# The Complete Mitochondrial Genomes of the Fenton's Wood White, Leptidea morsei, and the Lemon Emigrant, Catopsilia pomona

Authors: Hao, Juan-Juan, Hao, Jia-Sheng, Sun, Xiao-Yan, Zhang, Lan-Lan, and Yang, Qun

Source: Journal of Insect Science, 14(130) : 1-22

Published By: Entomological Society of America

URL: https://doi.org/10.1673/031.014.130

Journal of Insect Science
www.insectscience.org

# The complete mitochondrial genomes of the Fenton's wood white, *Leptidea morsei*, and the lemon emigrant, *Catopsilia pomona*

Juan-Juan Hao,[1a] Jia-Sheng Hao,[1b*] Xiao-Yan Sun,[2c] Lan-Lan Zhang,[1d] and Qun Yang[2e*]

[1]College of Life Sciences, Anhui Normal University, 241000 Wuhu, China
[2]State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing 210008, China

## Abstract

The complete mitochondrial genomes of *Leptidea morsei* Fenton (Lepidoptera: Pieridae: Dismorphiinae) and *Catopsilia pomona* (F.) (Lepidoptera: Pieridae: Coliadinae) were determined to be 15,122 and 15,142 bp in length, respectively, with that of *L. morsei* being the smallest among all known butterflies. Both mitogenomes contained 37 genes and an A+T-rich region, with the gene order identical to those of other butterflies, except for the presence of a tRNA-like insertion, $tRNA^{Leu}$ (UUR), in *C. pomona*. The nucleotide compositions of both genomes were higher in A and T (80.2% for *L. morsei* and 81.3% for *C. pomona*) than C and G; the A+T bias had a significant effect on the codon usage and the amino acid composition. The protein-coding genes utilized the standard mitochondrial start codon ATN, except the *COI* gene using CGA as the initiation codon, as reported in other butterflies. The intergenic spacer sequence between the $tRNA^{Ser}$ (UCN) and *ND1* genes contained the ATACTAA motif. The A+T-rich region harbored a poly-T stretch and a conserved ATAGA motif located at the end of the region. In addition, there was a triplicated 23 bp repeat and a microsatellite-like $(TA)_9(AT)_3$ element in the A+T-rich region of the *L. morsei* mitogenome, while in *C. pomona,* there was a duplicated 24 bp repeat element and a microsatellite-like $(TA)_9$ element. The phylogenetic trees of the main butterfly lineages (Hesperiidae, Papilionidae, Pieridae, Nymphalidae, Lycaenidae, and Riodinidae) were reconstructed with maximum likelihood and Bayesian inference methods based on the 13 concatenated nucleotide sequences of protein-coding genes, and both trees showed that the Pieridae family is sister to Lycaenidae. Although this result contradicts the traditional morphologically based views, it agrees with other recent studies based on mitochondrial genomic data.

**Keywords:** mitochondrial genome, Pieridae, phylogenetic analysis
**Correspondence:** [a] haojuan523@126.com, [b] jshaonigpas@sina.com, [c] ssd6sun@yahoo.com.cn, [d] zhanglanlan0531@163.com, [e] qunyang@nigpas.ac.cn, *Corresponding authors
**Editor:** Yoonseong Park was editor of this paper.
**Received:** 8 October 2012 **Accepted:** 5 June 2013 **Published:** 1 October 2014

**ISSN:** 1536-2442 | Vol. 14, Number 130

**Cite this paper as:**
Hao J-J, Hao J-S, Sun X-Y, Zhang L-L, Yang Q. 2014. The complete mitochondrial genomes of the Fenton's wood white, *Leptidea morsei*, and the lemon emigrant, *Catopsilia pomona*. *Journal of Insect Science* 14(130). Available online: http://www.insectscience.org/14.130

## Introduction

The animal mitochondrial genomes (mitogenomes) are usually circular molecules of 14-19 kb in size, containing 37 genes (including 13 protein-coding genes, two rRNA genes, and 22 tRNA genes) and a non-coding A+T-rich region that regulates the transcription and replication of the mitogenome (Clayton 1992). Due to its simple and compact structure, fast evolutionary rate, and maternal inheritage, it has been used frequently in the studies of population genetics, molecular evolution, phylogenetics, phylogeography, and evolutionary biology (Simon et al. 1994). In recent years, as the DNA sequencing technology has been progressing rapidly, more and more complete animal mitogenome sequences have been determined. To date, more than 240 complete or near-complete mitochondrial DNA sequences have been identified from insects, including 67 from lepidopterans. Of these, the available sequences are mainly from six superfamilies (Bombycoidea, Geometroidea, Papilionoidea, Noctuoidea, Tortricoidea, and Pyraloidea). In total, 33 of these lepidopteran mitogenomes are from butterflies (Table 1).

Pieridae is one of the largest families of Papilionoidea, containing 76 genera and approximately 1,100 species worldwide, mostly distributed in tropical Africa and Asia (Courtney 1986, Watt et al. 1996, Brunton 1998, Stavenga et al. 2004, Kemp et al. 2005). Their adults are generally of medium size and typically white, orange, and yellow in color (Chapman 1895). Taxonomically, they are currently divided into four subfamilies (Dismorphiinae, Pierinae, Coliadinae, and Pseudopontiinae). In addition, phylogenetically, they may stand as a key group to clarify the intra-familiar butterfly relationships. For example, they were traditionally considered to be the sister to the Papilionidae (Ehrlich 1958, Scott 1985). However, more and more evidence indicated that they were sister to the grouping of (Nymphalidae (Riodinidae, Lycaenidae)) (Kristensen 1976, de Jong et al. 1996, Weller et al. 1996, Ackery et al. 1999, Wahlberg et al. 2005) or sister to (Riodinidae + Lycaenidae) (Kim, M. J. et al. 2010, Chai et al. 2012). To our dissatisfaction, up to the present, only four mitogenomes of pierid species (*Artogeia melete* [Menetries], *Pieris rapae* [L.], *Delias hyparete* [L.], and *Aporia crataegi* [L.]) are available, and thus more pierid mitogenome data are needed to enrich the taxon sampling for use in phylogenetic studies.

The Fenton's wood white, *Leptidea morsei* Fenton, and the lemon emigrant, *Catopsilia pomona* (F.), are the two representative species of the family Pieridae. *Leptidea morsei* is distributed mainly throughout Europe, Siberia, Ussuri, Korea, northern China, and Japan. It is found occasionally in damp, grassy vegetation at the sunny edges of woods, as well as in grassy woodland. Its larvae feed on peas, and adults are seen twice per year from April to May and June to July. *Catopsilia pomona* is ubiquitously distributed from areas of southeast Asia (Sikkim, Malaysia, Philippines) to Australia. It is found often in secondary forests, along river courses, and even in the hot arid deserts throughout the year. Its colors are usually variable, chiefly lemon-yellow with an apical black margin (Rienks 1985).

In this study, we determined and analyzed the complete mitogenome sequences of these two pierid species and compared these sequences with those of other butterfly species available to clarify the phylogenetic relationships among the main butterfly lineages. The new sequence data will provide valuable information for the studies of lepidopteran comparative genomics, molecular evolution, and other relevant areas.

## Materials and Methods

### Sample collection

Adult individuals of *L. morsei* and *C. pomona* were collected from Shanxi and Hainan Provinces, China, in August 2008 and July 2009, respectively. After sample collection, the fresh materials were placed into 100% ethanol immediately for DNA fixation and stored at -20°C until used for genomic DNA extraction.

### DNA extraction and amplification by PCR

Total genomic DNA of *L. morsei* and *C. pomona* was extracted from the thoracic muscle of an adult individual by using the glass bead method after Hao et al. (2005). Insect universal primers were used for the amplification of the *COI*, *CytB*, *16S rRNA*, and *12S rRNA* genes (Simon et al. 1994). Primers for the *ND2*, *ND4*, *COIII*, and *ND5* amplification were designed via the alignment of the respective sequences from all the butterflies available by using Clustal X1.8 and Primer Premier 5.0 softwares (Thompson et al. 1997, Singh et al. 1998). Seven long fragments (*COI–COIII*, *COIII–ND5*, *ND5–CytB*, *CytB–16S*, *16S–12S*, *12S–ND2*, *ND2–COI*) were amplified via long PCR using Takara LA Taq™ (Takara Co., www.takara-bio.com). The long PCR conditions were as follows: an initial denaturation at 95°C for 5 min, followed by 30 cycles of denaturation at 95°C for 50 sec, annealing at 50–55°C (depending on primer pairs) for 50 sec, and elongation at 68°C for 150 sec during the first 15 cycles and then an additional 5 sec per cycle during the last 15 cycles, and a final extension at 68°C for 10 min. All PCR fragments were sequenced directly in both strands after purification with the QIA quick PCR Purification Kit (QIAGEN, www.qiagen.com), except for the *12S–ND2* fragment of *L. morsei*, which was sequenced after cloning. All of the long PCR fragments were sequenced by using the primer walking strategy.

### Sequence analysis

The raw sequences from the overlapping fragments were proofread and assembled in BioEdit version 7.0 (Hall 1999). Protein-coding genes, rRNA genes, and A+T-rich regions were determined via the alignment of the sequences by using Clustal X1.8 software (Thompson et al. 1997). The nucleotide sequences of the protein-coding genes were translated based on the invertebrate mtDNA genetic code. The tRNAs were identified by tRNAscan-SE software version 1.21 (Lowe and Eddy 1997). The putative tRNAs that could not be found by tRNAscan-SE were confirmed by sequence comparisons between the Pieridae and other butterfly tRNAs. Nucleotide composition and codon usage were calculated by using MEGA5.1 software (Kumar et al. 2004), and the tandem repeats in the A+T-rich regions were predicted by the Tandem Repeats Finder available online (http://tandem.bu.edu/trf/trf.html) (Benson et al. 1999). Sequence data were deposited in the GenBank database under the accession numbers JX274648 for *L. morsei* and JX274649 for *C. pomona*.

### Phylogenetic analysis

For the phylogenetic analysis, 13 concatenated nucleotide sequences of protein-coding genes of 33 available butterfly mitogenome sequences (two newly sequenced in this study and 31 extracted from GenBank, Table 1) were aligned by using Clustal X1.8 (Thompson 1997). The phylogenetic trees were then reconstructed with the maximum likelihood and Bayesian inference methods using the moth species *Adoxophyes honmai* Yasuda (Lepidoptera: Tortricidae) (GenBank accession number DQ073916) as the outgroup.

In the maximum likelihood and Bayesian inference analyses, the third position of all the codons was excluded, and the best fitting substitution model GTR + I + G (Lanave et al. 1984) was selected via a comparison of Akaike Information Criterion scores (Akaike 1974), calculated by using the Modeltest software version 3.7 (Posada and Crandal 1998). The maximum likelihood analyses were conducted in PAUP version 4.0b8 (Swofford 2002) under the following conditions: tree searching by TBR (tree bisection and reconnection) branch swapping (10 random-addition sequences); specifying the number of substitution rate categories as four; and using a BIONJ distance-based tree as the starting tree. The confidence values of each node of the maximum likelihood tree were evaluated via the bootstrapping test with 1,000 iterations. Bayesian analyses were performed by using the program MrBayes 3.1 (Huelsenbeck and Ronquist 2001). Two independent runs of four incrementally heated MCMC chains (one cold chain and three hot chains) were simultaneously run for one million generations in all datasets. Each set was sampled every 100 generations with a burn-in of 25%, and when the average standard deviation of split frequencies was less than 0.01, stationarity was considered to be reached. The confidence values of the Bayesian inference tree were presented as the Bayesian posterior probabilities in percentages.

## Results and Discussion

### General features

The complete mtDNA sequences of *L. morsei* and *C. pomona* were 15,122 and 15,142 bp in length, respectively, with that of *L. morsei* being the shortest among all known sequences of butterfly species (Table 2). Each genome was composed of the typical 13 protein-coding genes, 22 tRNA genes, two rRNA genes, and one major non-coding A+T-rich region. The gene order was identical to that of other sequences of butterflies but different from that found in the ancestral insects with respect to the location of $tRNA^{Met}$. That is to say, the $tRNA^{Met}$ was located between the control region and $tRNA^{Ile}$, giving the derived control region (CR)-Met (M)-Ile (I)-Glu (Q) arrangement instead of that of the insect ground plan CR-I-Q-M (Fig. 1). The total sizes of protein-coding genes, rRNAs, and tRNAs were all well within the corresponding ranges of those found in other butterfly species (Table 2). The size proportions of coding genes to the whole genome of these two and four other pierid species were slightly higher than those of butterflies of other families (Table 3). By contrast, their non-coding sequences, including intergenic spacers and the A+T-rich region, were slightly shorter than those of other family taxa. The A+T-rich region of *C. pomona* was the shortest, whereas that of *Papilio maraho* Shiraki & Sonan (Lepidoptera: Papilionidae) was the longest (Table 4). The majority strand coding nine protein-coding genes and 14 tRNAs were 7,804 and 7,836 bp, respectively, whereas the minority strand coding four protein-coding genes, eight tRNAs, and two rRNA genes were 6,901 and 6,934 bp, respectively, for the two pierid species.

The nucleotide compositions of the two entire mitogenome sequences were biased significantly toward A and T (Table 5). These A+T contents were generally consistent with those of other butterfly mitogenomes, which ranged from 79.1% in *Eumenis autonoe* Esper (Lepidoptera: Nymphalidae) (Kim, M. J. et al. 2010) to 82.7% in *Coreana raphaelis* Oberthür (Lepidoptera: Lycaenidae) (Kim, I. et al. 2006). The base composition bias of an individual strand can be described by A+T skewness, caculated by (A%−T%)/(A%+T%),

and G+C skewness, calculated by (G%−C%)/(G%+C%). The A+T and G+C skewness values in majority strands were calculated to be -0.122 and -0.121, respectively, for *L. morsei* , and -0.119 and -0.087, respecitively, for *C. pomona*.

**Protein-coding genes**

All the protein-coding genes in *L. morsei* and *C. pomona* started with a typical ATN codon, with the only exception represented by the CGA start codon of the *COI* gene. For *L. morsei*, three genes (*ND2*, *ND5*, *ND6*) started with ATT, one (*ATP8*) with ATC, two (*ND3*, *ND4*) with ATA, and six (*ATP6*, *ND1*, *COII*, *COIII*, *ND4L*, *CytB*) with ATG. In comparison with *L. morsei, ATP8*, *ND3*, and *ND6* in *C. pomona* possessed different start codons, namely ATT, ATT, and ATC, respectively. These start codons were well-conserved in the sequenced butterfly mitogenomes. For instance, the *ND2* gene usually used ATT as the start codon, whereas the *COII* and *ATP6* genes frequently used ATG as the start codon (Table 6).

Nine protein-coding genes were terminated with the standard stop codon TAA, whereas the *COI*, *COII*, *ND4*, and *ND5* genes used T as a truncated stop codon. The only difference in stop codons between the two pierid species was found in the *ND3* gene, that is, *L. morsei* used TAA instead of TAG, which appeared in *C. pomona*. Furthermore, incomplete stop codons were detected frequently in the *COI*, *COII*, and *ND5* genes in most insects, including all sequenced butterfly species (Table 6). Incomplete stop codons would produce functional stop codons after polycistronic transcript cleavage and polyadenylation (Ojala et al. 1981).

Previous studies reported that most lepidopterans used the codon CGA as the start codon for *COI* (Fig. 2). However, exceptions have been reported; for example, TTG was proposed as the start codon in *Acraea issoria* Hübner (Lepidoptera: Nymphalidae) (Hu et al. 2010), *Caligula boisduvalii* Eversmann (Lepidoptera: Saturniidae) (Kim, I. et al. 2006), and *Fabriciana nerippe* Felder (Lepidoptera: Nymphalidae) (Kim, M. J. et al. 2011a); ATT in *A. crataegi* (Park et al. 2012) and *Ctenoptilum vasava* Moore (Lepidoptera: Hesperiidae) (Hao et al. 2012); TTAG in *Bombyx mori* L. (Lepidoptera: Bombycidae) (Yukuhiro et al. 2002) and *C. raphaelis* (Hong, G. Y. et al. 2009); and ATTTAG in *Ostrinia nubilalis* Hübner and *Ostrinia furnacalis* Guenée (Lepidoptera: Crambidae) (Coates et al. 2005). In this study, a typical ATN initiator for *COI* in *L. morsei* and *C. pomona* was not detected at their starting sites. The putative ATT start codon is commonly located upstream of the *COI* gene and frequently followed immediately by the TAG or TAA stop codon; thus, the CGA, not ATT, probably acted as the start codon here as in most butterflies.

Excluding stop codons, the A+T contents of protein-coding genes in *L. morsei* and *C. pomona* were 79.16% and 80.02% (Table 5), respectively, and these values were similar to those detected in other butterflies, which ranged from 76.8% in *E. autonoe* (Kim, M. J. et al. 2010) to 81.5% in *C. raphaelis* (Kim, I. et al. 2006) (Table 2). When the first, second, and third codon positions were considered separately, the highest A+T contents were in the third positions for *L. morsei* and *C. pomona*. In addition, the highest T contents were detected in the second positions, and the lowest G contents in the third positions (Table 5).

Exclusive of the stop codon, 3,713 and 3,724 amino acids were encoded by the mitogenomes of *L. morsei* and *C. pomona*, respectively (Table 2). The amino acid numbers were well within the size range of 3,586

in *Sasakia charonda kuriyamaensis* Shirozu (Lepidoptera: Nymphalidae) (Hakozaki et al., unpublished, GenBank accession number NC_014223.1) to 3,740 in *Abisara fylloides* Moore (Lepidoptera: Riodinidae) (Shi et al. unpublished, College of Life Sciences, Anhui Normal University, China) detected in other butterflies. Among the amino acids, UUU (Phe), UUA (Leu), AUU (Ile), AUA (Met), and AAU (Asn) were the most frequently used codons in *L. morsei* and *C. pomona* (Table 7), and similar codons were found in other butterfly species, such as *C. vasava* (Hao et al. 2012), *A. crataegi* (Park et al. 2012), and *Sericinus montela* Gray (Lepidoptera: Papilionidae) (Ji et al. 2012).

**Transfer RNAs and ribosomal RNAs**

There were 22 tRNA genes (two each for serine and leucine, and one for each of the other amino acids) identified within the two pierid mitogenomes. An additional tRNA-like sequence ($tRNA^{Leu}$ [UUR]) located within the *16S rRNA* gene was detected in the mitogenome of *C. pomona*. The 22 tRNA genes ranging in size from 60 to 71 bp were interspersed throughout the two whole mitogenomes. The total sizes of the *L. morsei* and *C. pomona* tRNAs were 1,416 and 1,446 bp, respectively, with 80.68 and 81.05% A+T contents, respectively. All tRNAs could be folded into the typical clover leaf secondary structure, whereas $tRNA^{Ser}$ (AGN) in both mitochondrial genomes lacked the dihydrouridine (DHU) loop. This feature has been shown in the majority of metazoan mitogenomes, including all those sequenced from butterflies (Kim, I. et al. 2006; Hong, G. Y. et al. 2009; Hu et al. 2010; Kim, M. I. et al. 2009; Xia et al. 2011; Wang et al. 2011; Kim, M. J. et al. 2010, 2011a, 2011b; Chen et al. 2012; Shi et al. 2012; Tian et al. 2012). The anticodon for $tRNA^{Ser}$ (AGN) in mitogenomes of butterfly species was either TCT, GCT, or ACT, whereas only GCT was detected in all other sequenced mitogenomes of pierid species (*P. rapae, A. crataegi*, *D. hyparete*, and *A. melete*) (Hong, G. Y. et al. 2009, Mao et al. 2010, Park et al. 2012, Shi et al. 2012). The tRNA-like structure ($tRNA^{Leu}$ [UUR]) was detected in the *16S rRNA* gene of *C. pomona*, and a similar observation had been made in *C. vasava* (Hao et al. 2012). Interestingly, the 81 bp insertion of the tRNA-like sequence was made up completely of A and T, without G and C nucleotides. The *L. morsei* and *C. pomona* anticodon sequences of each tRNA isotype were identical to those of all other sequenced butterfly mitogenomes. As in other insects, unmatched base pairs were also detected in the stems of tRNAs. For *L. morsei*, there were 32 unmatched base pairs, consisting of 25 G-U, one A-A, and six U-U mismatches, whereas in *C. pomona*, 21 G-U, two A-A, and six U-U mismatches were identified.

Both of the two pierid mitogenomes harbored a large and a small ribosomal RNA subunit (*16S rRNA* and *12S rRNA*), located between $tRNA^{Leu}$ (CUN) and $tRNA^{Val}$, and between $tRNA^{Val}$ and the A+T-rich region, respectively. The length of the *16S rRNA* and *12S rRNA* genes in *L. morsei* were 1,337 and 764 bp, respectively, with A+T contents of 84.29 and 83.25%, respectively; the *16S rRNA* and *12S rRNA* in *C. pomona* were 1,332 and 779 bp in length, respectively, with A+T contents of 85.21 and 85.11%, respectively (Table 2).

**Intergenic spacers and overlapping sequences**

The mitogenomes of *L. morsei* and *C. pomona* harbored 11 and 15 intergenic spacers, ranging from 1 to 39 bp (94 bp in total) and 1 to 24 bp (87 bp in total), respectively (Table 3). Among these, only three intergenic spacers were longer than 10 bp in both pierid species

(Table 3). The longest intergenic spacers, located between the *tRNA*$^{Gln}$ and *ND2* genes in *L. morsei* and *C. pomona*, were 39 and 24 bp in length, respectively, with A+T contents of 90.35 and 91.67% respectively. This spacer is present in all of the butterfly mitogenomes sequenced to date, whereas absent in all non-lepidopteran insects. Another long spacer harboring the 7 bp ATACTAA motif, located between the *tRNA*$^{Ser}$ (UCN) and *ND1* genes, has been observed commonly in most insect groups including all other butterflies.

In addition, there were 33 overlapping nucleotides scattered over 13 locations in *L. morsei*, and 26 nucleotide overlaps scattered over eight locations in *C. pomona* (Table 2). Among these overlaps in the two pierid species, the longest one was 8 bp in length and located between *tRNA*$^{Trp}$ and *tRNA*$^{Cys}$ with the 7 bp motif AGCCTTA; the second longest one was 7 bp in length and located between *ATP8* and *ATP6* with the 7 bp motif ATGATAA. Both of these motifs have been observed in the mitogenomes of many butterfly species, including all the other pierids sequenced.

**The A+T-rich region**

The A+T-rich regions of *L. morsei* and *C. pomona* were 356 and 313 bp in length, respectively, with A+T contents of 89.60 and 97.13%, respectively. Among the A+T-rich regions of all the butterfly mitogenomes sequenced, that of *C. pomona* was the shortest in length and the highest in A+T content (Table 2). The A+T-rich regions of *L. morsei* and *C. pomona* contained the motif ATAGA, followed by a 19 and 18 bp poly-T stretch, respectively (Fig. 3). Besides, the regions also included microsatellite-like elements, such as $(TA)_9(AT)_3$ in *L. morsei* and $(TA)_9$ in *C. pomona*, which were preceded by the ATTA motif characteristic of lepidopteran mitogenomes. Additionally, a triplicated 23 bp and a duplicated 24 bp repeat element of unknown function were found in *L. morsei* and *C. pomona*, respectively (Fig. 3), and similar repeat elements were detected in other butterflies, such as *A. melete* (Hong, G. Y. et al. 2009), *E. autonoe* (Kim, M. J. et al. 2010), and *Agehana maraho* (Shiraki and Sonan) (Lepidoptera: Papilionidae) (syn. *Papilio maraho*) (Wu et al. 2010).

**Phylogenetic analysis**

Several competing hypotheses exist on the phylogenetic family relationships in butterflies. Ehrlich (1967), Ehrlich and Ehrlich (1967), and Scott (1985) demonstrated the close relationship between the Nymphalidae and Lycaenidae and that between the Pieridae and Papilionidae via numerical taxonomic methods and morphological characters. Son and Kim (2011) and Wahlberg et al. (2005) obtained results consistent with the traditional view of the sister relationship between the Pieridae and the Nymphalidae + Lycaenidae group, with Papilionidae as the basal lineage, in agreement with the earlier hypothesis of Kristensen (1976) and supported by several recent studies (Akaike 1974, Kristensen 1976, de Jong et al. 1996, Ackery et al. 1999, Wahlberg et al. 2005, Liao et al. 2010, Kim, M. I. et al. 2010, Kim, M. J. et al. 2011b, Son et al. 2011). However, Chai et al. (2012) and Zhang et al. (2012), on the basis of mitochondrial genomic data, proposed a close relationship between the Pieridae and Lycaenidae, with the Nymphalidae being the sister group, in agreement with the hypothesis of Robbins (1988). Therefore, controversy exists regarding the relationships among the Nymphalidae, Pieridae, and Lycaenidae.

In this study, we conducted phylogenetic analyses via Bayesian inference and maximum likelihood methods, using concatenated

nucleotide datasets of 13 protein-coding genes (6,582 aligned sites, 910 gaps, and 3,746 excluded positions), resulting in similar tree topologies of the butterfly families Papilionidae, Pieridae, Lycaenidae, and Nymphalidae (Fig. 4 A and B). The presented trees showed two major clusters (Fig. 4 A and B). The first one had the Papilionidae as the basal lineage, and the other one included the rest of the butterfly families. Maximum likelihood and Bayesian inference trees suggest a close relationship between Pieridae and Lycaenidae (Fig. 4 A and B), in agreement with the prevailing phylogeny of butterfly families (Hesperiidae (Papilionidae (Nymphalidae (Pieridae, Lycaenidae)))). Although the close relationship of the Pieridae and Lycaenidae proposed herein is contradictory to the traditional view (Pieridae (Nymphalidae, Lycaenidae)), the result is consistent with those of recent studies (Kim, M. J. et al. 2010, Chai et al. 2012, Hao et al. 2012, Zhang et al. 2012). However, uncertainty does exist regarding the sister relationship of the Pieridae and Lycaenidae as shown in the maximum likelihood tree (Fig. 4A). We also note that this relationship has been derived mainly from the protein-coding genes of the mitochondrial genome of butterflies.

In conclusion, although it may still be immature to suggest that the phylogeny of butterfly families is resolved, we do suggest that the sister relationship of the Pieridae and Lycaenidae is supported at the mitogenomic level in this study. We believe that the problem of butterfly phylogeny should be resolved step by step as gene sequence data and morphologic characters are being accumulated, and hopefully more sophisticated analytic tools will become available.

## Acknowledgements

This work was supported by the National Science Foundation of China (Grant No.41172004), the Chinese Academy of Sciences (Grant No. KZCX22YW2JC104), and a grant from the State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences (Grant No. 104143).

## References

**Ackery, P. R., R. de Jong, and R. I. Vane-Wright. 1999.** The butterflies: Hedyloidea, Hesperioidea, and Papilionoidea, pp. 263–300. *In* N. P. Kirstensen (ed.), *Lepidoptera, moths and butterflies, vol. 1, Evolution, systematics, and biogeography* (Handbook of zoology series). De Gruyter, Berlin, Germany.

**Akaike, H. 1974.** A new look at the statistical model identification. *IEEE Control Systems Society* 19: 716–723.

**Benson, G. 1999.** Tandem Repeats Finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27: 573–580.

**Brunton, C.F.A. 1998.** The evolution of ultraviolet patterns in European *Colias* butterflies (Lepidoptera, Pieridae): a phylogeny using mitochondrial DNA. *Heredity* 80: 611–616.

**Chai, H. N., Y. Z. Du, and B. P. Zhai. 2012.** Characterization of the complete mitochondrial genomes of *Cnaphalocrocis medinalis* and *Chilo suppressalis* (Lepidoptera: Pyralidae). *International Journal of Biological Sciences* 8: 561–579.

**Chapman, T. A. 1895.** Notes on butterfly pupae, with some remarks on the phylogenesis

of the Rhopalocera. *Entomologist's Record and Journal of Variation* 6: 101–107, 125–131, 147–152.

**Chen, M., L. L. Tian, Q. H. Shi, T. W. Cao, and J. S. Hao. 2012.** Complete mitogenome of the lesser purple emperor *Apatura ilia* (Lepidoptera: Nymphalidae: Apaturinae) and comparison with other nymphalid butterflies. *Zoological Research* 33: 191–201.

**Clayton, D. A. 1992.** Transcription and replication of animal mitochondrial DNA. *International Review of Cytology* 141: 217–232.

**Coates, B. S., D. V. Sumerford, R. L. Hellmich, and L. C. Lewis. 2005.** Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis*. *International Journal of Biological Sciences* 1: 13–18.

**Courtney, S. P. 1986.** The ecology of pierid butterflies: dynamics and interactions. *Advances in Ecological Research* 15: 51–131.

**de Jong, R., R. I. Vane-Wright, and P. R. Ackery. 1996.** The higher classification of butterflies (Lepidoptera): problems and prospects. *Insect Systematics and Evolution* 27: 65–101.

**Ehrlich, P. R. 1958.** The comparative morphology, phylogeny and higher classification of the butterflies (Lepidoptera: Papilionoidea). *University of Kansas Science Bulletin* 39: 305–364.

**Ehrlich, P. R., and A. H. Ehrlich. 1967.** The phenetic relationships of the butterflies. I. Adult taxonomy and the non-specificity hypothesis. *Systematic Zoology* 16: 301–317.

**Hall, T. A. 1999.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

**Hao, J. S., C. X. Li, X. Y. Sun, and Q. Yang. 2005.** The phenetic relationships of the butterflies based on mitochondrial *16S rRNA* sequences. *Chinese Science Bulletin* 12: 1205–1211.

**Hao, J. S., Q. Q. Sun, H. B. Zhao, X. Y. Sun, Y. H. Gai, and Q. Yang. 2012.** The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperiidae: Pyrginae) and its phylogenetic implication. *Comparative and Functional Genomics* 2012: 328049. doi:10.1155/2012/328049.

**Hong, G. Y., S. T. Jiang, M. Yu, Y. Yang, F. Li, F. S. Xue, and Z. J. Wei. 2009.** The complete nucleotide sequence of the mitochondrial genome of the cabbage butterfly, *Artogeia melete* (Lepidoptera: Pieridae). *Acta Biochimica et Biophysica Sinica* 41: 446–455.

**Hong, M. Y., E. M. Lee, Y. H. Jo, H. C. Park, S. R. Kim, J. S. Hwang, B. R. Jin, P. D. Kang, K. G. Kim, Y. S. Han, and I. Kim. 2008.** Complete nucleotide sequence and organization of the mitogenome of the silk moth *Caligula boisduvalii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects. *Gene* 413: 49–57.

**Hu, J., J. S. Hao, D. X. Zhang, D. Y. Huang, S. Cameron, and C. D. Zhu. 2010.** 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* 37: 3431–3438.

**Huelsenbeck, J. P., and F. Ronquist. 2001.** MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17: 754-755.

**Ji, L. W., J. S. Hao, Y. Wang, D. Y. Huang, J. L. Zhao, and C. D. Zhu. 2012.** The mitochondrial genome of the dragon swallowtail, *Sericinus montela* Gray (Lepidoptera: Papilionidae), and its phylogenetic implication. *Acta Entomologica Sinica* 55: 91–100.

**Kemp, D. J., R. L. Rutowski, and M. Mendoza. 2005.** Colour pattern evolution in butterflies: a phylogenetic analysis of structural ultraviolet and melanic markings in North American sulphurs. *Evolutionary Ecology Research* 7: 133–141.

**Kim, I., E. M. Lee, K. Y. Seol, E. Y. Yun, Y. B. Lee, J. S. Hwang, and B. R. Jin. 2006.** The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). *Insect Molecular Biology* 15: 217–225.

**Kim, M. I., J. Y. Baek, M. J. Kim, H. C. Jeong, K. G. Kim, C. H. Bae, Y. S. Han, B. R. Jin, and I. Kim. 2009.** 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* 28: 347–363.

**Kim, M. I., X. L. Wan, M. J. Kim, H. C. Jeong, N. H. Ahn, K. G. Kim, Y. S. Han, and I. Kim. 2010.** Phylogenetic relationships of true butterflies (Lepidoptera: Papilionoidea) inferred from *COI*, *16S rRNA* and *EF-1α* sequences. *Molecules and Cells* 30: 409–425.

**Kim, M. J., X. L. Wan, K. G. Kim, J. S. Hwang, and I. Kim. 2010.** Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). *African Journal of Biotechnology* 9: 735–754.

**Kim, M. J., H. C. Jeong, S. R. Kim, and I. Kim. 2011a.** Complete mitochondrial genome of the nerippe fritillary butterfly, *Argynnis nerippe* (Lepidoptera: Nymphalidae). *Mitochondrial DNA* 22: 86–88.

**Kim, M. J., A. R. Kang, H. C. Jeong, K. G. Kim, and I. Kim. 2011b.** Reconstructing intraordinal relationships in Lepidoptera using mitochondrial genome data with the description of two newly sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera: Lycaenidae). *Molecular Phylogenetics and Evolution* 61: 436–445.

**Kristensen, N. P. 1976.** Remarks on the family level phylogeny of butterflies (Insecta, Lepidoptera, Rhopalocera). *Zeitschrift für Zoologische Systematik und Evolutionsforschung* 14: 25-33.

**Kumar, S., K. Tamura, and M. Nei. 2004.** MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* 5: 150-163.

**Lanave, C., G. Preparata, C. Saccone, and G. Serio. 1984.** A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution* 20: 86–93.

**Liao F., L. Wang, S. Wu, Y. P. Li, L. Zhao, G. M. Huang, C. J. Niu, Y. Q. Liu, and M. G. Li. 2010.** The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *International Journal of Biological Sciences* 6: 172–186.

**Lowe, T. M., and S. R. Eddy. 1997.** tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* 25: 955–964.

**Mao, Z. H., J. S. Hao, G. P. Zhu, J. Hu, M. M. Si, and C. D. Zhu. 2010.** Sequencing and analysis of the complete mitochondrial genome of *Pieris rapae* Linnaeus (Lepidoptera: Pieridae). *Acta Entomologica Sinica* 5: 1295–1304.

**Ojala, D., J. Montoya, and G. Attardi. 1981.** tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290: 470–474.

**Park, J. S., Y. Cho, M. J. Kim, S. H. Nam, and I Kim. 2012.** Description of complete mitochondrial genome of the black-veined white, *Aporia crataegi* (Lepidoptera: Papilionoidea), and comparison to papilionoid species. *Journal of Asia-Pacific Entomology* 15: 331–341.

**Posada, D., and K. A. Crandal. 1998.** Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

**Qin, X. M., Q. X. Guan, D. L. Zeng, F. Qin, and H. M. Li. 2012.** Complete mitochondrial genome of *Kallima inachus* (Lepidoptera: Nymphalidae: Nymphalinae): comparison of *K. inachus* and *Argynnis hyperbius*. *Mitochondrial DNA* 23(4):318–320.

**Rienks, J. H. 1985.** Phenotypic response to photoperiod and temperature in a tropical pierid butterfly. *Australian Journal of Zoology* 33: 837–847.

**Robbins, R. K. 1988.** Comparative morphology of the butterfly foreleg coxa and trochanter (Lepidoptera) and its systematic implications. *Proceedings of the Entomological Society of Washington* 90: 133–154.

**Scott, J. A. 1985.** The phylogeny of butterflies (Papilionoidea and Hesperioidea). *Journal of Research on the Lepidoptera* 23: 241–281.

**Shi Q. H., J. Xia, X. Y. Sun, J. S. Hao, and Q. Yang. 2012.** Complete mitogenome of the Painted Jezebel, *Delis hyparete* Linnaeus (Lepidoptera: Pieridae) and its comparison with other butterfly species. *Zoological Research* 33: 111–120.

**Simon, C., F. Frati, A. Beckenbach, B. Cresp, H. Liu, and P. Flook. 1994.** Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* 87: 651–701.

**Singh, V. K., A. K. Mangalam, S. Dwivedi, and S. Naik. 1998.** Primer Premier: program for design of degenerate primers from a protein sequence. *Bio Techniques* 24: 318–319.

**Son, Y., and Y. Kim. 2011.** The complete mitochondrial genome of *Grapholita molesta* (Lepidoptera: Tortricidae). *Annals of the Entomological Society of America* 104: 788–799.

**Stavenga, D. G., S. Stowe, K. Siebke, J. Zeil, and K. Arikawa. 2004.** Butterfly wing colours: scale beads make white pierid wings brighter. *Proceedings of the Royal Society of London B* 271: 1577–1584.

**Swofford, D. L. 2002.** *PAUP: Phylogenetic analysis using parsimony (and other methods) version 4.10*. Sinauer Associates, Sunderland, MA.

**Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997.** The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882.

**Tian, L. L., X. Y. Sun, M. Chen, Y. Gai, J. S. Hao, and Q. Yang. 2012.** Complete mitochondrial genome of the five-dot sergeant *Parathyma sulpitia* (Nymphalidae: Limenitidinae) and its phylogenetic implications. *Zoological Research* 33: 133–143.

**Wahlberg, N., M. F. Braby, A. V. Z. Brower, R. de Jong, M. M. Lee, S. Nylin, N. Pierce, F. A. Sperling, R. Vila, A. D. Warren. 2005.** Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. *Proceedings of the Royal Society B: Biological Sciences* 272: 1577-1586.

**Wang, X. C., X. Y. Sun, Q. Q. Sun, D. X. Zhang, J. Hu, Q. Yang, and J. S. Hao. 2011.** The complete mitochondrial genome of the laced fritillary *Argyreus hyperbius* (Lepidoptera: Nymphalidae). *Zoological Research* 32: 465–475.

**Watt, W. B., K. Donohue, and P. A. Carter. 1996.** Adaptation at specific loci. VI. Divergence vs. parallelism of polymorphic allozymes in molecular function and fitness component effects among *Colias* species (Lepidoptera, Pieridae). *Molecular Biology and Evolution* 13: 699–709.

**Weller, S. J., D. P. Pashley, and J. A. Martin. 1996.** Reassessment of butterfly family relationships using independent genes and morphology. *Annals of the Entomological Society of America* 89: 184–192.

**Wu, L. W., D. C. Lees, S. H. Yen, C. C. Lu, and Y. F. Hsu. 2010.** The complete mitochondrial genome of the near-threatened swallowtail, *Agehana maraho* (Lepidoptera: Papilionidae): evaluating sequence variability and suitable markers for conservation genetic studies. *Entomological News* 121: 267–280.

**Xia, J., J. Hu, G. P. Zhu, C. D. Zhu, and J. S. Hao. 2011.** Complete mitochondrial DNA sequence of *Calinaga davidis*. *Acta Entomologica Sinica* 54: 555–565.

**Yukuhiro, K., H. Sezutsu, M. Itoh, K. Shimizu, and Y. Banno. 2002.** Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkmoth, *Bombyx mandarina* , and its close relative, the domesticated silkmoth, *Bombyx mori*. *Molecular Biology and Evolution* 19: 1385–1389.

**Zhang, M., X. P. Nie, T. W. Cao, J. P. Wang, T. Li, X. N. Zhang, Y. P. Guo, E. B. Ma, and Y. Zhong. 2012.** The complete mitochondrial genome of the butterfly *Apatura metis* (Lepidoptera: Nymphalidae). *Molecular Biology Reports* 39: 6529–6536.

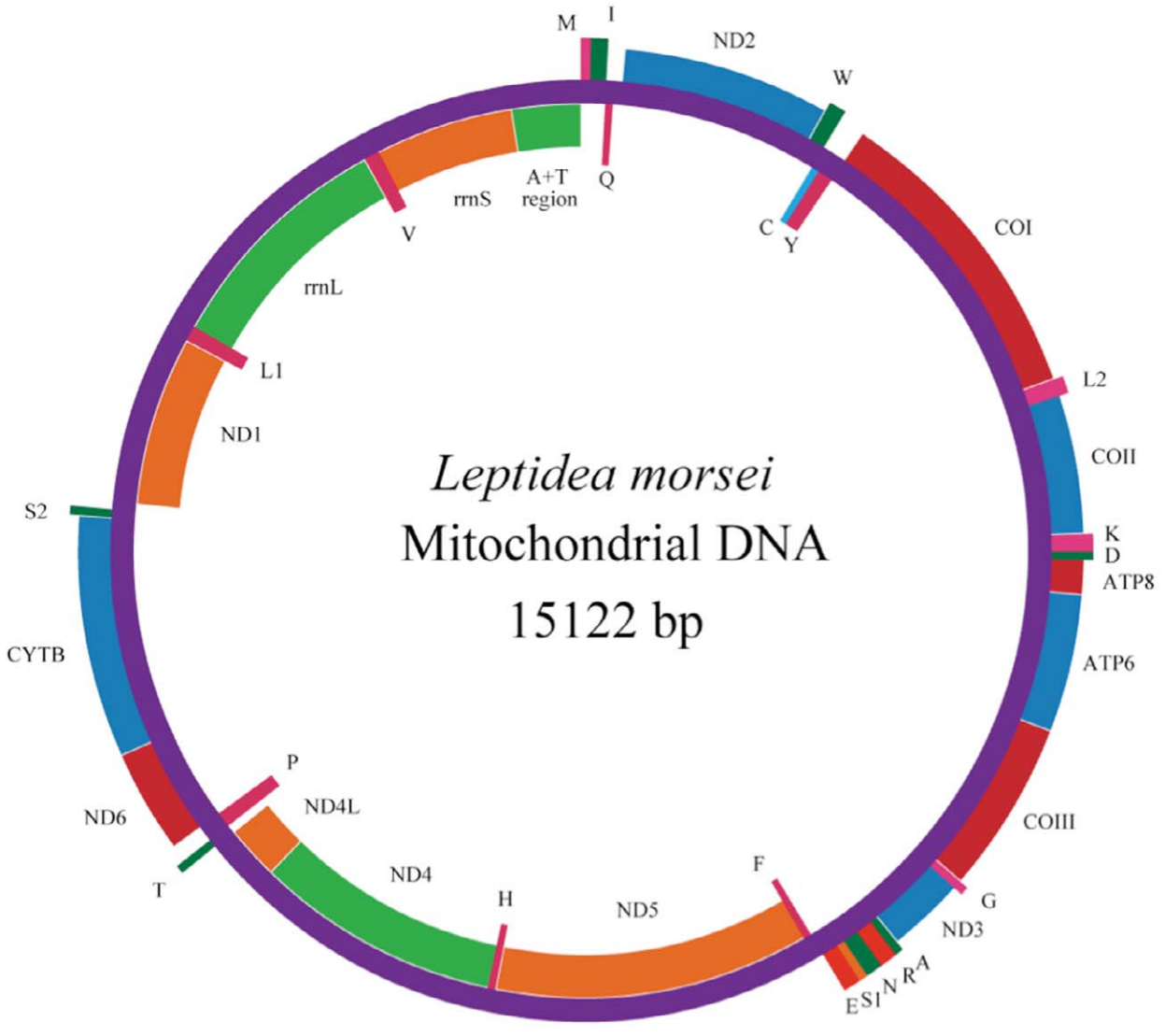


**Figure 1.** Map of the circular mitochondrial genome of *L. morsei.* Genes encoded in the H-strand (clockwise orientation) are colored in red or blue. Genes encoded in the L-strand (anticlockwise orientation) are colored in orange or green. Abbreviations for the genes: *COI-III* for cytochrome oxidase subunits, *CYTB* for cytochrome b, and *ND1-6* for NADH dehydrogenase components. tRNAs are denoted as one-letter symbols according to the IUPAC-IUB single-letter amino acid codes. High quality figures are available online.


**Figure 2.** Alignment of the initiation codons of the *COI* genes of lepidopterans, including those of *L. morsei* and *C. pomona*. The first four or five codons for *COI* and their amino acids are shown on the right-hand side of the figure. Underlined nucleotides indicate the adjacent partial sequence of *tRNA*$^{Tyr}$. Arrows indicate the transcriptional direction. Boxed nucleotides indicate the currently proposed translation initiators for the *COI* gene of lepidopteran insects. The start codon for *L. morsei* and *C. pomona* is designated as CGA. High quality figures are available online.

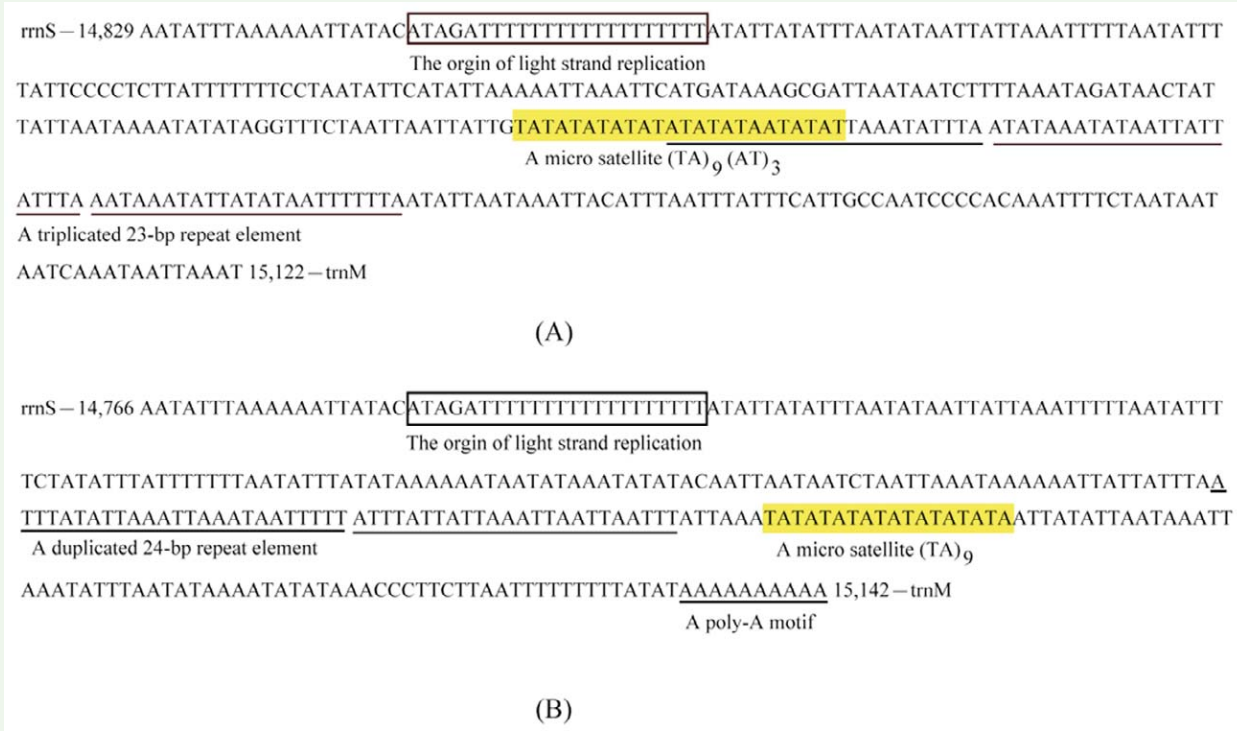


**Figure 3.** Structures of the A+T-rich regions in the mitogenomes of *L. morsei* (A) and *C. pomona* (B). High quality figures are available online.

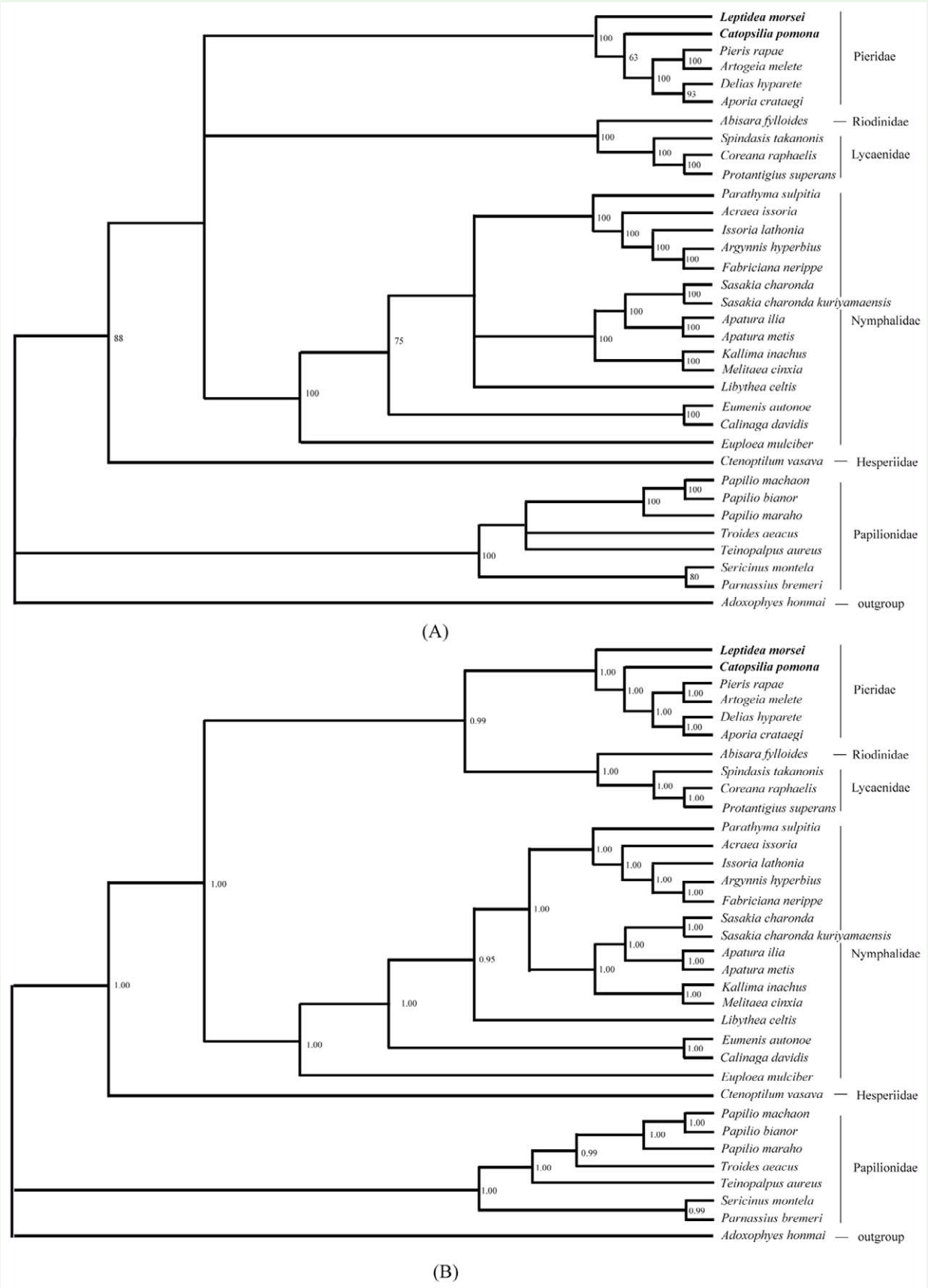


**Figure 4.** Phylogenetic trees of the butterflies in this study based on the nucleotide sequences of 13 protein-coding genes. (A) Maximum likelihood tree. (B) Bayesian inference tree. Numbers at each node indicate bootstrap percentage of maximum likelihood analysis and posterior probability of Bayesian inference analysis. High quality figures are available online.

**Table 1.** Mitogenomes of the 33 butterfly species used in this study.

| Family | Subfamily | Species | GenBank Acc. No. | References |
|---|---|---|---|---|
| Nymphalidae | Argynninae | *Argyreus hyperbius* | JF439070 | Wang *et al.*, 2011 |
| Nymphalidae | Apaturinae | *Apatura ilia* | JF437925 | Chen *et al.*, 2012 |
| Nymphalidae | Apaturinae | *Apatura metis* | JF801742 | Zhang et al., 2012 |
| Nymphalidae | Apaturinae | *Sasakia charonda kuriyamaensis* | AP011825 | Hakozaki *et al.*, unpublished |
| Nymphalidae | Apaturinae | *Sasakia charonda* | AP011824 | Wang *et al.*, 2012 |
| Nymphalidae | Calinaginae | *Calinaga davidis* | HQ658143 | Xia *et al.*, 2011 |
| Nymphalidae | Danainae | *Euploea mulciber* | HQ378507 | Sun *et al.*, unpublished |
| Nymphalidae | Heliconiinae | *Acraea issoria* | GQ376195 | Hu *et al.*, 2010 |
| Nymphalidae | Heliconiinae | *Issoria lathonia* | HM243590 | Qin *et al.*, unpublished |
| Nymphalidae | Heliconiinae | *Fabriciana nerippe* | JF504707 | Kim M.J. et al., 2011a |
| Nymphalidae | Limenitinae | *Parathyma sulpitia* | JQ347260 | Tian *et al.*, 2012 |
| Nymphalidae | Libytheinae | *Libythea celtis* | HQ378508 | Sun *et al.*, unpublished |
| Nymphalidae | Nymphalinae | *Melitaea cinxia* | GQ398377 | Xu *et al.*, unpublished |
| Nymphalidae | Nymphalinae | *Kallima inachus* | JN857943 | Qin *et al.*, 2012 |
| Nymphalidae | Satyrinae | *Eumenis autonoe* | GQ868707 | Kim M.J. *et al.*, 2010 |
| Papilionidae | Papilioninae | *Teinopalpus aureus* | HM563681 | Qin *et al.*, 2012 |
| Papilionidae | Papilioninae | *Papilio maraho* | GQ868707 | Wu *et al.*, 2010 |
| Papilionidae | Papilioninae | *Papilio machaon* | HM243594 | Xu *et al.*, unpublished |
| Papilionidae | Papilioninae | *Papilio bianor* | NC018040 | Xu *et al.*, unpublished |
| Papilionidae | Papilioninae | *Troides aeacus* | EU625344 | Jiang *et al.*, unpublished |
| Papilionidae | Parnassiinae | *Parnassius bremeri* | FJ871125 | Kim *et al.*, 2009 |
| Papilionidae | Parnassiinae | *Sericinus montela* | HQ259122 | Ji *et al.*, 2012 |
| Lycaenidae | Aphnaeinae | *Spindasis takanonis* | HQ184266 | Kim M.J. *et al.*, 2011b |
| Lycaenidae | Theclinae | *Coreana raphaelis* | DQ102703 | Kim I. *et al.*, 2006 |
| Lycaenidae | Theclinae | *Protantigius superans* | HQ184265 | Kim M.J. *et al.*, 2011b |
| Riodinidae | Nemeobiinae | *Abisara fylloides* | HQ259069 | Zhao *et al.*, unpublished |
| Pieridae | Coliadinae | *Catopsilia Pomona* | JX274649 | This study |
| Pieridae | Dismorphiinae | *Leptidea morsei* | JX274648 | This study |
| Pieridae | Pierinae | *Artogeia melete* | EU597124 | Hong G.Y. *et al.*, 2009 |
| Pieridae | Pierinae | *Pieris rapae* | HM156697 | Mao *et al.*, 2010 |
| Pieridae | Pierinae | *Delias hyparete* | JX094279 | Shi *et al.*, 2012 |
| Pieridae | Pierinae | *Aporia crataegi* | JN796473 | Kim *et al.*, 2012 |
| Hesperiidae | Pyrginae | *Ctenoptilum vasava* | JF713818 | Hao *et al.*, 2012 |

**Table 2.** Characteristics of mitogenomes of the 33 butterfly species available.

| Taxon | Size (bp) | A+T content (%) | No. codons[a] | PCG[b] A+T (%) | 16S rRNA | | 12S rRNA | | A+T-rich region | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Size (bp) | A+T (%) | Size (bp) | A+T (%) | Size (bp) | A+T (%) |
| *Argyreus hyperbius* | 15,156 | 80.8 | 3,718 | 79.4 | 1,330 | 84.5 | 778 | 85.2 | 349 | 95.4 |
| *Apatura ilia* | 15,242 | 80.5 | 3,711 | 78.9 | 1,333 | 86.0 | 776 | 84.9 | 403 | 92.5 |
| *Apatura metis* | 15,236 | 80.5 | 3,707 | 78.9 | 1,333 | 84.5 | 779 | 84.8 | 394 | 92.9 |
| *Sasakia charonda kuriyamaensis* | 15,222 | 79.9 | 3,586 | 78.2 | 1,311 | 84.2 | 775 | 85.0 | 380 | 91.8 |
| *Sasakia charonda* | 15,244 | 79.9 | 3,695 | 78.2 | 1,323 | 84.4 | 775 | 85.0 | 380 | 91.8 |
| *Calinaga davidis* | 15,267 | 80.4 | 3,737 | 78.8 | 1,337 | 83.8 | 773 | 85.9 | 389 | 92.0 |
| *Euploea mulciber* | 15,166 | 81.5 | 3,713 | 80.2 | 1,314 | 84.6 | 776 | 85.3 | 399 | 93.5 |
| *Acraea issoria* | 15,245 | 79.7 | 3,717 | 78.1 | 1,331 | 83.9 | 788 | 83.7 | 430 | 96.0 |
| *Issoria lathonia* | 15,172 | 81.2 | 3,718 | 79.9 | 1,319 | 84.4 | 771 | 85.1 | 361 | 96.1 |
| *Fabriciana nerippe* | 15,140 | 80.9 | 3,719 | 79.6 | 1,321 | 84.4 | 773 | 84.9 | 329 | 95.4 |
| *Parathyma sulpitia* | 15,282 | 81.9 | 3,729 | 80.6 | 1,319 | 84.7 | 779 | 85.7 | 349 | 94.6 |
| *Libythea celtis* | 15,164 | 81.2 | 3,732 | 80.1 | 1,335 | 84.7 | 774 | 85.4 | 328 | 96.3 |
| *Melitaea cinxia* | 15,170 | 80.0 | 3,718 | 78.6 | 1,336 | 84.6 | 772 | 84.6 | 338 | 92.9 |
| *Kallima inachus* | 15,183 | 80.3 | 3,704 | 79.2 | 1,335 | 82.8 | 774 | 85.1 | 376 | 92.0 |
| *Eumenis autonoe* | 15,489 | 79.1 | 3,728 | 76.8 | 1,335 | 83.7 | 775 | 85.3 | 678 | 94.5 |
| *Teinopalpus aureus* | 15,242 | 79.8 | 3,720 | 78.3 | 1,320 | 82.4 | 781 | 85.6 | 395 | 93.1 |
| *Papilio maraho* | 16,094 | 80.5 | 3,717 | 78.1 | 1,333 | 83.7 | 779 | 85.5 | 1,270 | 93.7 |
| *Papilio machaon* | 15,185 | 80.3 | 3,720 | 79.0 | 1,319 | 83.5 | 772 | 84.2 | 362 | 92.5 |
| *Papilio bianor* | 15,340 | 80.6 | 3,719 | 79.0 | 1,324 | 83.4 | 773 | 85.4 | 498 | 94.0 |
| *Troides aeacus* | 15,263 | 80.2 | 3,636 | 79.1 | 1,234 | 83.2 | 784 | 84.9 | 418 | 89.7 |
| *Parnassius bremeri* | 15,389 | 81.3 | 3,734 | 80.2 | 1,344 | 83.8 | 773 | 85.1 | 504 | 93.6 |
| *Sericinus montela* | 15,242 | 80.9 | 3,691 | 79.8 | 1,338 | 83.6 | 760 | 84.6 | 408 | 94.1 |
| *Spindasis takanonis* | 15,349 | 82.4 | 3,719 | 81.0 | 1,333 | 85.6 | 777 | 84.7 | 371 | 94.6 |
| *Coreana raphaelis* | 15,314 | 82.7 | 3,708 | 81.5 | 1,330 | 85.3 | 777 | 85.8 | 375 | 94.1 |
| *Protantigius superans* | 15,248 | 81.7 | 3,712 | 80.3 | 1,331 | 85.0 | 739 | 85.5 | 361 | 93.6 |
| *Abisara fylloides* | 15,301 | 81.2 | 3,740 | 79.8 | 1,334 | 85.4 | 771 | 85.6 | 423 | 91.0 |
| ***Catopsilia pomona*** | **15,142** | **81.3** | **3,724** | **80.0** | **1,332** | **85.2** | **779** | **85.1** | **313** | **97.1** |
| ***Leptidea morsei*** | **15,122** | **80.2** | **3,713** | **79.2** | **1,337** | **84.3** | **764** | **83.2** | **356** | **89.3** |
| *Artogeia melete* | 15,140 | 79.8 | 3,715 | 78.4 | 1,319 | 83.4 | 777 | 86.9 | 351 | 88.0 |
| *Pieris rapae* | 15,157 | 79.7 | 3,721 | 78.2 | 1,330 | 84.0 | 764 | 85.0 | 393 | 91.6 |
| *Delias hyparete* | 15,186 | 79.8 | 3,703 | 78.4 | 1,336 | 83.7 | 774 | 85.1 | 377 | 92.0 |
| *Aporia crataegi* | 15,140 | 81.3 | 3,708 | 79.9 | 1,326 | 85.4 | 779 | 85.5 | 354 | 95.2 |
| *Ctenoptilum vasava* | 15,468 | 80.5 | 3,698 | 78.9 | 1,343 | 84.1 | 774 | 86.4 | 429 | 88.1 |

[a] Termination codons were excluded from total codon count.
[b] Protein coding genes.

**Table 3.** Summarized characteristics of the mitogenomes of *L. morsei* (*Lm*) and *C. pomona* (*Cp*).

| Gene | Direction | Nucleotide No. | | Size (bp) | | IGNc | | Start codon/Stop codon | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Lm* | *Cp* | *Lm* | *Cp* | *Lm* | *Cp* | *Lm* | *Cp* |
| $tRNA^{Met}$ | F | 1-64 | 1-67 | 64 | 67 | 0 | 3 | | |
| $tRNA^{Ile}$ | F | 65-128 | 71-134 | 64 | 64 | -3 | 1 | | |
| $tRNA^{Gln}$ | R | 126-194 | 136-204 | 69 | 69 | 39 | 24 | | |
| ND2 | F | 234-1250 | 229-1260 | 1017 | 1032 | -2 | 6 | ATT/TAA | ATT/TAA |
| $tRNA^{Trp}$ | F | 1249-1314 | 1267-1332 | 66 | 66 | -8 | -8 | | |
| $tRNA^{Cys}$ | R | 1307-1368 | 1325-1386 | 62 | 62 | 4 | 5 | | |
| $tRNA^{Tyr}$ | R | 1373-1436 | 1392-1456 | 64 | 65 | 2 | 2 | | |
| COI | F | 1439-2969 | 1459-2989 | 1531 | 1531 | 0 | 0 | CGA/T-tRNA | CGA/T-tRNA |
| $tRNA^{Leu}$(UUR) | F | 2970-3036 | 2990-3055 | 67 | 66 | 0 | 0 | | |
| COII | F | 3037-3712 | 3056-3731 | 676 | 676 | 0 | 0 | ATG/T-tRNA | ATG/T-tRNA |
| $tRNA^{Lys}$ | F | 3713-3783 | 3732-3801 | 71 | 70 | -1 | -1 | | |
| $tRNA^{Asp}$ | F | 3783-3848 | 3801-3868 | 66 | 68 | 0 | 0 | | |
| ATP8 | F | 3849-4007 | 3869-4030 | 159 | 162 | -7 | -7 | ATC/TAA | ATT/TAA |
| ATP6 | F | 4001-4684 | 4024-4701 | 684 | 678 | -1 | -1 | ATG/TAA | ATG/TAA |
| COIII | F | 4684-5475 | 4701-5492 | 792 | 792 | 2 | 3 | ATG/TAA | ATG/TAA |
| $tRNA^{Gly}$ | F | 5478-5542 | 5496-5561 | 65 | 66 | 0 | 0 | | |
| ND3 | F | 5543-5896 | 5562-5915 | 354 | 354 | 3 | -2 | ATA/TAA | ATT/ TAG |
| $tRNA^{Ala}$ | F | 5900-5965 | 5914-5982 | 66 | 69 | -1 | -1 | | |
| $tRNA^{Arg}$ | F | 5965-6027 | 5982-6048 | 65 | 67 | 11 | 16 | | |
| $tRNA^{Asn}$ | F | 6039-6103 | 6065-6130 | 65 | 66 | -1 | 1 | | |
| $tRNA^{Ser}$(AGN) | F | 6103-6163 | 6132-6191 | 61 | 60 | -1 | 1 | | |
| $tRNA^{Glu}$ | F | 6163-6228 | 6193-6259 | 66 | 67 | -2 | 0 | | |
| $tRNA^{Phe}$ | R | 6227-6290 | 6260-6326 | 64 | 67 | 0 | 0 | | |
| ND5 | R | 6291-8028 | 6327-8064 | 1738 | 1738 | -3 | 0 | ATT/T-tRNA | ATT/ T-tRNA |
| $tRNA^{His}$ | R | 8026-8089 | 8065-8130 | 64 | 66 | 0 | 0 | | |
| ND4 | R | 8090-9413 | 8131-9466 | 1325 | 1336 | 12 | -4 | ATA/T-tRNA | ATA/T-tRNA |
| ND4L | R | 9426-9704 | 9463-9744 | 279 | 282 | 2 | 3 | ATG/TAA | ATG/TAA |
| $tRNA^{Thr}$ | F | 9707-9769 | 9748-9811 | 63 | 64 | 0 | 0 | | |
| $tRNA^{Pro}$ | R | 9770-9835 | 9812-9877 | 66 | 66 | 2 | 2 | | |
| ND6 | F | 9838-10365 | 9880-10407 | 528 | 528 | -1 | 3 | ATT/TAA | ATC/TAA |
| CytB | F | 10365-11513 | 10411-11565 | 1149 | 1155 | -2 | -2 | ATG/TAA | ATG/TAA |
| $tRNA^{Ser}$(UCN) | F | 11512-11578 | 11564-11629 | 67 | 66 | 16 | 16 | | |
| ND1 | R | 11595-12533 | 11646-12584 | 939 | 939 | 1 | 1 | ATG/TAA | ATG/TAA |
| $tRNA^{Leu}$(CUN) | R | 12535-12601 | 12586-12652 | 67 | 67 | 0 | 0 | | |
| lrRNA | R | 12602-13938 | 12653-13984 | 1337 | 1332 | 0 | 0 | | |
| $tRNA^{Val}$ | R | 13939-14002 | 13985-14050 | 64 | 66 | 0 | 0 | | |
| srRNA | R | 14003-14766 | 14051-14829 | 764 | 779 | 0 | 0 | | |
| A+T-rich region | R | 14767-15122 | 14830-15142 | 356 | 313 | 356 | 313 | | |

IGNc: intergenic nucleotide length,the positive number indicates interval nucleotides (base pairs) between genes, while the negative number indicates the overlapped nucleotides (base pairs) between genes.

**Table 4.** Size proportion of coding genes, intergenic spacers, and the A+T-rich region to the whole genome of the butterflies in this study.

| Taxon | Total | Gene | | Intergenic spacer | | A+T-rich region | |
|---|---|---|---|---|---|---|---|
| | Size (bp) | Size (bp) | % | Size (bp) | % | Size (bp) | % |
| *Argynnis hyperbius* | 15,156 | 14,707 | 97.0 | 97 | 0.6 | 349 | 2.3 |
| *Apatura ilia* | 15,242 | 14,718 | 96.6 | 155 | 1.0 | 403 | 2.6 |
| *Apatura metis* | 15,236 | 14,676 | 96.3 | 161 | 1.1 | 394 | 2.6 |
| *Sasakia charonda kuriyamaensis* | 15,222 | 14,340 | 93.6 | 608 | 4.0 | 380 | 2.5 |
| *Sasakia charonda* | 15,244 | 14,643 | 96.1 | 321 | 2.1 | 380 | 2.5 |
| *Calinaga davidis* | 15,267 | 14,775 | 96.8 | 130 | 0.9 | 389 | 2.5 |
| *Euploea mulciber* | 15,166 | 14,682 | 96.6 | 117 | 0.8 | 399 | 2.6 |
| *Acraea issoria* | 15,245 | 14,746 | 96.7 | 88 | 0.6 | 430 | 2.8 |
| *Issoria lathonia* | 15,172 | 14,733 | 97.0 | 107 | 0.7 | 361 | 2.4 |
| *Fabriciana nerippe* | 15,140 | 14,748 | 97.2 | 98 | 0.6 | 329 | 2.2 |
| *Parathyma sulpitia* | 15,268 | 14,919 | 96.4 | 213 | 1.4 | 349 | 2.3 |
| *Libythea celtis* | 15,164 | 14,766 | 97.2 | 97 | 0.6 | 328 | 2.2 |
| *Melitaea cinxia* | 15,170 | 14,751 | 97.1 | 96 | 0.6 | 338 | 2.2 |
| *Kallima inachus* | 15,183 | 14,756 | 97.0 | 88 | 0.6 | 376 | 2.5 |
| *Euploea mulciber* | 15,166 | 14,682 | 96.6 | 117 | 0.8 | 399 | 2.6 |
| *Teinopalpus aureus* | 15,242 | 14,713 | 96.5 | 107 | 0.7 | 395 | 2.6 |
| *Papilio maraho* | 16,094 | 14,705 | 91.4 | 111 | 0.7 | 1270 | 7.9 |
| *Papilio machaon* | 15,185 | 14,733 | 97.0 | 125 | 0.8 | 362 | 2.4 |
| *Papilio bianor* | 15,340 | 14,742 | 96.1 | 127 | 0.8 | 498 | 3.2 |
| *Troides aeacus* | 15,263 | 14,433 | 94.4 | 446 | 2.9 | 418 | 2.7 |
| *Parnassius bremeri* | 15,389 | 14,745 | 95.8 | 138 | 0.9 | 504 | 3.3 |
| *Sericinus montela* | 15,241 | 14,667 | 96.2 | 166 | 1.1 | 408 | 2..7 |
| *Spindasis takanonis* | 15,349 | 14,743 | 96.1 | 219 | 1.4 | 371 | 2.4 |
| *Coreana raphaelis* | 15,314 | 14,689 | 95.9 | 177 | 1.2 | 375 | 2.4 |
| *Protantigius superans* | 15,248 | 14,663 | 96.2 | 217 | 1.4 | 361 | 2.4 |
| *Abisara fylloides* | 15,301 | 14,820 | 96.5 | 118 | 0.8 | 423 | 2.8 |
| ***Catopsilia pomona*** | **15,142** | **14,770** | **97.4** | **87** | **0.6** | **313** | **2.1** |
| ***Leptidea morsei*** | **15,122** | **14,707** | **97.0** | **94** | **0.6** | **356** | **2.3** |
| *Artogeia melete* | 15,140 | 14,677 | 96.9 | 117 | 0.8 | 351 | 2.3 |
| *Pieris rapae* | 15,157 | 14,682 | 96.9 | 112 | 0.7 | 393 | 2.6 |
| *Delias hyparete* | 15,186 | 14,699 | 96.8 | 153 | 1.0 | 377 | 2.5 |
| *Aporia crataegi* | 15,140 | 14,677 | 96.9 | 111 | 0.7 | 354 | 2.3 |
| *Ctenoptilum vasava* | 15,468 | 14,887 | 96.1 | 179 | 1.2 | 429 | 2.8 |

**Table 5.** Nucleotide compositions in *L. morsei* (*Lm*) and *C. pomona* (*Cp*).

| Feature | A (%) | | C (%) | | G (%) | | T (%) | | A+T (%) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Lm* | *Cp* | *Lm* | *Cp* | *Lm* | *Cp* | *Lm* | *Cp* | *Lm* | *Cp* |
| Whole genome | 38.51 | 39.45 | 11.83 | 11.07 | 7.93 | 7.64 | 41.73 | 41.84 | 80.24 | 81.29 |
| Protein-coding genes[a] | 34.09 | 34.35 | 10.39 | 9.56 | 10.46 | 10.36 | 45.07 | 45.67 | 79.16 | 80.02 |
| 1st codon positions | 37.95 | 37.65 | 10.48 | 9.56 | 15.35 | 15.84 | 36.22 | 36.95 | 74.17 | 74.60 |
| 2st codon positions | 22.25 | 22.13 | 15.84 | 16.03 | 13.14 | 13.13 | 48.77 | 48.71 | 71.02 | 70.84 |
| 3st codon positions | 42.07 | 43.29 | 4.85 | 3.22 | 2.88 | 2.09 | 50.20 | 51.40 | 92.27 | 94.69 |
| tRNA genes | 39.69 | 41.24 | 10.93 | 7.83 | 8.36 | 10.86 | 41.02 | 40.07 | 80.71 | 81.31 |
| rrnL genes | 45.62 | 43.84 | 5.46 | 5.18 | 10.25 | 9.61 | 38.67 | 41.37 | 84.29 | 85.21 |
| rrnS genes | 44.50 | 46.34 | 5.63 | 4.88 | 11.13 | 10.01 | 38.74 | 38.77 | 83.24 | 85.11 |
| A+T-rich region | 41.29 | 46.65 | 7.58 | 2.56 | 2.81 | 3.19 | 48.31 | 50.48 | 89.60 | 97.13 |

[a] Stop codons excluded.

**Table 6.** The 13 protein-coding gene initiation and termination codons in the mitogenomes of the 33 butterfly species in this study.

| Species | Predicted initiation and termination codons | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ND2 | COI | COII | ATP8 | ATP6 | COIII | ND3 | ND5 | ND4 | ND4L | ND6 | CytB | ND1 |
| *Argyreus hyperbius* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATA/TAA | ATG/TAA | ATG/TAA | ATA/TAA | ATG/TAA | ATA/TAA |
| *Apatura ilia* | ATT/TAA | CGA/TAA | ATG/T | ATT/TAA | ATG/TAG | ATG/TAA | ATA/T | ATT/T | ATG/T | ATG/TAG | ATA/TAA | ATG/TAA | ATG/TAA |
| *Apatura metis* | ATT/TAA | CGA/TAA | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATT/T | ATG/T | ATG/TAA | ATA/TAA | ATG/TAA | ATG/TAA |
| *Sasakia charonda kuriyamaensis* | ATA/TAA | ATA/TAA | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATA/TAG | ATT/TAA | ATA/TAA | ATA/TAA | ATA/TAA | ATG/TAG | ATG/TAA |
| *Sasakia charonda* | ATA/TAA | TTG/TAA | GTG/T | ATC/TAA | ATG/TA | ATG/TAA | ATT/T | ATT/TAA | ATG/TAA | ATG/TA | ATA/TAA | ATG/TAA | ATG/TAA |
| *Calinaga davidis* | ATT/TAA | CGA/T | ATG/TAA | ATC/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/TAA | ATG/TA | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TAA |
| *Euploea mulciber* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATT/T | ATG/T | ATG/TAA | ATA/TAA | ATG/TAA | ATG/TAA |
| *Acraea issoria* | ATT/TAA | TTG/T | ATG/TAA | ATG/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATA/TAA | ATG/TAA | ATG/TAA | ATA/TAA | ATG/TAA | ATG/TAA |
| *Issoria lathonia* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/T | ATT/T | ATT/TAA | ATG/TA | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TA |
| *Fabriciana nerippe* | ATT/T | TTG/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/TA | ATG/TA | ATG/TAA | ATT/TAA | ATG/TAA | ATA/TAA |
| *Parathyma sulpitia* | ATT/TAA | CGA/TAA | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/TAA | ATG/T | ATG/TAA | ATA/TAA | ATG/TAA | ATG/TAA |
| *Libythea celtis* | ATC/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/T | ATG/T | ATG/TAG | ATT/TAA | ATG/TAA | ATG/TAA |
| *Melitaea cinxia* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/TAA | ATG/TA | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TA |
| *Kallima inachus* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATA/TAG | ATG/TAA | ATG/TA | ATTG/TAA | ATT/TAA | ATG/TAA | ATG/TAG |
| *Eumenis autonoe* | ATT/TAA | CGA/TAA | ATG/T | ATC/TAA | ATG/TAA | ATG/TAA | ATC/TAA | ATC/T | ATG/T | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TAA |
| *Teinopalpus aureus* | ATT/T | TAGCGA/T | ATG/T | ATT/TAA | ATG/TAA | ATA/T | ATT/T | ATA/TAA | ATG/T | ATG/TAG | ATT/TAA | ATA/T | ATG/TAG |
| *Papilio maraho* | ATT/T | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATA/TAA | ATT/TAA | ATG/T | ATG/TAA | ATC/TAA | ATA/T | ATG/TAG |
| *Papilio machaon* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATT/TAA | ATG/TAA | ATG/TAA | ATA/TAA | ATG/TAA | ATG/TAG |
| *Papilio bianor* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATC/TAG | ATT/TAA | ATG/TAA | ATG/TAA | ATG/TAA | ATG/TAA | ATG/TAG |
| *Troides aeacus* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TGG | ATA/TAG | ATT/TAA | ATC/TAT | ATG/TAA | ATT/TAA | ATA/TAA | ATG/TAA |
| *Parnassius bremeri* | ATT/TAA | CGA/TAA | ATG/T | ATA/TAA | ATG/TAA | ATG/TAA | ATA/TAA | ATT/TAA | ATG/TAA | ATG/TAA | ATC/TAA | ATA/TAA | ATG/TAA |
| *Sericinus montela* | ATT/TAA | CGA/T | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATT/TAA | ATG/TAA | ATA/TAG | ATT/TAA | ATA/TAA | ATA/TAA |
| *Spindasis takanonis* | ATT/TAA | CGA/TAA | ATG/T | ATC/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/TAA | ATG/TAG | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TA |
| *Coreana raphaelis* | ATT/TAA | TTAG/T | ATG/T | ATC/TAA | ATG/TA | ATG/TAA | ATT/TAA | ATT/T | ATG/T | ATG/TA | ATA/TA | ATG/T | ATG/TAA |
| *Protantigius superans* | ATT/TAA | CGA/TAA | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/T | ATG/T | ATG/TAA | ATA/TAA | ATG/TAA | ATG/TAA |
| *Abisara fylloides* | ATT/TAA | CGA/T | ATG/TAA | ATC/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/T | ATG/T | ATG/T | ATA/TAA | ATG/TAA | ATG/TAA |
| ***Catopsilia pomona*** | **ATT/TAA** | **CGA/T** | **ATG/T** | **ATT/TAA** | **ATG/TAA** | **ATG/TAA** | **ATT/TAG** | **ATT/T** | **ATA/T** | **ATG/TAA** | **ATC/TAA** | **ATG/TAA** | **ATG/TAA** |
| ***Leptidea morsei*** | **ATT/TAA** | **CGA/T** | **ATG/T** | **ATC/TAA** | **ATG/TAA** | **ATG/TAA** | **ATA/TAA** | **ATT/T** | **ATA/T** | **ATG/TAA** | **ATT/TAA** | **ATG/TAA** | **ATG/TAA** |
| *Artogeia melete* | ATT/T | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAA | ATT/T | ATG/TAA | ATA/TAA | ATT/TAA | ATG/TAA | ATA/TAA |
| *Pieris rapae* | ATT/TAA | TTAAAG/T | ATG/TAA | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATT/TAG | ATG/TAA | ATG/TAA | ATT/TAA | ATG/TAA | ATA/TAA |
| *Delias hyparete* | ATT/TAA | CGA/T | ATG/T | ATT/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATT/T | ATG/T | ATA/TAA | ATT/TAA | ATG/TAA | ATA/TAA |
| *Aporia crataegi* | ATT/TAA | ATT/TAA | ATG/T | ATC/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATT/T | ATG/TA | ATG/T | ATC/TAA | ATG/TAA | ATG/TAA |
| *Ctenoptilum vasava* | ATT/TAG | ATT/TAA | ATG/T | ATA/TAA | ATG/TAA | ATG/TAA | ATT/TAG | ATA/TAA | ATG/TAA | ATG/TAA | ATA/TAA | ATG/TAA | ATG/TAA |

**Table 7.** The codon usage in the mitogenomes of *L. morsei* (*Lm*) and *C. pomona* (*Cp*).

| Amino acid | Codon | N | | RSCU | | Amino acid | Condon | N (%) | | RSCU | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *Lm* | *Cp* | *Lm* | *Cp* | | | *Lm* | *Cp* | *Lm* | *Cp* |
| Phe | UUU | 338 | 349 | 1.82 | 1.87 | Tyr | UAU | 169 | 177 | 1.81 | 1.84 |
| | UUC | 34 | 25 | 0.18 | 0.13 | | UAC | 18 | 15 | 0.19 | 0.16 |
| Leu | UUA | 454 | 470 | 4.96 | 5.31 | His | CAU | 59 | 59 | 1.79 | 1.79 |
| | UUG | 11 | 9 | 0.12 | 0.10 | | CAC | 7 | 7 | 0.21 | 0.21 |
| | CUU | 52 | 33 | 0.57 | 0.37 | Gln | CAA | 61 | 63 | 1.94 | 2.00 |
| | CUC | 4 | 3 | 0.04 | 0.03 | | CAG | 2 | 0 | 0.06 | 0.00 |
| | CUA | 28 | 16 | 0.31 | 0.18 | Asn | AAU | 247 | 241 | 1.84 | 1.91 |
| | CUG | 0 | 0 | 0.00 | 0.00 | | AAC | 21 | 12 | 0.16 | 0.09 |
| Ile | AUU | 435 | 444 | 1.86 | 1.93 | Lys | AAA | 93 | 103 | 1.82 | 1.87 |
| | AUC | 33 | 15 | 0.14 | 0.07 | | AAG | 9 | 7 | 0.18 | 0.13 |
| Met | AUA | 283 | 298 | 1.88 | 1.86 | Asp | GAU | 52 | 57 | 1.73 | 1.78 |
| | AUG | 18 | 22 | 0.12 | 0.14 | | GAC | 8 | 7 | 0.27 | 0.22 |
| Val | GUU | 68 | 66 | 2.25 | 2.03 | Glu | GAA | 68 | 71 | 1.70 | 1.87 |
| | GUC | 1 | 3 | 0.03 | 0.09 | | GAG | 12 | 5 | 0.30 | 0.13 |
| | GUA | 47 | 58 | 1.55 | 1.78 | Cys | UGU | 21 | 28 | 1.62 | 1.81 |
| | GUG | 5 | 3 | 0.17 | 0.09 | | UGC | 5 | 3 | 0.38 | 0.19 |
| Ser | UCU | 102 | 110 | 2.56 | 2.83 | Trp | UGA | 89 | 97 | 1.85 | 2.00 |
| | UCC | 10 | 5 | 0.25 | 0.13 | | UGG | 7 | 0 | 0.15 | 0.00 |
| | UCA | 85 | 86 | 2.13 | 2.21 | Arg | CGU | 15 | 19 | 1.09 | 1.46 |
| | UCG | 2 | 2 | 0.05 | 0.05 | | CGC | 0 | 1 | 0.00 | 0.08 |
| Pro | CCU | 72 | 74 | 2.38 | 2.41 | | CGA | 38 | 31 | 2.76 | 2.38 |
| | CCC | 9 | 9 | 0.30 | 0.29 | | CGG | 2 | 1 | 0.15 | 0.08 |
| | CCA | 39 | 40 | 1.29 | 1.30 | Ser | AGU | 41 | 32 | 1.03 | 0.82 |
| | CCG | 1 | 0 | 0.03 | 0.00 | | AGC | 3 | 1 | 0.08 | 0.03 |
| Thr | ACU | 76 | 95 | 2.03 | 2.50 | | AGA | 76 | 72 | 1.91 | 1.85 |
| | ACC | 9 | 7 | 0.24 | 0.18 | | AGG | 0 | 3 | 0.00 | 0.08 |
| | ACA | 65 | 50 | 1.73 | 1.32 | Gly | GGU | 47 | 55 | 0.98 | 1.09 |
| | ACG | 0 | 0 | 0.00 | 0.00 | | GGC | 5 | 1 | 0.10 | 0.02 |
| Ala | GCU | 70 | 75 | 2.37 | 2.52 | | GGA | 102 | 120 | 2.14 | 2.39 |
| | GCC | 13 | 6 | 0.44 | 0.20 | | GGG | 37 | 25 | 0.77 | 0.50 |
| | GCA | 34 | 37 | 1.15 | 1.24 | Stop | UAA | 0 | 0 | 0.00 | 0.00 |
| | GCG | 1 | 1 | 0.03 | 0.03 | | UAG | 0 | 0 | 0.00 | 0.00 |

Start and stop codons excluded from total codon counts; N, frequency of codon use; RSCU, relative synonymous codon usage.