

## Research Article

# The Complete Mitochondrial Genome of *Ctenoptilum vasava* (Lepidoptera: Hesperioidea: Pyrginae) and Its Phylogenetic Implication

Jiasheng Hao,<sup>1,2</sup> Qianqian Sun,<sup>1</sup> Huabin Zhao,<sup>3</sup> Xiaoyan Sun,<sup>2</sup> Yonghua Gai,<sup>2</sup> and Qun Yang<sup>2</sup>

<sup>1</sup> College of Life Sciences, Anhui Normal University, Wuhu 241000, China

<sup>2</sup> SKLPS, LPS, Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing 210008, China

<sup>3</sup> Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

Correspondence should be addressed to Jiasheng Hao, jshaonigpas@sina.com and Qun Yang, qunyang@nigpas.ac.cn

Received 6 November 2011; Accepted 19 December 2011

Academic Editor: Elena Pasyukova

Copyright © 2012 Jiasheng Hao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We here report the first complete mitochondrial (mt) genome of a skipper, *Ctenoptilum vasava* Moore, 1865 (Lepidoptera: Hesperioidea: Pyrginae). The mt genome of the skipper is a circular molecule of 15,468 bp, containing 2 ribosomal RNA genes, 24 putative transfer RNA (tRNA), genes including an extra copy of trnS (AGN) and a tRNA-like insertion trnL (UUR), 13 protein-coding genes and an AT-rich region. All protein-coding genes (PCGs) are initiated by ATN codons and terminated by the typical stop codon TAA or TAG, except for COII which ends with a single T. The intergenic spacer sequence between trnS (AGN) and ND1 genes also contains the ATACTAA motif. The AT-rich region of 429 bp is comprised of nonrepetitive sequences, including the motif ATAGA followed by an 19 bp poly-T stretch, a microsatellite-like (AT)<sub>3</sub> (TA)<sub>9</sub> element next to the ATTTA motif, an 11 bp poly-A adjacent to tRNAs. Phylogenetic analyses (ML and BI methods) showed that Papilionoidea is not a natural group, and Hesperioidea is placed within the Papilionoidea as a sister to ((Pieridae + Lycaenidae) + Nymphalidae) while Papilionoidea is paraphyletic to Hesperioidea. This result is remarkably different from the traditional view where Papilionoidea and Hesperioidea are considered as two distinct superfamilies.

## 1. Introduction

The taxonomic status and the phylogenetic position of skippers (Hesperioidea) within Lepidoptera remain a controversial issue [1–3]. Due to the distinct differences between the skippers and the typical butterflies/moths in terms of morphological and behavioral characteristics, such as the short stout bodies, hooked antennae, and rapid skipping flight, the skippers were previously proposed to represent a separate group that is distinct from butterflies/moths in lepidopterans. More specifically, the skippers are assigned to the family Hesperioidea in a monotypic superfamily Hesperioidea, a sister lineage to the typical rhopaloceran butterflies, which mostly belong to superfamily Papilionoidea (true butterflies) [2, 4, 5]. In addition, the three superfamilies Hesperioidea, Papilionoidea, and Hedyloidea share numerous morphological characteristics, particularly in their egg,

larval, and pupal stages, and thus were considered to be a large natural group [2].

The Lepidoptera is one of the largest groups of insects, accounting for more than 160,000 species. Despite of the huge taxonomic diversity, the current information on the lepidopteran mt genomes is very limited. Only 40 lepidopteran mt genomes were sequenced, including 10 butterfly species such as *Coreana raphaelis* [6], *Artogeia melete* [7], *Parnassius bremeri* [8], *Acraea issoria* [9], *Pieris rapae* [10], *Eumenis autonoe* [11], *Calinaga davidis* [12], and nearly thirty moth species such as *Bombyx mori*, *Bombyx mandarina* [13], *Adoxophyes honmai* [14], *Antheraea pernyi* [15], *Ochrogaster lunifer* [16], *Manduca sexta* [17], *Phthonandria atrilineata* [18], *Eriogyna pyretorum* [19], *Antheraea yamamai* [20], and *Caligula boisduvalii* [21]. The sampling is restricted to only six superfamilies (Papilionoidea, Totricioidea, Bombycoidea, Geometroidea, Noctuoidea, and Pyraloidea) to date,

TABLE 1: Primer sequences for the long PCR amplification used in this study.

Primers	Upper primer sequence (5'-3')	Lower primer sequence (5'-3')
COI-COIII	GGAAATTGACTTGTGCCT	TTGTATGTTTACCTTGGA
COIII-ND4	AAAGGATTACGATGAGGT	GGTCTTGTTATTTGGTGGGA
ND4-Cytb	CGTCTATGTAAACGCTCA	ATAAGGGTTTTCTACTGGT
Cytb-12S	TTTTACATCAAACAGGA	ACTAGGATTAGATACCC
12S-COI	GAAACACTTTCAGTACCT	CTAAACCAATTCAACATCC

a complete mt genome sequence of a skipper from the family Hesperidae is lacking, despite of a huge diversity of the skippers (>3500 species). The lack of mt genome data from skippers dampens phylogenetic and population genetic studies in skippers and the related species.

The Tawny Angle, *Ctenoptilum vasava*, is a typical skipper commonly found in southern East Asia, such as China, India, Burma, Thailand, and Vietnam. In this study, we sequenced the complete mitochondrial genome of the skipper, representing the first mt genome sequence from the family Hesperidae (superfamily Hesperioidea). We next compared this sequence with other lepidopteran mt genomes sequences and examined the phylogenetic relationships within lepidopterans and reevaluated the phylogenetic position of skippers. We show that the skipper shares the general organization and structure of the mt genome with other species from the order Lepidoptera. By examining currently available mt genomes in lepidopterans, we find the Hesperioidea is placed within the Papilionoidea, which may be a paraphyletic group.

## 2. Materials and Methods

**2.1. Sample Collection.** Adult individuals of *Ctenoptilum vasava* (Lepidoptera: Hesperidae: Pyrginae: *Ctenoptilum*) were captured from National Natural Conservation Areas of Jiu Lianshan Mountain, Jiangxi Province, China, in August, 2009. After a brief examination for species identification, the fresh tissues were preserved in 100% ethanol immediately for DNA fixation and stored at  $-20^{\circ}\text{C}$  until further use for genomic DNA isolation.

**2.2. DNA Extraction and PCR Amplification.** The whole genomic DNA of *C. vasava* was isolated from the thoracic muscle of an adult individual using the proteinase K-SiO<sub>2</sub> method as described by Hao et al. [32]. Partial sequences of COI, COIII, Cytb, ND4, 16S rRNA, and 12S rRNA genes were amplified using insect universal primers [33]. Polymerase chain reactions (PCRs) were performed under the following condition: 5 minutes of initial denaturation at  $95^{\circ}\text{C}$ , 35 cycles of denaturation at  $95^{\circ}\text{C}$  for 50 seconds, annealing at  $45\text{--}55^{\circ}\text{C}$  (depending on primer pairs) for 50 seconds, extension at  $72^{\circ}\text{C}$  for 1 minute, and a final extension at  $72^{\circ}\text{C}$  for 10 minutes. Based on the sequences from the newly acquired gene fragments, the long PCR primers were designed (Table 1) according to the conserved regions by the program Primer premier 5.0 [34], and the entire mt genome of the skipper was in turn amplified in five

long fragments (12S-COI, COI-COIII, COIII-ND4, ND4-Cytb, Cytb-12S) by using Takara LA Taq<sup>TM</sup> (Takara). The long PCR condition is as follow: an initial denaturation at  $95^{\circ}\text{C}$  for 5 min, 15 cycles of denaturation at  $95^{\circ}\text{C}$  for 50 seconds, annealing at  $50\text{--}55^{\circ}\text{C}$  (depending on primer pairs) for 50 seconds, extension at  $68^{\circ}\text{C}$  for 150 seconds, additional 15 cycles of denaturation at  $95^{\circ}\text{C}$  for 50 seconds, annealing at  $50\text{--}55^{\circ}\text{C}$  for 50 seconds, extension at  $68^{\circ}\text{C}$  for 150 seconds, and a final extension at  $68^{\circ}\text{C}$  for 10 minutes. PCR products were examined by electrophoresis on a 1% agarose gel and purified using a DNA gel extraction kit (Takara). All PCR fragments were directly sequenced in both strands after purification with QIA quick PCR Purification Kit (Qiagen). Long PCR fragments were sequenced using the primer walking strategy (walking primer information will be provided upon request).

**2.3. Sequence Analysis.** The raw sequence files were proof-read and assembled in BioEdit version 7.0 [35]. The concatenated amino acid sequences of the 13 PCGs were obtained and analysed by the ClustalX [36] and the MEGA 3.0 [37] softwares. The structures of the 23 tRNAs and a tRNA-like gene of the skipper were identified by the software tRNAscan-SE version 1.21 [38]. The putative tRNAs, which were not found by tRNAscan-SE, were identified by sequence comparisons between the skipper and other lepidopteran tRNAs. Nucleotide composition was calculated using MEGA 3.0 software [37], and the tandem repeats in the AT-rich region were predicted by the Tandem Repeats Finder available online (<http://tandem.bu.edu/trf/trf.html>) [39]. The mt genome sequence has been submitted to GenBank database under the accession number JF713818.

**2.4. Phylogenetic Analysis.** The multiple alignment of the 13 PCG concatenated nucleotide sequences of the 32 lepidopteran mt genome sequences (one is from this study, and 31 were extracted from GenBank, see Table 2) was conducted using Clustal X1.8 [36] and was checked by eye. The phylogenetic trees were reconstructed with the maximum likelihood (ML) and bayesian inference (BI) methods using a hymenopteran species *Apis cerana* (GenBank accession number NC.014295) as the outgroup. In both phylogenetic analyses, the third position of all the codons was excluded. The ML analyses were conducted in PAUP (version 4.0b8) [40] with searching method of TBR branch swapping (10 random addition sequences), the general time reversible model with invariant sites and among-site variation (GTR+I+ $\Gamma$ ) was selected as the best fit model using

TABLE 2: Lepidopteran mt genome sequences used in this study.

Species	Superfamily and family		GenBank acc. no.	References
<i>Hyphantria cunea</i>	Noctuoidea	Arctiidae	NC.014058	[22]
<i>Lymantria dispar</i>		Lymantriidae	NC.012893	[23]
<i>Helicoverpa armigera</i>		Noctuidae	NC.014668	[24]
<i>Ochrogaster lunifer</i>		Notodontidae	NC.011128	[16]
<i>Phalera flavescens</i>		Notodontidae	JF440342	(Sun et al., unpublished)
<i>Phthonandria atrilineata</i>	Geometroidea	Geometridae	NC.010522	[18]
<i>Spilonota lechriaspis</i>	Tortricoidea	Tortricidae	NC.014294	[25]
<i>Adoxophyes honmai</i>		Tortricidae	NC.008141	[14]
<i>Grapholita molesta</i>		Tortricidae	NC.014806	[26]
<i>Antheraea pernyi</i>	Bombycoidea	Saturniidae	NC.004622	[15]
<i>Antheraea yamamai</i>		Saturniidae	NC.012739	[20]
<i>Eriogyna pyretorum</i>		Saturniidae	NC.012727	[19]
<i>Caligula boisduvalii</i>		Saturniidae	NC.010613	[21]
<i>Chinese Bombyx mandarina</i>		Bombycidae	AY301620	[27]
<i>Bombyx mandarina</i>		Bombycidae	NC.003395	[13]
<i>Bombyx mori</i>		Bombycidae	NC.002355	[13]
<i>Manduca sexta</i>		Sphingidae	NC.010266	[17]
<i>Diatraea saccharalis</i>	Pyraloidea	Crambidae	NC.013274	[28]
<i>Ostrinia furnacalis</i>		Crambidae	NC.003368	[29]
<i>Ostrinia nubilalis</i>		Crambidae	NC.003367	[29]
<i>Eumemis autonoe</i>	Papilionoidea	Nymphalidae	NC.014587	[11]
<i>Acraea issoria</i>		Nymphalidae	GQ376195	[9]
<i>Sasakia charonda kuriyamaensis</i>		Nymphalidae	NC.014223	(Hakozaki et al., unpublished)
<i>Argyreus hyperbius</i>		Nymphalidae	JF439070	[30]
<i>Calinaga davidis</i>		Nymphalidae	HQ658143	[12]
<i>Papilio maraho</i>		Papilionidae	NC.014055	[31]
<i>Parnassius bremeri</i>		Papilionidae	NC.014053	[8]
<i>Teinopalpus aureus</i>		Papilionidae	NC.014398	(Qin et al., unpublished)
<i>Troides aeacus</i>		Papilionidae	EU625344	(Jiang et al., unpublished)
<i>Coreana raphaelis</i>		Lycaenidae	NC.007976	[6]
<i>Pieris raphaelis</i>		Pieridae	HM156697	[10]
<i>Ctenoptilum vasava</i>	Hesperoidea	Hesperiidae	JF713818	This study

Modeltest (version 3.06) [41] under the AIC criteria, and the bootstrap values of the ML tree were evaluated via the bootstrap test with 1000 iterations. The Bayesian analysis was performed using MrBayes 3.1.2 [42] with the partitioned strategy, and the best fit substitution model was selected as in the ML analysis. MrBayes 3.1.1 simultaneously initiates 2 Markov Chain Monte Carlo (MCMC) that runs to provide additional confirmation for the convergence of posterior probability distribution. Four simultaneous chains were run, 3 hot and 1 cold. Analyses were run for one million generations until the average standard deviation of split frequencies to be less than 0.01, which means convergence was reached. Chains were sampled every 1,000 generations. Starting trees were random.

### 3. Results and Discussion

**3.1. Genome Organization, Gene Arrangement, and Base Composition.** The organization of the skipper mt genome was

shown in Figure 1. The complete mt genome is 15,468 bp in length, containing 13 protein-coding genes (ND1-6, ND4L, COI-III, Cytb, ATP6, ATP8), 2 ribosomal RNAs (12S and 16S), 24 putative tRNAs, and an AT-rich region, same characteristics as with those of other butterfly species available (Tables 3 and 4). The size of so far sequenced mt genomes in lepidopteran insects ranges from 15,140 bp in *Artogeia melete* [7] to 15,928 bp in *Bombyx mandarina* [13]; our newly sequenced mt genome of the skipper is within the above size range. Mitochondrial genes of the skipper are arranged in the same order and orientation as those of other lepidopterans, except for the presence of an extra copy of trnS (AGN) and a tRNA-like insertion trnL(UUR). Similar to many insect mt genomes, the heavy strand (H-strand) encodes more genes (9 PCGs and 14 tRNAs), whereas the light strand (L-strand) encodes less (4 PCGs, 8 tRNAs, and 2 rRNA genes).

The composition of A, T, G, C nucleotides in the skipper mt genome L-strand indicating a remarkably high AT

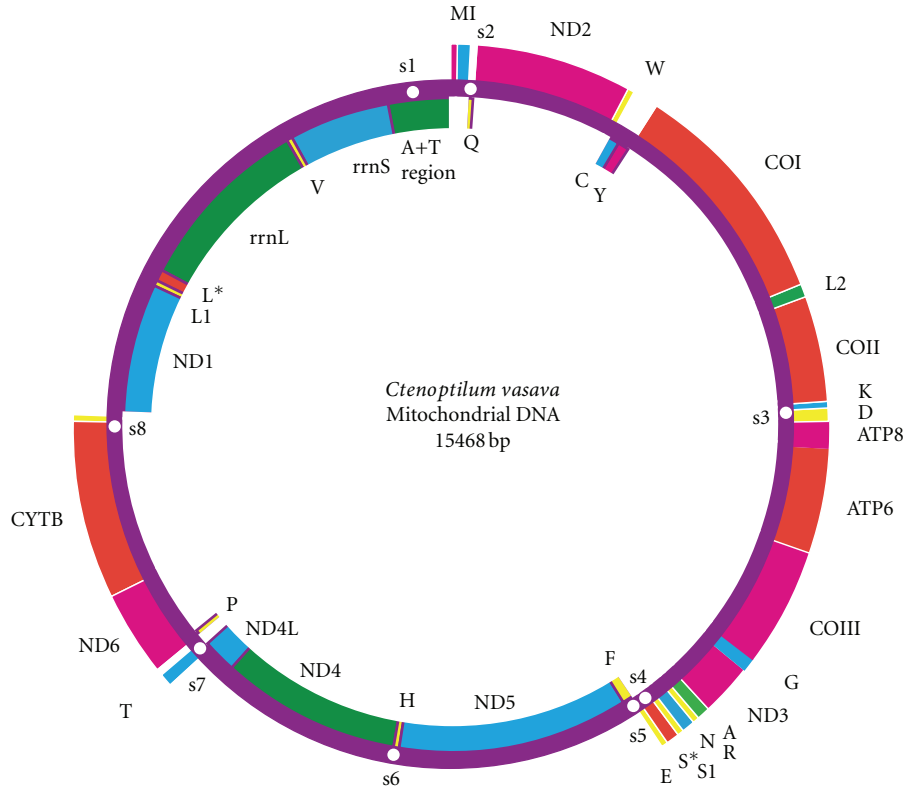


FIGURE 1: Map of the circular mitochondrial genome of *Ctenoptilum vasava*. Genes encoded in the H-strand (clockwise orientation) are colored in red or pink. Genes encoded in the L-strand (anticlockwise orientation) are colored in blue or green. The abbreviations for the genes are as follows: *COI-III* stands for cytochrome oxidase subunits, *Cytb* for cytochrome b, and *ND1-6* for NADH dehydrogenase components. *tRNAs* are denoted as one-letter symbol according to the IUPAC-IUB single-letter amino acid codes.

content (80.55%) and a low GC content (19.45%). This result is consistent with the AT bias generally observed in other lepidopteran mt genomes, given that the composition of AT is ranging from 77.9% in *Ochrogaster lunifer* [16] to 82.7% in *Coreana raphaelis* [6]. The AT and GC skew values in H-strand of the skipper mt genome are  $-0.0293$  and  $-0.216$ , respectively, indicating that T and C are favored over A and G, respectively.

**3.2. Protein-Coding Genes (PCG).** The 13 PCGs of the skipper mt genome include 7 NADH dehydrogenase subunits, 3 cytochrome c oxidase subunits, 2 ATPase subunits, and one cytochrome b gene. The 13 PCGs are 11,172 bp in length which accounts for 72.23% of the whole mitochondrial genome, encoding 3,724 amino acid residues. All the protein-coding genes except for the *COI* start with a canonical start codon ATN. More specifically, 7 PCGs (*COII*, *ATP6*, *COIII*, *ND4*, *ND4L*, *Cytb*, and *ND1*) start with ATG, 2PCGs (*ND2*, *ND3*) with ATT, additional 3 PCGs (*ATP8*, *ND5*, *ND6*) with ATA. For the stop codon, 3 PCGs (the *ND2*, *ND3*, and *ND5*) terminate with ATG, 2 PCGs (*COI* and *COII*) with a single T, the remaining 8 PCGs with the typical stop codon TAA.

The start codons for *COI* gene of the lepidopteran insects are not uniform. All lepidopteran species examined to date

use Arginine (R), which correspond to the codon CGA, as the initial amino acid for *COI* [43]. In this study, the CGA is not conserved across all lepidopteran mt genomes, but it has been suggested to serve as the *COI* start codon [8]. However, in other insect groups, some other canonical codons, such as the TTA [44], TCG [45], TTG [46], ACG [47], were reported as the *COI* start codons. Additionally, some tetranucleotides, such as the ATAA, ATCA, and ATTA [48], as well as the hexanucleotides, such as the ATTTAA [49–51], TATCTA [52], TTTTATG [13], and TATTAG [29], were also proposed as the start codons for this gene.

The 13 protein-coding genes all possess complete stop codons except for the *COI* and *COII* in this study. Most stop codons for *COII* are single T in most sequenced lepidopteran mt genomes. The single T stop codon is usually located nearby the *trnK*, the structure of which is recognized by endonucleases processing the polycistronic pre-mRNA transcription, and produces functional stop codons by polyadenylation from its contiguous protein-coding genes [53–55]. Accordingly, partial stop codons observed in most lepidopteran species would be one strategy for the selection of stop codon; in other words, it would minimize the intergenic spacers and gene overlaps from an evolutionary economic perspective.

TABLE 3: Comparison of the mt genome characteristics of all butterfly species available. All genomes share the features of 13 protein-coding genes (ND1-6, ND4L, COI-III, Cytb, ATP6, ATP8), 2 ribosomal RNAs (lsu rRNA (12S) and lsu rRNA (16S)), and an AT-rich region.

Taxon	Whole genomic DNA				lrRNA		srRNA		A+T-rich region		GenBank
	Size (bp)	A+T %	No. codons <sup>a</sup>	PCG <sup>b</sup>	A+T %	Size (bp)	A+T %	Size (bp)	A+T %	Size (bp)	A+T %
<i>Adoxophyes honmai</i>	15,680	80.4	3,748	78.5	1,387	83.6	779	85.4	489	94.3	DQ073916
<i>Antheraea pernyi</i>	15,575	80.2	3,732	78.5	1,369	83.9	775	84.1	552	90.4	AY242996
<i>Antheraea yamamai</i>	15,338	80.2	3,714	79.3	1,380	83.5	776	85.9	334	90.4	EU726630
<i>Caligula boisduvalii</i>	15,360	80.6	3,734	79.1	1,391	84.8	774	84.1	330	91.5	EF622227
<i>Eriogyna pyretorum</i>	15,327	80.8	3,711	79.3	1,338	84.6	762	84.5	374	91.7	FJ685653
<i>Bombyx mori</i>	15,656	81.4	3,720	79.5	1,378	84.4	783	85.6	494	95.5	AB070264
<i>Manduca sexta</i>	15,516	81.8	3,718	80.2	1,391	84.7	777	86.8	324	95.1	EU286785
<i>Phthonandria atrilineata</i>	15,499	81.1	3,724	79.0	1,400	85.1	803	87.5	457	98.5	NC010522
<i>Parnassius bremeri</i>	15,389	81.3	3,734	80.2	1,344	83.8	773	85.1	504	93.6	FJ871125
<i>Artogeia melete</i>	15,140	79.8	3,715	78.4	1,319	83.4	777	86.9	351	88.0	NC010568
<i>Coreana raphaelis</i>	15,314	82.7	3,708	81.5	1,330	85.3	777	85.8	375	94.1	DQ102703
<i>Eumenis autonoe</i>	15,489	79.1	3,728	76.8	1,335	83.7	775	85.3	678	94.5	GQ868707
<i>Apatura ilia</i>	15,242	80.5	3,711	78.9	1,333	86.0	776	84.9	403	92.5	JF437925
<i>Calinaga davidis</i>	15,267	80.4	3,737	78.8	1,337	83.8	773	85.9	403	92.5	HQ658143
<i>Argyreus hyperbius</i>	15,156	80.8	4,164	79.7	1,330	84.5	778	85.2	349	95.4	JF439070
<i>Acraea issoria</i>	15,245	79.7	3,745	78.0	1,331	83.8	788	83.8	430	96.0	GQ376195
<i>Ochrogaster lunifer</i>	15,593	77.9	3,746	75.7	1,351	81.4	806	84.3	319	92.7	AM946601
<i>Lymantria dispar</i>	15,569	79.9	3,742	77.8	1,351	84.2	799	85.2	435	96.1	FJ617240
<b><i>Ctenoptilum vasava</i></b>	<b>15,468</b>	<b>80.5</b>	<b>3,698</b>	<b>78.9</b>	<b>1,343</b>	<b>84.1</b>	<b>774</b>	<b>86.4</b>	<b>429</b>	<b>88.1</b>	<b>JF713818</b>

The amino acids composition and the codon usage of the skipper mt genome are shown in the Tables 5 to 6. The results showed that the codon usage of all the genes has a strong bias, and the RSCU (relative synonymous codon usage) of NNU and NNA codons are greater than 1, indicating that the third positions of the UA have a high frequency of codon usage. The codon usage bias in protein-coding genes and the third position of AT bias (92.1%) are positively correlated. In addition, our analysis also showed that UUU (Phe), UUA (Leu), AUU (Ile), AUA (Met), and AAU (Asn) are the most frequently used codons (45.45%). Together, all observations suggest the strong AT bias of the protein-coding genes in the skipper mt genome.

**3.3. Ribosomal and Transfer RNA Genes.** The srRNA (12S rRNA) and lrRNA (16S rRNA) genes in the skipper mt genome are 774 bp and 1343 bp in size, respectively. Similar to other lepidopteran rRNAs [8, 22], the skipper 12S rRNA is located between trnL (CUN) and trnV, whereas its 16S rRNA resides between trnV and the A+T-rich region. The A+T contents of the two rRNA genes are 86.43% and 84.14%, respectively, which is consistent with those observed in other lepidopterans as well.

The skipper mt genome contains 23 transfer RNAs and a tRNA-like gene (trnL (UUR)), and these genes are interspersed throughout the whole genome and ranged in

size from 61 to 79 bp. Twenty-three of them are shown to be folded into the expected clover-leaf secondary structure; however, the trnS (AGN) lacks the dihydrouridine (DHU) loop (Figure 2) as shown in the other insect mt genomes examined to date. Therefore, the incomplete trnS should be considered as one of common features in most, if not all, insect mt genomes. A total of 32 unmatched base pairs were detected in the skipper tRNAs, and 18 of them are GU pairs, which form a weak bond in the tRNAs, the remaining 14 are atypical pairs, including 11 UU pairs, two AC pairs, and one AA pairs (Figure 2).

The extra trnS was detected between trnS (AGN) and trnE in the skipper mt genome (Figure 2), and the tRNA insertions were also observed in other lepidopterans, such as *C. raphaelis* and *A. issoria*. The high similarity between the two tandemly repeated copies of the trnS in primary and secondary structures may suggest a recent duplication event [6, 9]. Interestingly, a 79 bp insertion of a tRNA-like gene was detected between the trnL (CUN) and lrRNA genes, and this observation was not shown in any lepidopteran mt genomes to date, despite that tRNA translocation is a frequent event in the evolution of lepidopteran mt genomes [56].

**3.4. Intergenic Spacers and Overlapping Sequences.** The skipper mt genome contains a total of 607 bp intergenic spacers, which distribute in 11 regions, ranging from 2 to 61 bp in

TABLE 4: Summary of the skipper mt genome. IGNC stands for intergenic nucleotides, and the positive number indicates interval nucleotides (base pairs) between genes, while the negative number indicates the overlapped nucleotides (base pairs) between genes.

Gene	Direction	Nucleotide no.	Size	IGNc	Start codon	Stop codon
tRNA <sup>Met</sup>	F	1–68	68	0		
tRNA <sup>Ile</sup>	F	69–133	65	0		
tRNA <sup>Gln</sup>	R	136–204	69	2		
ND2	F	266–1270	1005	61	ATT	TAG
tRNA <sup>Trp</sup>	F	1269–1335	67	–2		
tRNA <sup>Cys</sup>	R	1331–1395	65	–5		
tRNA <sup>Tyr</sup>	R	1407–1471	65	11		
COI	F	1469–2990	1522	–3	ATT	T-tRNA
tRNA <sup>Leu</sup> (UUR)	F	2991–3057	67	0		
COII	F	3058–3731	674	0	ATG	T-tRNA
tRNA <sup>Lys</sup>	F	3732–3801	70	0		
tRNA <sup>Asp</sup>	F	3816–3884	69	14		
ATP8	F	3885–4052	168	0	ATA	TAA
ATP6	F	4046–4723	678	–7	ATG	TAA
COIII	F	4723–5505	783	–1	ATG	TAA
tRNA <sup>Gly</sup>	F	5508–5574	67	2		
ND3	F	5575–5928	354	0	ATT	TAG
tRNA <sup>Ala</sup>	F	5927–5990	64	–2		
tRNA <sup>Arg</sup>	F	5996–6061	66	5		
tRNA <sup>Asn</sup>	F	6061–6125	65	–1		
tRNA <sup>Ser</sup> (AGN)	F	6133–6194	62	7		
tRNA <sup>Ser</sup> (AGN)	F	6208–6267	60	13		
tRNA <sup>Glu</sup>	F	6279–6344	66	11		
tRNA <sup>Phe</sup>	R	6343–6409	67	–2		
ND5	R	6409–8145	1737	–1	ATA	TAA
tRNA <sup>His</sup>	R	8161–8228	68	15		
ND4	R	8228–9568	1341	–1	ATG	TAA
ND4L	R	9574–9855	282	5	ATG	TAA
tRNA <sup>Thr</sup>	F	9869–9933	65	13		
tRNA <sup>Pro</sup>	R	9934–9998	65	0		
ND6	F	10001–10531	531	2	ATA	TAA
Cytb	F	10531–11679	1149	–1	ATG	TAA
tRNA <sup>Ser</sup>	F	11679–11745	67	–1		
ND1	R	11763–12704	942	17	ATG	TAA
tRNA <sup>Leu</sup> (CUN)	R	12706–12773	68	1		
tRNA <sup>Leu</sup> (UUR)	F	12774–12852	79	0		
16S	R	12853–14195	1343	0		
tRNA <sup>Val</sup>	R	14196–14265	70	0		
12S	R	14266–15039	774	0		
A+T-rich region		15040–15468	429	0		

length. Most of the spacers are shorter than 10 bp, only 7 spacers are longer (Table 2). The longest intergenic spacer (61 bp) is located between the trnQ and ND2 genes, with a high AT content (98.4%). This spacer is usually considered as a constitutive synapomorphic feature of lepidopteran mt genomes because of the absence in nonlepidopteran insects [17]. The second longest one (ATACTAAAAATATATTA) is inserted between the trnS2 and ND1 and shared the

ATACTAA motif observed in most insects, including lepidopterans [17], and possibly fundamental to site recognition by the transcription termination peptide (mtTERM protein) [57]. The third longest spacer (CAATTTCTTTT) is inserted between the trnC and trnY, which is the same as *O. lunifer*. The fourth longest spacer (AAATTATTAATTT) is located between trnK and trnD, which is also found in other lepidopterans. The fifth longest one (TTTTCTTTTCTTT) is

TABLE 5: Amino acid composition of the mt genome of the skipper. Start codons and stop codons were excluded in total codon counts.

	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Total
ND2	2.10	1.20	0.30	1.50	17.72	3.30	0.60	13.81	5.11	11.11	8.41	8.41	1.50	1.20	0.90	12.01	3.00	1.20	2.40	4.20	333
COI	5.72	0.39	2.76	1.58	8.88	7.89	2.96	10.65	1.78	11.64	6.34	4.93	5.13	1.97	1.38	9.47	5.52	4.14	2.56	4.34	507
COII	3.59	0.9	3.59	4.48	7.17	2.69	2.24	13.9	2.24	12.11	4.48	9.87	4.48	4.04	3.14	6.73	4.48	3.59	2.69	3.59	223
ATP8	0.00	0.00	0.00	0.00	12.96	0.00	0.00	20.37	9.26	5.56	9.26	12.96	5.56	1.85	0.00	3.70	1.85	0.00	5.56	11.11	54
ATP6	2.23	0.00	0.45	1.79	9.82	4.46	2.23	12.95	0.89	17.41	4.91	7.14	4.91	2.23	1.34	11.16	7.14	4.02	2.23	2.68	224
COIII	5.02	0.77	1.54	2.70	11.20	6.56	5.02	12.74	1.54	10.42	2.70	4.25	3.86	2.32	1.54	6.18	7.72	3.47	5.02	5.41	259
ND1	2.88	1.28	2.24	3.21	10.26	5.13	0.00	9.62	3.53	18.59	5.13	4.49	1.92	1.28	1.60	11.86	3.21	4.49	1.92	7.37	312
Cytb	4.46	0.79	1.84	1.05	9.45	5.77	2.10	14.17	2.89	14.44	3.67	7.09	5.77	2.36	1.84	5.51	5.25	3.67	3.15	4.72	381
ND6	1.14	0.57	0.57	0.57	12.57	2.86	0.57	15.43	4.57	18.29	7.43	10.86	1.14	1.14	0.57	6.29	3.43	2.86	1.14	8.00	175
ND4L	1.09	1.09	1.09	4.35	10.87	4.35	3.26	15.22	3.26	18.48	11.96	3.26	0.00	1.09	1.09	9.78	0.00	5.43	0.00	4.43	92
ND4	3.15	2.02	0.90	2.02	9.66	6.07	1.35	9.66	0.36	17.30	10.56	4.94	2.47	1.35	1.12	8.99	1.80	3.15	2.70	7.19	445
ND5	3.29	0.87	2.25	1.91	9.36	4.16	0.69	10.75	3.81	16.64	9.88	5.72	1.91	1.04	0.87	11.44	2.60	4.51	1.56	6.76	577
ND3	1.72	1.72	2.59	3.45	11.21	1.72	1.72	19.83	6.03	12.96	5.17	6.03	3.45	0.86	1.72	10.34	3.45	1.72	2.59	1.72	116
Total	36.39	11.60	20.12	28.61	141.13	54.96	22.74	179.10	45.27	185.0	89.90	89.95	42.10	22.73	17.11	103.12	46.00	42.25	33.52	71.52	3698

TABLE 6: The codon usage of the skipper mt genome.  $n$ : frequency of codon used; RSCU: relative synonymous codon usage; \* stop codon. Start codons and stop codons were excluded in total codon counts.

Codon (Aa)	$n$ (RSCU)	Codon (Aa)	$n$ (RSCU)	Codon (a)	$n$ (RSCU)	Codon (Aa)	$n$ (RSCU)
UUU (F)	351.0 (1.81)	UCU (S)	116.0 (2.71)	UAU (Y)	192.0 (1.90)	UGU (C)	34.0 (1.94)
UUC (F)	37.0 (0.19)	UCC (S)	16.0 (0.37)	UAC (Y)	10.0 (0.10)	UGC (C)	1.0 (0.06)
UUA (L)	458.0 (5.07)	UCA (S)	84.0 (1.96)	UAA (*)	0.0 (0.00)	UGA (W)	86.0 (1.87)
UUG (L)	13.0 (0.14)	UCG (S)	7.0 (0.16)	UAG (*)	0.0 (0.00)	UGG (W)	6.0 (0.13)
CUU (L)	45.0 (0.50)	CCU (P)	74.0 (2.45)	CAU (H)	55.0 (1.72)	CGU (R)	15.0 (1.20)
CUC (L)	6.0 (0.07)	CCC (P)	17.0 (0.56)	CAC (H)	9.0 (0.28)	CGC (R)	0.0 (0.00)
CUA (L)	19.0 (0.21)	CCA (P)	29.0 (0.96)	CAA (Q)	63.0 (1.97)	CGA (R)	31.0 (2.48)
CUG (L)	1.0 (0.01)	CCG (P)	1.0 (0.03)	CAG (Q)	1.0 (0.03)	CGG (R)	4.0 (0.32)
AUU (I)	431.0 (1.89)	ACU (T)	94.0 (2.54)	AAU (N)	209.0 (1.79)	AGU (S)	36.0 (0.84)
AUC (I)	26.0 (0.11)	ACC (T)	7.0 (0.19)	AAC (N)	25.0 (0.21)	AGC (S)	3.0 (0.07)
AUA (M)	245.0 (1.91)	ACA (T)	45.0 (1.22)	AAA (K)	106.0 (1.77)	AGA (S)	74.0 (1.73)
AUG (M)	12.0 (0.09)	ACG (T)	2.0 (0.05)	AAG (K)	14.0 (0.23)	AGG (S)	6.0 (0.14)
GUU (V)	61.0 (1.86)	GCU (A)	77.0 (2.44)	GAU (D)	58.0 (1.81)	GGU (G)	46.0 (1.00)
GUC (V)	1.0 (0.03)	GCC (A)	13.0 (0.41)	GAC (D)	6.0 (0.19)	GGC (G)	5.0 (0.11)
GUA (V)	61.0 (1.86)	GCA (A)	33.0 (1.05)	GAA (E)	71.0 (1.84)	GGA (G)	109.0 (2.37)
GUG (V)	8.0 (0.24)	GCG (A)	3.0 (0.10)	GAG (E)	6.0 (0.16)	GGG (G)	24.0 (0.52)

resided between the two trnS, which is almost identical to that in *C. raphaelis*.

The skipper mt genes are overlapped in a total of 51 bp at 12 locations, with the longest overlap of 7 bp, which is located between the ATP8 and ATP6. The overlapped nucleotides are shared in all the lepidopteran mt genomes so far examined and are probably helpful to form the structure of hairpin loop for posttranslation modifications [6, 43]. Other overlap regions include the 5 bp nucleotides between the trnW and trnC, 3 bp nucleotides between the trnY and COI, and the remaining 9 overlapping nucleotides are shorter than 3 bp (Table 4). Of the overlaps, the one between the trnF and ND5 is only 1 bp in size, substantially shorter than those of other lepidopteran species, such as the *A. pernyi* (15 bp), *C. boisduvalii* (17 bp), *E. pyretorum* (17 bp), *O. furnacalis* (15 bp), *O. nubilalis* (16 bp), *E. autonoe* (16 bp), and *A. honmai* (29 bp).

**3.5. The AT-Rich Region.** The 429 bp AT-rich region with the AT content of 88.11% is located between the srRNA and trnM. This region includes the O<sub>N</sub> (origin of minority or light strand replication) which has the motif ATAGA and is followed by 19 bp poly-T that is conserved across all lepidopterans. The AT-rich region also includes a few short microsatellite-like repeat regions such as poly-A, poly-T, and (AT)<sub>3</sub> (TA)<sub>9</sub> prior to the structural motif ATTTA or ATTA that is common to all the lepidopteran mt genomes. In addition, for the first time, we found two large repeated elements, the 20 bp repeat (TATAATTTAATAAAAAAATA), in the AT-rich region in a lepidopteran mt genome (Figure 3), similar observation was also reported in satyrid *Eumenis autonoe*, despite that their functions still remain unknown [11].

#### 4. Phylogenetic Analysis

A total of 32 lepidopteran species represent seven lepidopteran superfamilies (Tortricoidea, Pyraloidea, Papilionoidea, Hesperioidea, Bombycoidea, Geometroidea, and Noctuoidea). Of them, the superfamilies Papilionoidea and Hesperioidea are usually referred to as the macrolepidopterans together with the superfamilies Bombycoidea, Geometroidea, and Noctuoidea. Their phylogenetic relationships have long been a subject of controversy [58–61]. One of the most compelling hypotheses is that the Hesperioidea and Papilionoidea are closely related to each other [58, 60].

The ML and BI phylogenetic trees in this study based on 13 protein coding genes of the mitochondrial genomes revealed two major clusters: the first includes Bombycoidea, Geometroidea, Noctuoidea, Tortricoidea, and Cramboidea, and the second contains Papilionoidea and Hesperioidea with strong support (Figure 4(a)). Because the first cluster was also revealed by Kim et al. [11] and Mutanen et al. [62], we here focus on the second cluster, in which the superfamily Papilionoidea appears to be paraphyletic (Figure 4(a)) or polyphyletic (Figure 4(b)), and the superfamily Hesperioidea is placed within the Papilionoidea as a sister to ((Pieridae + Lycaenidae) + Nymphalidae). This result is consistent with those reported earlier based on combined data sets of a few nuclear and mitochondrial genes [62–64]. It is, however, remarkably different from the traditional view where Papilionoidea and Hesperioidea are considered as two distinct superfamilies. In conclusion, on the contrary to the common view that Papilionoidea and Hesperioidea are two separate groups in light of morphological and behavioral characteristics, our new mt genome data support that, while the majority of the Papilionoidea lineages is a sister group



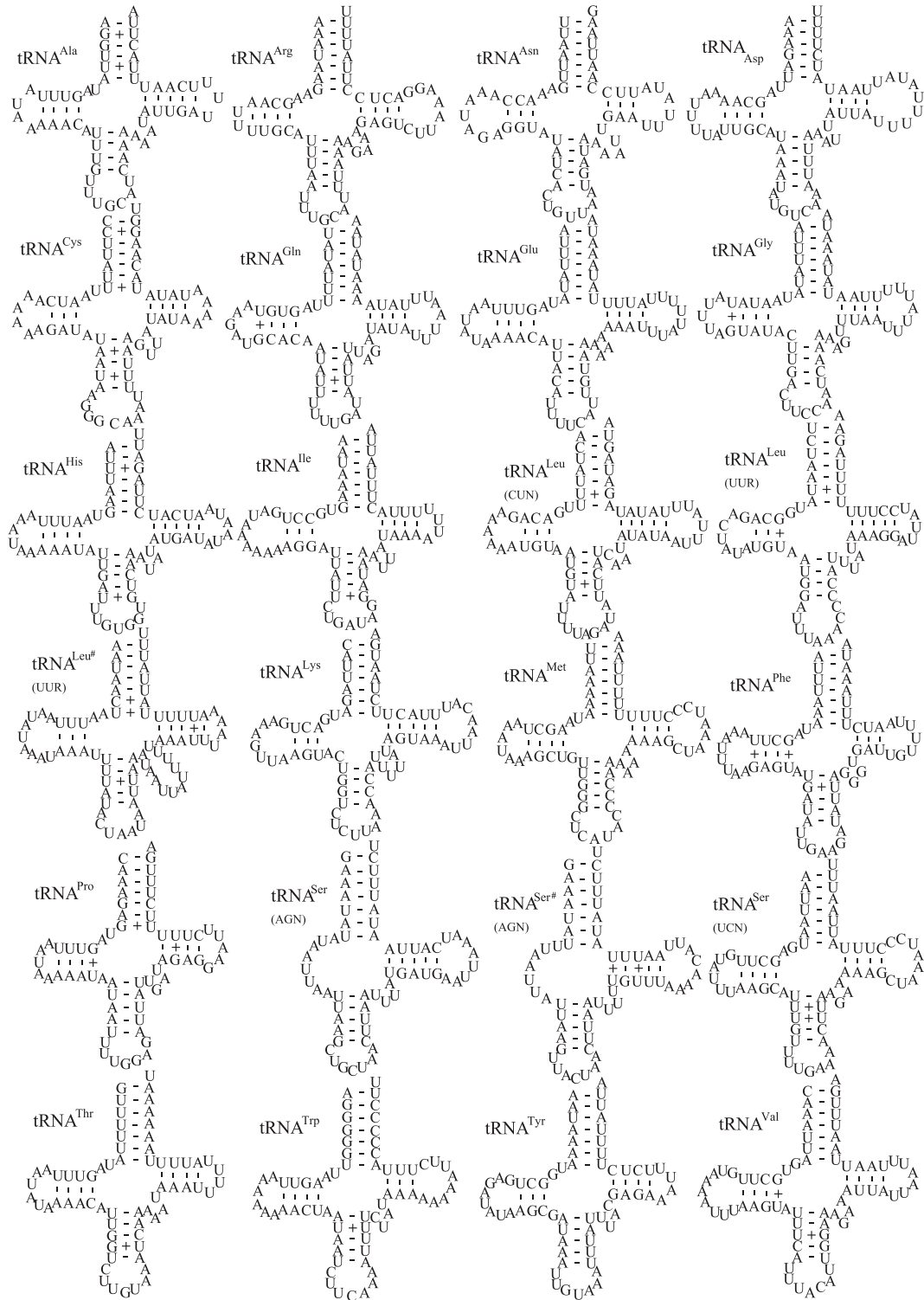


FIGURE 2: Predicted secondary clover-leaf structure for the tRNA genes of *Ctenoptilum vasava*. Dashes: Watson-Crick base pairing; centered dots (+): unmatched base pairing.

to Hesperioidea, the Papilionoidea is a separate clade among the two superfamilies. Our results call for more data from Hesperioidea to establish the accurate phylogeny of skippers (Hesperioidea) and true butterflies (Papilionoidea).

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants no. 41172004 and 40902004),

CTATATATAAAATTTTAAATAGAAATTTTTTTTTTTTATTTAAATAAAATATATTTAATACAACTTTACAACATATA  
 TCCAGTATAGGATTATTTCACTGATTCTCAAATTTTATGTCTGATATTAATAATATATATTTATCATAATATTAATTTACT  
 TATATAATGAAATTAATAATTTATAAATTTTATAATTTATAATTTATAATTTATAATTTATAATTTATAATTTATA  
 TTAATATATATATATATATATATATTTAACTAAGCCTAAATAGCTCTTTTGTTTAAATTTAAATAACAGAGACCTTATTTGAA  
 GATTTTATCTGAAATAAAGTAAGGAATACCTAAATTTATAGATCAGCTTAATAATTTTCTTCTAAATATAAA  
 Repeat1 15254 TATAATTTAAATAAAAAATA 15273  
 Repeat2 15274 TATATATAATAAAAAATA 15293

FIGURE 3: Repeated segments in A+T-rich region in the skipper mt genome. The consensus pattern (20 bp) repeat unit (TATAATT-TAATAAAAAAATA) is highlighted in red and purple. Numbers are shown as the positions of starting nucleotides.

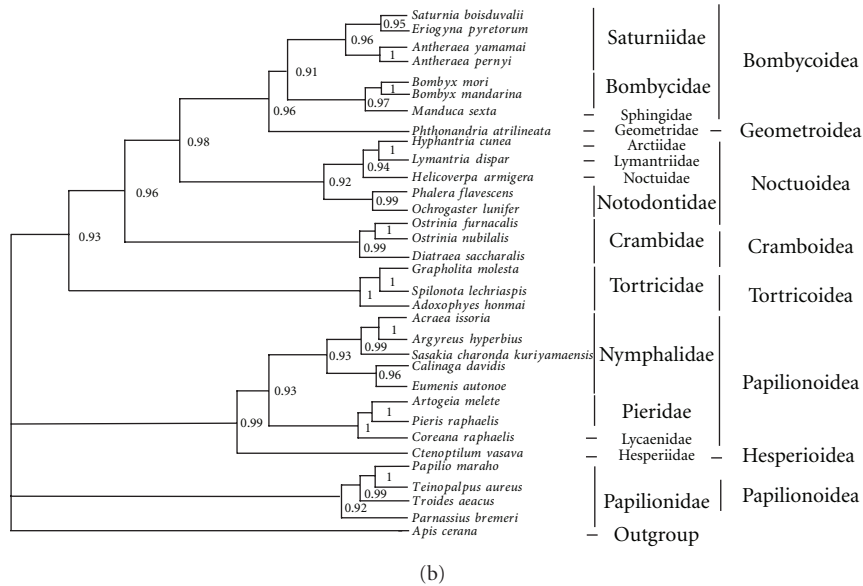
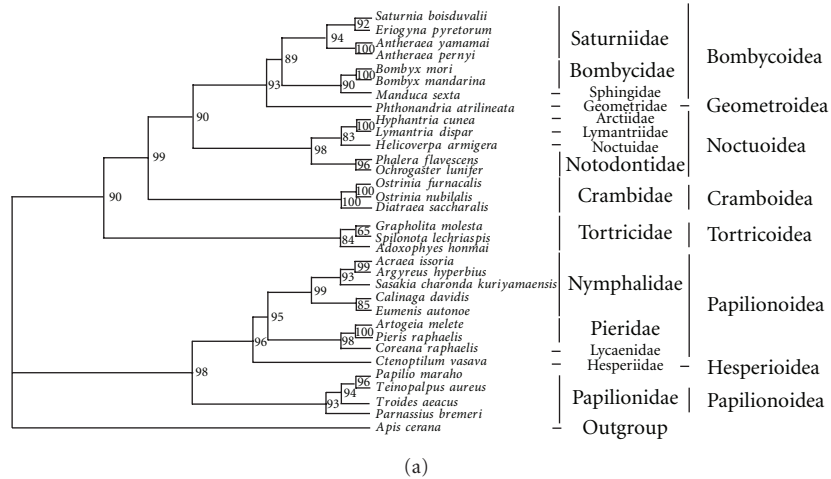


FIGURE 4: Phylogenetic trees of the lepidopterans based on 13 PCG nucleotide sequences. (a) ML tree. (b) BI tree. Numbers at each node indicate bootstrap percentages of 100 pseudoreplicates for ML analysis and posterior probability for BI analysis.

the CAS/SAFEA International Partnership Program for Creative Research Teams, Chinese Academy of Sciences (Grant no. KZCX22YW2JC104), the Provincial Key Project of the Natural Science Foundation from the Anhui Province, China (Grant no. KJ2010A142), and the Opening Funds from the State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences. The authors also thank Dr. Gang Liu

(School of Resources and Environmental Engineering, Anhui University) for his assistance in the laboratory and valuable comments on an earlier version of this paper.

## References

- [1] C. A. Bridges, *Catalogue of Hesperidae (Lepidoptera: Rhopalocera)*, Urbana, Ill, USA, 1988.

- [2] P. R. Ackery, R. de Jong, and R. I. Vane-Wright, "The butterflies: hedyloidea, hesperioidea and papilionoidea," in *Handbook of Zoology, A Natural History of the Phyla of the Animal Kingdom Vol. 4, Arthropoda: Insecta—Part 35, Lepidoptera, Moths and Butterflies, Evolution, Systematics and Biogeography*, N. P. Kristensen, Ed., vol. 1, pp. 263–300, 1999.
- [3] P. Y. Gorbunov, *The butterflies of Russia: classification, genitalia, keys for identification (Lepidoptera: Hesperioidea and Papilionoidea)*, Ph.D. thesis, Tezis Izdatel Stvo, Ekaterinburg, Russia, 2001.
- [4] H. Chiba and J. N. Eliot, "A review of the genus *Parnara* Moore (Lepidoptera: Hesperidae), with special reference to the Asian species," *Tyo to Ga*, vol. 42, no. 3, pp. 179–194, 1991.
- [5] I. O. Chou, *Classification and Identification of Chinese Butterflies*, Henan Scientific and Technological Publishing House, Zhengzhou, China, 1998.
- [6] I. Kim, E. M. Lee, K. Y. Seol et al., "The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae)," *Insect Molecular Biology*, vol. 15, no. 2, pp. 217–225, 2006.
- [7] G. Y. Hong, S. T. Jiang, M. Yu et al., "The complete nucleotide sequence of the mitochondrial genome of the cabbage butterfly, *Artogeia melete* (Lepidoptera: Pieridae)," *Acta Biochimica et Biophysica Sinica*, vol. 41, no. 6, pp. 446–455, 2009.
- [8] M. I. Kim, J. Y. Baek, M. J. Kim et al., "Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects," *Molecules and Cells*, vol. 28, no. 4, pp. 347–363, 2009.
- [9] J. Hu, J. S. Hao, D. X. Zhang, D. Huang, S. Cameron, and C. D. Zhu, "The complete mitochondrial genome of the yellow coaster, *Acraea issoria* (Lepidoptera: Nymphalidae: Heliconiinae: Acraeini): Sequence, gene organization and a unique tRNA translocation event," *Molecular Biology Reports*, vol. 37, no. 7, pp. 3431–3438, 2010.
- [10] Z. H. Mao, J. S. Hao, G. P. Zhu, J. Hu, M. M. Si, and C. D. Zhu, "Sequencing and analysis of the complete mitochondrial genome of *Pieris rapae* Linnaeus (Lepidoptera: Pieridae)," *Acta Entomologica Sinica*, vol. 53, no. 11, pp. 1295–1304, 2010.
- [11] M. J. Kim, X. L. Wan, K. G. Kim, J. S. Hwang, and I. Kim, "Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae)," *African Journal of Biotechnology*, vol. 9, no. 5, pp. 735–754, 2010.
- [12] J. Xia, J. Hu, G. P. Zhu, C. D. Zhu, and J. S. Hao, "Complete mitochondrial DNA sequence of the *Calinaga davidis*," *Acta Entomologica Sinica*, vol. 54, no. 5, pp. 555–565, 2011.
- [13] K. Yukuhiro, H. Sezutsu, M. Itoh, K. Shimizu, and Y. Banno, "Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkworm, *Bombyx mandarina*, and its close relative, the domesticated silkworm, *Bombyx mori*," *Molecular Biology and Evolution*, vol. 19, no. 8, pp. 1385–1389, 2002.
- [14] E. S. Lee, K. S. Shin, M. S. Kim, H. Park, S. Cho, and C. B. Kim, "The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae)," *Gene*, vol. 373, no. 1-2, pp. 52–57, 2006.
- [15] Y. Liu, Y. Li, M. Pan et al., "The complete mitochondrial genome of the Chinese oak silkworm, *Antheraea pernyi* (Lepidoptera: Saturniidae)," *Acta Biochimica et Biophysica Sinica*, vol. 40, no. 8, pp. 693–703, 2008.
- [16] P. Salvato, M. Simonato, A. Battisti, and E. Negrisolo, "The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae)," *BMC Genomics*, vol. 9, article 331, 2008.
- [17] S. L. Cameron and M. F. Whiting, "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*, vol. 408, no. 1-2, pp. 112–123, 2008.
- [18] L. Yang, Z. J. Wei, G. Y. Hong, S. T. Jiang, and L. P. Wen, "The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae)," *Molecular Biology Reports*, vol. 36, no. 6, pp. 1441–1449, 2009.
- [19] S. T. Jiang, G. Y. Hong, M. Yu et al., "Characterization of the complete mitochondrial genome of the giant silkworm moth, *Eriogyna pyretorum* (Lepidoptera: Saturniidae)," *International Journal of Biological Sciences*, vol. 5, no. 4, pp. 351–365, 2009.
- [20] S. R. Kim, M. I. Kim, M. Y. Hong et al., "The complete mitogenome sequence of the Japanese oak silkworm, *Antheraea yamamai* (Lepidoptera: Saturniidae)," *Molecular Biology Reports*, vol. 36, no. 7, pp. 1871–1880, 2009.
- [21] M. Y. Hong, E. M. Lee, Y. H. Jo et al., "Complete nucleotide sequence and organization of the mitogenome of the silk moth *Caligula boisduvalii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects," *Gene*, vol. 413, no. 1-2, pp. 49–57, 2008.
- [22] L. Fang, L. Wang, S. Wu et al., "The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae)," *International Journal of Biological Sciences*, vol. 6, no. 2, pp. 172–186, 2010.
- [23] Y. J. Zhu, R. Fang, G. L. Zhou, J. Ye, and J. P. Yi, "The complete sequence determination and analysis of *Lymantria dispar* mitochondrial genome," *National Center for Biotechnology Information*, vol. 24, no. 4, pp. 6–11, 2009.
- [24] J. Yin, G. Y. Hong, A. M. Wang, Y. Z. Cao, and Z. J. Wei, "Mitochondrial genome of the cotton bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae) and comparison with other Lepidopterans," *Mitochondrial DNA*, vol. 21, no. 5, pp. 160–169, 2010.
- [25] J. L. Zhao, Y. Y. Zhang, A. R. Luo, G. F. Jiang, S. L. Cameron, and C. D. Zhu, "The complete mitochondrial genome of *Spilonota lechriaspis* Meyrick (Lepidoptera: Tortricidae)," *Molecular Biology Reports*, vol. 38, pp. 3757–3764, 2011.
- [26] Y. J. Gong, B. C. Shi, Z. J. Kang, F. Zhang, and S. J. Wei, "The complete mitochondrial genome of the oriental fruit moth *Grapholita molesta* (Busck) (Lepidoptera: Tortricidae)," *Molecular Biology Reports*, vol. 39, no. 3, pp. 2893–2900, 2011.
- [27] M. H. Pan, Q. Y. Yu, Y. L. Xia et al., "Characterization of mitochondrial genome of Chinese wild mulberry silkworm, *Bombyx mandarina* (Lepidoptera: Bombycidae)," *Science in China C*, vol. 51, no. 8, pp. 693–701, 2008.
- [28] W. W. Li, X. Y. Zhang, and Z. X. Fan, "Structural characteristics and phylogenetic analysis of the mitochondrial genome of the Sugarcane Borer, *Diatraea saccharalis* (Lepidoptera: Crambidae)," *DNA and Cell Biology*, vol. 30, no. 1, pp. 3–8, 2011.
- [29] B. S. Coates, D. V. Sumerford, R. L. Hellmich, and L. C. Lewis, "Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis*," *International Journal of Biological Sciences*, vol. 1, no. 1, pp. 13–18, 2005.
- [30] X. C. Wang, X. Y. Sun, Q. Q. Sun et al., "The complete mitochondrial genome of the laced fritillary *Argyreus hyperbius* (Lepidoptera: Nymphalidae)," *Zoology Research*, vol. 32, no. 5, pp. 465–475, 2011.
- [31] X. Feng, D. F. Liu, N. X. Wang, C. D. Zhu, and G. F. Jiang, "The mitochondrial genome of the butterfly *Papilio xuthus* (Lepidoptera: Papilionidae) and related phylogenetic analyses,"

- Molecular Biology Reports*, vol. 37, no. 8, pp. 3877–3888, 2010.
- [32] J. S. Hao, C. X. Li, X. Y. Sun, and Q. Yang, “Phylogeny and divergence time estimation of cheilostome bryozoans based on mitochondrial 16S rRNA sequences,” *Chinese Science Bulletin*, vol. 12, pp. 1205–1211, 2005.
- [33] C. Simon, F. Frati, A. Bekenbach, B. Crespi, H. Liu, and P. Flook, “Evolution, weighting, and phylogenetic utility of mitochondrial genes and a compilation of conserved polymerase chain reaction primers,” *Annals of the Entomological Society of America*, vol. 87, no. 6, pp. 651–701, 1994.
- [34] V. K. Singh, A. K. Mangalam, S. Dwivedi, and S. Naik, “Primer premier: Program for design of degenerate primers from a protein sequence,” *BioTechniques*, vol. 24, no. 2, pp. 318–319, 1998.
- [35] T. A. Hall, “BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT,” *Nucleic Acids Symposium Series*, no. 41, pp. 95–98, 1999.
- [36] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins, “The clustal X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools,” *Nucleic Acids Research*, vol. 24, pp. 4876–4882, 1997.
- [37] S. Kumar, K. Tamura, and M. Nei, “MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment,” *Briefings in Bioinformatics*, vol. 5, no. 2, pp. 150–163, 2004.
- [38] T. M. Lowe and S. R. Eddy, “tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence,” *Nucleic Acids Research*, vol. 25, no. 5, pp. 955–964, 1997.
- [39] G. Benson, “Tandem repeats finder: a program to analyze DNA sequences,” *Nucleic Acids Research*, vol. 27, no. 2, pp. 573–580, 1999.
- [40] D. L. Swofford, *PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods), Version 4.10*, Sinauer Associates, Sunderland, Mass, USA, 2002.
- [41] D. Posada and K. A. Crandall, “Modeltest: testing the model of DNA substitution,” *Bioinformatics*, vol. 14, no. 9, pp. 817–818, 1998.
- [42] J. P. Huelsenbeck and F. Ronquist, “MrBayes: bayesian inference of phylogenetic trees,” *Bioinformatics*, vol. 17, no. 8, pp. 754–755, 2001.
- [43] J. D. Fenn, S. L. Cameron, and M. F. Whiting, “The complete mitochondrial genome sequence of the Mormon cricket (*Anabrus simplex*: Tettigoniidae: Orthoptera) and an analysis of control region variability,” *Insect Molecular Biology*, vol. 16, no. 2, pp. 239–252, 2007.
- [44] M. M. Yamauchi, M. U. Miya, and M. Nishida, “Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda),” *Gene*, vol. 295, no. 1, pp. 89–96, 2002.
- [45] F. Nardi, A. Carapelli, R. Dallai, and F. Frati, “The mitochondrial genome of the olive fly *Bactrocera oleae* two haplotypes from distant geographical locations,” *Insect Molecular Biology*, vol. 12, no. 6, pp. 605–611, 2003.
- [46] S. Lutz-Bonengel, T. Sanger, and S. Pollak, “Different methods to determine length heteroplasmy within the mitochondrial control region,” *International Journal of Legal Medicine*, vol. 118, no. 5, pp. 274–281, 2004.
- [47] K. Ogoh and Y. Ohmiya, “Complete mitochondrial DNA sequence of the sea-firefly, *Vargula hilgendorffii* (Crustacea, Ostracoda) with duplicate control regions,” *Gene*, vol. 327, no. 1, pp. 131–139, 2004.
- [48] P. K. Flook, C. H. F. Rowell, and G. Gellissen, “The sequence, organization, and evolution of the *Locusta migratoria* mitochondrial genome,” *Journal of Molecular Evolution*, vol. 41, no. 6, pp. 928–941, 1995.
- [49] C. D. Beard, S. M. Hamm, and F. H. Collins, “The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects,” *Insect Molecular Biology*, vol. 2, no. 2, pp. 103–124, 1993.
- [50] S. E. Mitchell, A. F. Cockburn, and J. A. Seawright, “The mitochondrial genome of *Anopheles quadrimaculatus* species A: complete nucleotide sequence and gene organization,” *Genome*, vol. 36, no. 6, pp. 1058–1073, 1993.
- [51] L. Spanos, G. Koutroumbas, M. Kotsyfakis, and C. Louis, “The mitochondrial genome of the Mediterranean fruit fly, *Ceratitis capitata*,” *Insect Molecular Biology*, vol. 9, no. 2, pp. 139–144, 2000.
- [52] K. Wilson, V. Cahill, E. Ballment, and J. Benzie, “The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods?” *Molecular Biology and Evolution*, vol. 17, no. 6, pp. 863–874, 2000.
- [53] D. Ojala, C. Merkel, R. Gelfand, and G. Attardi, “The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA,” *Cell*, vol. 22, no. 2, pp. 393–403, 1980.
- [54] D. Ojala, J. Montoya, and G. Attardi, “tRNA punctuation model of RNA processing in human mitochondria,” *Nature*, vol. 290, no. 5806, pp. 470–474, 1981.
- [55] C. Lu, Y. Q. Liu, S. Y. Liao et al., “Complete sequence determination and analysis of *Bombyx mori* mitochondrial genome,” *Journal of Agricultural Biotechnology*, vol. 10, no. 2, pp. 163–170, 2002.
- [56] M. F. J. Taylor, S. W. Mckechnie, N. Pierce, and M. Kreitman, “The lepidopteran mitochondrial control region: structure and evolution,” *Molecular Biology and Evolution*, vol. 10, no. 6, pp. 1259–1272, 1993.
- [57] J. W. Taanman, “The mitochondrial genome: structure, transcription, translation and replication,” *Biochimica et Biophysica Acta*, vol. 1410, no. 2, pp. 103–123, 1999.
- [58] J. Minet, “Tentative reconstruction of the ditrysian phylogeny (Lepidoptera: Glossata),” *Entomologica Scandinavica*, vol. 22, pp. 69–95, 1991.
- [59] J. Minet, “The Bombycoidea: phylogeny and higher classification (Lepidoptera : Glossata),” *Entomologica Scandinavica*, vol. 25, no. 1, pp. 63–88, 1994.
- [60] E. S. Nielsen, “Phylogeny of major lepidopteran groups,” in *The Hierarchy of Life*, B. Fernholm, K. Bremer, and H. Jörnvall, Eds., pp. 281–294, Elsevier, Amsterdam, The Netherlands, 1989.
- [61] J. A. Scott, “On the monophyly of the macrolepidoptera, including a reassessment of their relationship to cossioidea and castnioidea, and a reassignment of mimallonidae to pyraloidea,” *The Journal of Research on the Lepidoptera*, vol. 25, no. 1, pp. 30–38, 1986.
- [62] M. Mutanen, N. Wahlberg, and L. Kaila, “Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies,” *Proceedings of the Royal Society B*, vol. 277, no. 1695, pp. 2839–2848, 2010.
- [63] N. Wahlberg, M. F. Braby, A. V. Z. Brower et al., “Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers,” *Proceedings of the Royal Society B*, vol. 272, no. 1572, pp. 1577–1586, 2005.

- [64] J. C. Regier, C. P. Cook, C. Mitter, and A. Hussey, "A phylogenetic study of the 'bombycoid complex' (Lepidoptera) using five protein-coding nuclear genes, with comments on the problem of macrolepidopteran phylogeny," *Systematic Entomology*, vol. 33, no. 1, pp. 175–189, 2008.