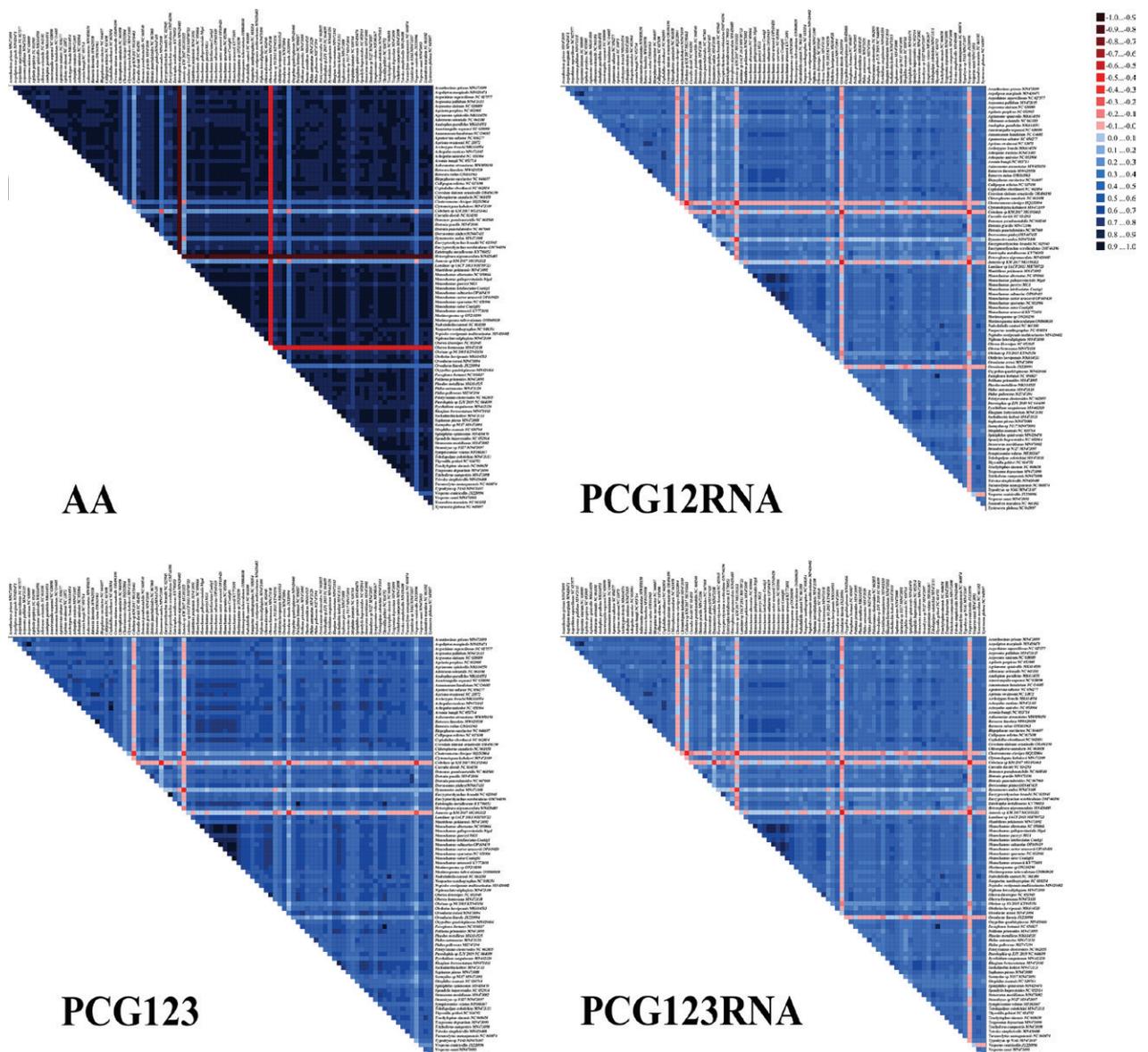


Figure 9. Heterogeneity of the sequence composition of the mitochondrial genomes in different datasets. The pairwise Aliscore values are indicated by colored squares. Darker colors indicate full random similarity, and lighter colors indicate the opposite. All taxon names are listed on the left side of the heat map.

Table 3. Substitution saturation test results.

Data partition	Iss	Iss.cSym [†]	Psym [‡]	Iss.cAsym [§]	Pasym [¶]
PCG1	0.458	0.818	0.0000	0.572	0.0000
PCG12	0.458	0.818	0.0000	0.572	0.0000
PCG123	0.462	0.818	0.0000	0.572	0.0000
PCG1+rRNA	0.461	0.818	0.0000	0.572	0.0000
PCG12+rRNA	0.455	0.818	0.0000	0.572	0.0000
PCG123+rRNA	0.602	0.819	0.0000	0.573	0.0000

3.9. Phylogenetic analyses

We selected a total of 86 species of longhorn beetles and analyzed four datasets using IQ-TREE and Phylobayes two methods. The results indicated that most nodes within the phylogenetic trees displayed strong support values, indicating robust branching patterns (Fig. 10, Fig. S2). The overall topology of the phylogenetic trees remained largely consistent, with only a few branches showing instability.

In these phylogenetic trees, results obtained from the AA dataset and the PCG123 dataset both constructed

using IQ-TREE are notably similar with relatively high node support (Fig. 10, Fig. S2). Through analyzing all the phylogenetic trees, we observed that Cerambycinae and Prioninae are sister groups in all of the tree. And the position of *V. conicicollis* within Vesperidae exhibited notable instability across all phylogenetic trees. Specifically, in the results generated by IQ-TREE, *V. conicicollis* appeared to be distantly related to the other sequenced longhorn beetles. In AA dataset, the relationships among Disteniinae, Cerambycinae and Prioninae remained consistent across both methods: (Disteniinae+(Cerambycinae+Prioninae))

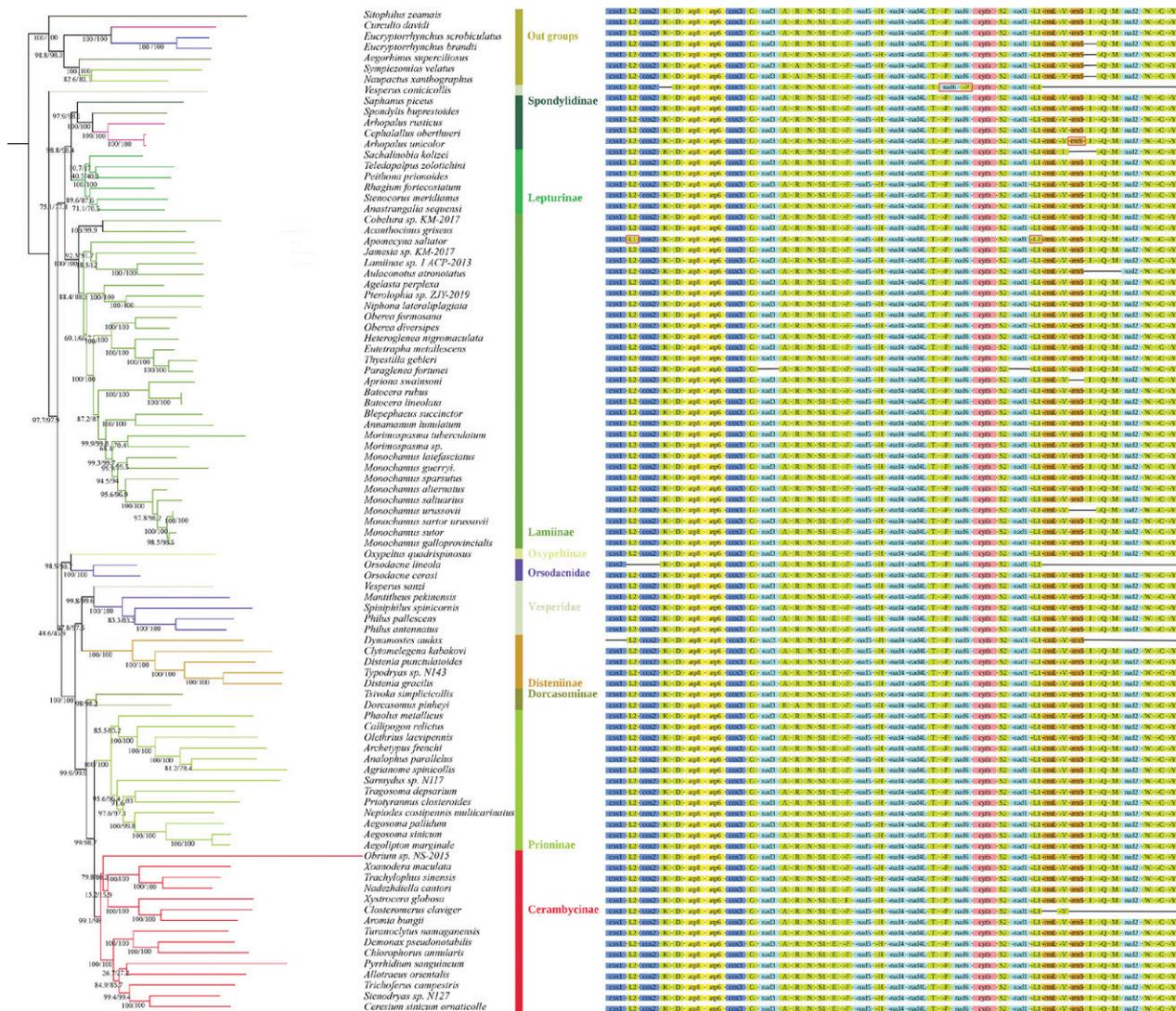


Figure 10. Phylogenetic tree inferred from the PCG123 and AA dataset using maximum likelihood (ML) and arrangements of genes in the mitochondrial genomes. Numbers beside each node are support values. The values indicate ML bootstrap values (BPs).

(BPs = 99, PPs = 0.99). The Orsodacnidae, Oxypeltinae and Vesperidae relationship is unstable. In PCG12RNA, Dorcasominae were highly supported, but the phylogenetic positions of them were unstable. This may be attributed to the extremely limited amount of data available. In PCG123 dataset and PCG123RNA dataset, Orsodacnidae and Oxypeltinae both form sister group but with different node support. In a comparative analysis of the results from eight phylogenetic trees, we observed that the tree constructed by the IQ-TREE method using the AA dataset is better, demonstrating relatively high node support and strong consistency of phylogenetic relationships. The results in this tree revealed four major clades, which are consistent with Nie et al., except for the 'lamiinae' clade in our study is ((Lepturinae + Spondylidinae) + Lamiinae), while it is ((Lamiinae + Spondylidinae) + Lepturinae) in Nie et al. (Nie et al. 2020) (Fig. 10, Fig. S2).

Through constructing phylogenetic tree and the analysis of mitochondrial genome data presented in this study, we have presented following perspectives on some contentious issues. The consistent clustering of Spondylidinae and Aseminae in the phylogenetic trees supports the idea that these two groups may be treated as a single entity, Spondylidinae, with some cases of synonymous nomenclature (Linsley et al. 1961; Gressitt et al. 1970; Nakamura et al. 1981; Napp et al. 1994). In addition, some other scholars believe that Oxypeltidae is a family at the same taxonomic level as Cerambycidae (Monne 2012). The morphological identification of Orsodacnidae suggests that it is an independent family within the order Coleoptera (Li 2018). In the phylogenetic results of current study, Oxypeltidae and Orsodacnidae consistently form a single branch. This indicates that they are relatively closely related and their taxonomic status may be the same. However, in our phylogenetic tree analysis, there are still some unresolved issues. For instance, the phylogenetic relationships of certain groups remain unresolved (Vesperidae, Oxypeltinae, Orsodacnidae, and Dorcasominae), and the support values at key nodes are not notably high. Despite employing various methods and constructing different datasets, the results are consistently similar. These challenges are also observed in the study by Nie et al. (Nie et al., 2020). Therefore, we propose that future research should focus on expanding mitochondrial genome sequencing to include a more extensive range of species, thereby enriching our sample size. Additionally, exploring alternative tree-building methods may be beneficial to identify an approach better suited for the phylogenetic analysis of the longhorn beetles. It is essential to note that, although we, like Nie et al. have related to research on phylogeny of the longhorn beetles, our methods for constructing the phylogenetic tree and the focal points of our research are different. Nie et al. chose a more extensive set of species, offering new insights into the phylogeny of longhorn beetles. In contrast, our objective is to employ diverse methods to analyze mitochondrial genome data from multiple perspectives. Our aim is to conduct comparative analyses of these datasets, providing additional support for the phylogeny of the longhorn beetles. And the results we obtained in the phylogenetic section are generally consistent.

In two methods employed for the analysis, the four newly sequenced beetles consistently grouped within the Lamiinae. This clustering demonstrates their close genetic affinity with other species within the Lamiinae. The proximity of closely related species within the same subfamily is evident in the phylogenetic trees, indicating a shared evolutionary history and genetic similarity among these species (Fig. 10, Fig. S2). In 2022, Shi et al. conducted a comparative analysis of mitochondrial genomes within Lamiinae. Although both studies focus on the mitochondrial genome analysis of longhorn beetles, it's crucial to note that our research targets differently. While the study by Shi et al. concentrates on Lamiinae, the scope of this article encompasses the entire longhorn beetles families or subfamilies. Additionally, new content has been incorporated into the analysis methods, including analyses such as CBI and ENC. In our study, the research results regarding Lamiinae align with those of Shi et al. Apart from a few individual species that are categorized under Lamiinae in the Shi et al., but our research indicates that they belong to Spondylidinae, such as *A. unicolor*. Furthermore, compared to other groups, we observed that the phylogenetic relationships within Lamiinae are relatively stable, and they exhibit a close affinity with Spondylidinae and Lepturinae.

4. Conclusion

This study sequenced four species of Cerambycidae and downloaded 82 sequences of longhorn beetles from platforms such as NCBI, resulting in a dataset comprising 86 sequences. These sequences were employed in a comparative investigation of mitochondrial genomes within the longhorn beetles. Various aspects of genetic codons were analyzed, coupled with the construction of phylogenetic trees using two distinct methodologies and four distinct datasets. The study yielded several noteworthy conclusions: Firstly, at the gene level, multiple analytical approaches consistently showed that *cox1* and *cox2* exhibited the least variation and were relatively conservative in the overall evolutionary process. On the contrary, genes associated with respiration and energy production, such as *atp6*, *atp8*, *nad4*, displayed comparatively heightened genetic variations, indicative of their adaptability throughout evolutionary history (Hancock et al. 2010). Secondly, at the species level, species within the Lepturinae, Spondylidinae displayed strong adaptability in specific ecological environments and exhibited a competitive advantage in adapting to specific habitats. Moreover, this study contributed valuable insights into the phylogeny of longhorn beetles. Based on the genetic data and the outcomes of phylogenetic analyses, some of the previously proposed viewpoints have been validated in phylogenetic relationships. In summation, the comparative analysis of mitochondrial genomes in Cerambycidae and related families in the Chrysomeloidea significantly contributes to a more comprehensive understanding of

the diversity of longhorn beetles and the complexity of their internal phylogenetic relationships. The foundation was laid for a more in-depth systematic study in longhorn beetles.

5. Acknowledgment

This study was carried out at Beijing Key Laboratory for Forest Pest Control, Beijing Forestry University. The work was supported by the National Key Research & Development Program of China (2021YFD1400900). We thank the editor and reviewer for the valuable comments and suggestions. We thank TopEdit (www.topeditsci.com) for linguistic assistance during the preparation of this manuscript.

6. References

- Andersson SGE, Sharp PM (1996) Codon usage and base composition in *Rickettsia prowazekii*. *Journal of molecular evolution* 42: 525–536. <https://doi.org/10.1007/BF02352282>
- Acuna-Hidalgo R, Veltman JA, Hoischen A (2016) New insights into the generation and role of de novo mutations in health and disease. *Genome biology* 17: 1–19. <https://doi.org/10.1186/s13059-016-1110-1>
- Behura SK (2006) Molecular marker systems in insects: current trends and future avenues. *Molecular ecology* 15(11): 3087–3113. <https://doi.org/10.1111/j.1365-294X.2006.03014.x>
- Boore JL (1999) Animal mitochondrial genomes. *Nucleic acids research* 27(8), 1767–1780
- Brewster D (1832) *The Edinburgh Encyclopedia*. Joseph and Edward Parker; William Brown, Vol. 15.
- Chen Y, Zhou YT, Sun H, Wang Y, Xu ZT, Li XD (2023) Occurrence of major forest pests in China in 2022 and trend prediction in 2023. *Forest Pest and Disease* 42(02): 51–54. <https://cnki.issn1671-0886.20230007>
- Clary DO, Wolstenholme DR (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *Journal of Molecular Evolution* 22(3): 252–271. <https://doi.org/10.1007/BF02099755>
- El-Shafie HA (2021) The longhorn beetle *Jebusaea hammerschmidti* Reiche (Coleoptera: Cerambycidae): An old serious pest undermining date palm plantations. *CABI Reviews*.
- Fabricius JC (1801) *Systema eleutheratorum secundum ordines, genera, species: adiectis, synonymis, locis, observationibus, descriptionibus*. Impensis Bib-liopolii Academici Novi, Vol. 1 2: 3–687
- Feng J, Guo Y, Yan C, Ye Y, Li J, Guo B, Lü Z (2020) Comparative analysis of the complete mitochondrial genomes in two limpets from Lottiidae (Gastropoda: Patellogastropoda): rare irregular gene rearrangement within Gastropoda. *Scientific Reports* 10(1): 19277. <https://doi.org/10.1038/s41598-020-76410-w>
- Folmer O, Black M, Hoeh W, Lutz R. & Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotech* 3(5): 294–299
- Footitt RG, Adler PH (2018) *Insect biodiversity: science and society*, volume II.
- Freeland JR (2020) *Molecular ecology*. John Wiley & Sons.
- Fonseca VG, Kirse A, Giebner H, Vause BJ, Drago T, Power DM, Peck LS, Clark MS (2022) Metabarcoding the Antarctic Peninsula biodiversity using a multi-gene approach. *ISME Communications* 2(1): 37. <https://doi.org/10.1038/s43705-022-00118-3>
- Garey JR, Wolstenholme DR (1989) Platyhelminth mitochondrial DNA: evidence for early evolutionary origin of a tRNA ser AGN that contains a dihydrouridine arm replacement loop, and of serine-specifying AGA and AGG codons. *Journal of Molecular Evolution* 28: 374–38. <https://doi.org/10.1007/BF02603072>
- Geoffroy ÉL (1762) *Histoire abrégée des insectes qui se trouvent aux environs de Paris: dans laquelle ces animaux sont rangés suivant un ordre méthodique*. Durand, Vol. 2.
- Gressitt JL, Rondon JA. & von Breuning S (1970) Cerambycid-beetles of Laos (Longicornes du Laos). *Pacific Insects* 24: 1–651 pp.
- Hancock L, Goff L, Lane C (2010) Red algae lose key mitochondrial genes in response to becoming parasitic. *Genome biology and evolution* 2: 897–910. <https://doi.org/10.1093/gbe/evq075>
- Hui M, Wang M, Sha Z (2018) The complete mitochondrial genome of the alvinocaridid shrimp *Shinkaicaris leurokolos* (Decapoda, Caridea): Insight into the mitochondrial genetic basis of deep-sea hydrothermal vent adaptation in the shrimp. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 25: 42–52. <https://doi.org/10.1016/j.cbd.2017.11.002>
- Ji YT, Zhou XJ, Yang Q, Lu YB, Wang J, Zou JX (2023) Adaptive evolution characteristics of mitochondrial genomes in genus *Aparapotamon* (Brachyura, Potamidae) of freshwater crabs. *BMC genomics* 24(1): 1–24. <https://doi.org/10.1186/s12864-023-09290-9>
- Johnson KP (2019) Putting the genome in insect phylogenomics. *Current opinion in insect science* 36: 111–117. <https://doi.org/10.1016/j.cois.2019.08.002>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* 30: 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermini LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* 14: 587–589. <https://doi.org/10.1038/nmeth.4285>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution* 35: 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Kück P, Meid SA, Groß C, Wägele JW, Misof B (2014) AliGROOVE – visualization of heterogeneous sequence divergence within multiple sequence alignments and detection of inflated branch support. *BMC Bioinformatics* 15: 294. <https://doi.org/10.1186/1471-2105-15-294>
- Lanfear R, Frandsen PB, Wright AM, Senfeld T and Calcott B (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular biology and evolution* 34: 772–773. <https://doi.org/10.1093/molbev/msw260>
- Lartillot N, Rodrigue N, Stubbs D, Richer J (2013) PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic biology* 62: 611–615. <https://doi.org/10.1093/sysbio/syt022>
- Letunic I, Bork P (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic acids research* 47(W1): W256–W259. <https://doi.org/10.1093/nar/gkz239>
- Latreille PA (1829) *Les crustacés, les arachnides et les insectes: distribués en familles naturelles: ouvrage formant les tomes 4 et 5 de celui de M. le baron Cuvier sur Le règne animal (deuxième édition)*. Loth ed 1: 54.

- Librado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, *Bioinformatics*, pp. 1451–1452.
- Li K, Liang H (2018) A comparative study of external female genitalia (including the 8th and 9th abdominal segments) in the family Megalopodidae and other related families of Chrysomeloidea. *Zookeys* 30(762): 69–104. <https://doi.org/10.3897/zookeys.762.22163>
- Li X, Yan L, Pape T, Gao Y, Zhang D (2020) Evolutionary insights into bot flies (Insecta: Diptera: Oestridae) from comparative analysis of the mitochondrial genomes. *International Journal of Biological Macromolecules* 149: 371–380. <https://doi.org/10.1016/j.ijbiomac.2020.01.249>
- Liu Q, Cai YD, Ma L, Liu HG, Linghu T, Guo SK, Wei SJ, Song F, Tian L, Cai WZ, Li H (2023) Relaxed purifying selection pressure drives accelerated and dynamic gene rearrangements in thrips mitochondrial genomes. *International Journal of Biological Macromolecules* 253: 126742. <https://doi.org/10.1016/j.ijbiomac.2023.126742>
- Li Q, Luo Y, Sha A, Xiao W, Xiong Z, Chen X, He J, Peng L, Zou L (2023) Analysis of synonymous codon usage patterns in mitochondrial genomes of nine *Amanita* species. *Frontiers in Microbiology* 14: 1134228. <https://doi.org/10.3389/fmicb.2023.1134228>
- Linnaeus C (1758) *Systema naturae*. Laurentii Salvii: Stockholm, pp. 824.
- Linsley EG, & Chemsak JA (1961) *The Cerambycidae of North America* Vol. 102. University of California Press.
- Lowe TM, Chan PP (2016) tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 44: W54–W57. <https://doi.org/10.1093/nar/gkw413>
- Lynch M & Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290(5494): 1151–1155. <https://doi.org/10.1126/science.290.5494.1151>
- MacLeod A, Evans HF, & Baker RHA (2002) An analysis of pest risk from an Asian longhorn beetle (*Anoplophora glabripennis*) to hardwood trees in the European community. *Crop Protection* 21(8): 635–645. [https://doi.org/10.1016/S0261-2194\(02\)00016-9](https://doi.org/10.1016/S0261-2194(02)00016-9)
- Meng G, Li Y, Yang C, Liu S (2019) MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic acids research* 47(11): e63. <https://doi.org/10.1093/nar/gkz173>
- Monné M, Wang Q (2017) *General morphology, classification and biology of Cerambycidae*. CRC Press.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler, A (2020) IQ-TREE, R. L.2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution* 37(5): 1530–1534. <https://doi.org/10.1093/molbev/msaa131>
- Monne MA (2012) Catalogue of the type-species of the genera of the Cerambycidae, Disteniidae, Oxypeltidae and Vesperidae (Coleoptera) of the Neotropical Region. *Zootaxa* 3213(1): 1–183. <https://doi.org/10.11646/ZOOTAXA.3213.1.1>
- Napp DS (1994) Phylogenetic relationships among the subfamilies of Cerambycidae (Coleoptera, Chrysomeloidea). *Revista Brasileira de Entomologia* 38(2): 265–419.
- Nakamura S (1981) Morphological and taxonomic studies of the cerambycid pupae of Japan: (Coleoptera: Cerambycidae). *Miscellaneous Reports of the Hiwa Museum for Natural History* 20: 1–159.
- Nie R, Vogler AP, Yang XK & Lin M (2021) Higher-level phylogeny of longhorn beetles (Coleoptera: Chrysomeloidea) inferred from mitochondrial genomes. *Systematic Entomology* 46(1): 56–70. <https://doi.org/10.1111/syen.12447>
- Nei M (1978). The theory of genetic distance and evolution of human races: The Japan society of human genetics award lecture. *Japanese Journal of Human Genetics* 23: 341–369
- Ovenden JR, Dudgeon C, Feutry P, Feldheim K & Maes GE (2018) Genetics and genomics for fundamental and applied research on elasmobranchs. *Shark research: Emerging technologies and applications for the Field and laboratory* 1: 235–254
- Patel RK, Jain M, Toolkit NGSQC (2012) A toolkit for quality control of next generation sequencing data. *PLoS One* 7: e30619. <https://doi.org/10.1371/journal.pone.0030619>
- Pogoda CS, Keepers KG, Lendemer JC, Kane NC, Tripp EA (2018) Reductions in complexity of mitochondrial genomes in lichen-forming fungi shed light on genome architecture of obligate symbioses. *Molecular ecology* 27(5): 1155–1169. <https://doi.org/10.1111/mec.14519>
- Ranwez V, Douzery EJP, Cambon C, Chantret N, Delsuc F (2018) MACSE v2: Toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Molecular biology and evolution* 35: 2582–2584. <https://doi.org/10.1093/molbev/msy159>
- Rubinoff D, & Holland, B. S (2005) Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic biology* 54(6): 952–961.
- Shi F, Ge S, Hou Z, Xu Y, Tao J, Wu H, & Zong S (2022) Species-specific primers for rapid detection of *Monochamus saltuarius*, an effective vector of *Bursaphelenchus xylophilus* in China. *Journal of Applied Entomology* 146(5): 636–647. <https://doi.org/10.1111/jen.12972>
- Shi F, Yu T, Xu Y, Zhang S, Niu Y, Ge S & Zong S (2023) Comparative mitochondrial genomic analysis provides new insights into the evolution of the subfamily Lamiinae (Coleoptera: Cerambycidae). *International Journal of Biological Macromolecules* 225: 634–647. <https://doi.org/10.1016/j.ijbiomac.2022.11.125>
- Silva-Pinheiro P, & Minczuk M (2022) The potential of mitochondrial genome engineering. *Nature Reviews Genetics* 23(4): 199–214. <https://doi.org/10.1038/s41576-021-00432-x>
- Song SN, Tang P, Wei SJ, Chen XX (2016) Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans. *Scientific Reports* 6(1): 20972. <https://doi.org/10.1038/srep20972>
- Sun CH, Liu HY, & Lu CH (2020) Five new mitogenomes of *Phylloscopus* (Passeriformes, Phylloscopidae): Sequence, structure, and phylogenetic analyses. *International journal of biological macromolecules* 146: 638–647. <https://doi.org/10.1016/j.ijbiomac.2019.12.253>
- Sun F (2019) *Fauna of Cerambycidae (Insecta: Coleoptera) from Dayaoshan, Guangxi, China*. Phd Thesis, University of Guangxi Normal, Guiyang, China.
- Tavakilian GL, Chevillotte H (2019) Titan: base de données internationales sur les Cerambycidae ou Longicornes. Internet resource (<http://titan.gbif.fr/index.html>), accessed January, 25, 2019.
- Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution* 38(7): 3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Wei L, He J, Jia X, Qi Q, Liang Z, Zheng H, Ping Y, Liu S, Sun J (2014) Analysis of codon usage bias of mitochondrial genome in *Bombyx mori* and its relation to evolution. *BMC evolutionary biology* 14: 1–12. <https://doi.org/10.1186/s12862-014-0262-4>
- Wei S (2009) *Characterization and Evolution of Hymenopteran Mitochondrial Genomes and Their Phylogenetic Utility*. Phd thesis. Zhejiang University, Zhejiang, China.
- Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, & Scheffler K (2015) RELAX: detecting relaxed selection in a phylogenetic

- ic framework. *Molecular biology and evolution* 32(3): 820–832. <https://doi.org/10.1093/molbev/msu400>
- Wolstenholme DR (1992) Animal mitochondrial DNA: structure and evolution, *International Review of Cytology* 141: 173–216. [https://doi.org/10.1016/S0074-7696\(08\)62066-5](https://doi.org/10.1016/S0074-7696(08)62066-5)
- Wu Y, Liu X, Zhang Y, Fang H, Lu J, Wang J (2023) Characterization of four mitochondrial genomes of Crambidae (Lepidoptera, Pyraloidea) and phylogenetic implications. *Archives of Insect Biochemistry and Physiology* 112(1): e21914. <https://doi.org/10.1002/arch.21914>
- Wu P, Xiao W, Luo Y, Xiong Z, Chen X, He J, Sha A, Gui M, Li Q (2023) Comprehensive analysis of codon bias in 13 *Ganoderma* mitochondrial genomes. *Frontiers in Microbiology* 14: 1170790. <https://doi.org/10.3389/fmicb.2023.1170790>
- Xiang CY, Gao F, Jakovčić I, Lei HP, Hu Y, Zhang H, Zou H, Wang G & Zhang, D (2023) Using PhyloSuite for molecular phylogeny and tree-based analyses. *iMeta* 2(1): e87. <https://doi.org/10.1002/imt2.87>
- Xia XH (2013) DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Molecular biology and evolution* 30: 1720–1728. <https://doi.org/10.1093/molbev/mst064>
- Xu H, Wu Y, Wang Y, Liu Z (2020) Comparative analysis of five mitochondrial genomes of Osmylinae (Neuroptera: Osmyliidae) and their phylogenetic implications. *International Journal of Biological Macromolecules* 164: 447–455. <https://doi.org/10.1016/j.ijbiomac.2020.07.150>
- Yan L, Xu W, Zhang D, Li J (2021) Comparative analysis of the mitochondrial genomes of flesh flies and their evolutionary implication. *International Journal of Biological Macromolecules* 174: 385–391. <https://doi.org/10.1016/j.ijbiomac.2021.01.188>
- Yang M, Li J, Su S, Zhang H, Wang Z, Ding W, & Li L (2021) The mitochondrial genomes of Tortricidae: nucleotide composition, gene variation and phylogenetic performance. *BMC genomics* 22(1): 1–12. <https://doi.org/10.1186/s12864-021-08041-y>
- Yamamoto Y, Ishikawa Y, & Uehara K (2022) Characteristics of Trees Infested by the Invasive Primary Wood-Borer *Aromia bungii* (Coleoptera: Cerambycidae). *Insects* 13(1): 54. <https://doi.org/10.3390/insects13010054>
- Yang M, Gong L, Sui J, & Li X (2019) The complete mitochondrial genome of *Calypptogena marissinica* (Heterodonta: Veneroidea: Vesicomidae): Insight into the deep-sea adaptive evolution of vesicomids. *PLoS one* 14(9): e0217952. <https://doi.org/10.1371/journal.pone.0217952>
- Yang T, Wang Y, Liao W, Zhang S, Wang S, Xu N, Xie W, Luo C, Wang Y, Wang Z, Zhang Y (2021) Down-regulation of EPB41L4A-AS1 mediated the brain aging and neurodegenerative diseases via damaging synthesis of NAD⁺ and ATP. *Cell & bioscience* 11: 1–14. <https://doi.org/10.1186/s13578-021-00705-2>
- Yévenes M, Núñez-Acuña G, Gallardo-Escárate C, & Gajardo G (2022) Adaptive mitochondrial genome functioning in ecologically different farm-impacted natural seedbeds of the endemic blue mussel *Mytilus chilensis*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 42: 100955. <https://doi.org/10.1016/j.cbd.2021.100955>
- Zhang KJ, Zhu WC, Rong X, Zhang YK, Ding XL, Liu J, Chen DS, Du Y, Hong XY (2013) The complete mitochondrial genomes of two rice planthoppers, *Nilaparvata lugens* and *Laodelphax striatellus*: conserved genome rearrangement in Delphacidae and discovery of new characteristics of *atp8* and tRNA genes. *BMC genomics* 14: 1–12. <https://doi.org/10.1186/1471-2164-14-417>
- Zhang SQ, Che LH, Li Y, Liang D, Pang H, Ślipiński A, Zhang P (2018) Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nature communications* 9(1): 205. <https://doi.org/10.1038/s41467-017-02644-4>
- Zheng B (2022) Comparative mitogenomics and phylogenetics of Ichneumonidae Hymenoptera: Aculeata. Phd thesis. Zhejiang University, Zhejiang, China.

Supplementary Material 1

Tables S1–S4

Authors: Niu YM, Shi FM, Li XY, Zhang SN, Xu YB, Tao J, Li M, Zhao YX and Zong SX (2024)

Data type: .xlsx

Explanation notes: **Table S1.** Samples and gene information of mitochondrial genomes used in the study. — **Table S2.** Best partitioning schemes and models based on different datasets calculated by ModelFinder and PartitionFinder for maximum likelihood and Bayesian inference analysis. — **Table S3.** Base composition and strand bias across mitochondrial genomes used in the study. — **Table S4.** Evaluation of codon bias across Cerambycidae mitogenomes. ENC, effective number of codons..

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/asp.82.e114299.suppl1>

Supplementary Material 2

Tables S5

Authors: Niu YM, Shi FM, Li XY, Zhang SN, Xu YB, Tao J, Li M, Zhao YX and Zong SX (2024)

Data type: .xlsx

Explanation notes: Sequence characteristics of the mitochondrial genomes of 86 longhorn beetles.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/asp.82.e114299.suppl2>

Supplementary Material 3

Tables S6

Authors: Niu YM, Shi FM, Li XY, Zhang SN, Xu YB, Tao J, Li M, Zhao YX and Zong SX (2024)

Data type: .xlsx

Explanation notes: The codon usage patterns of PCGs in the mitochondrial genomes of 86 longhorn beetles.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/asp.82.e114299.suppl3>

Supplementary Material 4

Figures S1, S2

Authors: Niu YM, Shi FM, Li XY, Zhang SN, Xu YB, Tao J, Li M, Zhao YX and Zong SX (2024)

Data type: .zip

Explanation notes: **Figure S1.** Nucleotide substitution saturation plots of mitochondrial genomes for different datasets. — **Figure S2.** Phylogenetic tree of Cerambycidae inferred from the AA, PCG123, PCG12RNA, PCG123R dataset using PhyloBayes (BI, A, B, C, E) and PCG12RNA, PCG123R dataset using IQ-TREE (ML, D, F), Branches are labeled with Bayesian posterior probabilities (PPs) more than 0.75 and parsimony bootstrap values (BPs) higher than 50%.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/asp.82.e114299.suppl4>