Genetics and Molecular Research

Molecular phylogenetic and dating analysis of pierid butterfly species using complete mitochondrial genomes

Y. Cao[1], J.S. Hao[1,2], X.Y. Sun[2], B. Zheng[1] and Q. Yang[2]

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

Corresponding authors: J.S. Hao / Q. Yang
E-mail: jshaonigpas@sina.com / qunyang@nigpas.ac.cn

Genet. Mol. Res. 15 (4): gmr15049196
Received September 6, 2016
Accepted October 31, 2016
Published December 2, 2016
DOI http://dx.doi.org/10.4238/gmr15049196

ABSTRACT. Pieridae is a butterfly family whose evolutionary history is poorly understood. Due to the difficulties in identifying morphological synapomorphies within the group and the scarcity of the fossil records, only a few studies on higher phylogeny of Pieridae have been reported to date. In this study, we describe the complete mitochondrial genomes of four pierid butterfly species (*Aporia martineti*, *Aporia hippia*, *Aporia bieti*, and *Mesapia peloria*), in order to better characterize the pierid butterfly mitogenomes and perform the phylogenetic analyses using all available mitogenomic sequence data (13PCGs, rRNAs, and tRNAs) from the 18 pierid butterfly species comprising the three main subfamilies (Dismorphiinae, Coliadinae and Pierinae). Our analysis shows that the four new mitogenomes share similar features

with other known pierid mitogenomes in gene order and organization. Phylogenetic analyses by maximum likelihood and Bayesian inference show that the pierid higher-level relationship is: Dismorphiinae + (Coliadinae + Pierinae), which corroborates the results of some previous molecular and morphological studies. However, we found that the *Hebomoia* and *Anthocharis* make a sister group, supporting the traditional tribe Anthocharidini; in addition, the *Mesapia peloria* was shown to be clustered within the *Aporia* group, suggesting that the genus *Mesapia* should be reduced to the taxonomic status of subgenus. Our molecular dating analysis indicates that the family Pieridae began to diverge during the Late Cretaceous about 92 million years ago (mya), while the subfamily Pierinae diverged from the Coliadinae at about 86 mya (Late Cretaceous).

Key words: Mitogenomes; Pieridae; Phylogeny; Divergence time

INTRODUCTION

The insect mitochondrial genomes are usually circular molecules, approximately 14-19 kb, containing 37 genes (including 13 protein-coding, 22 tRNA, and 2 rRNA genes) and a non-coding A+T-rich region harboring the initiation sites for transcription and replication (Clayton, 1992). Owing to their unique characteristics, such as more phylogenetic information than the individual genes, maternal inheritance, medium size, and little recombination, the mitogenome sequences have been widely used in studies on phylogenetics, molecular evolution, evolutionary genomics, and phylogeography (Cameron, 2014). With the advancements in PCR technique, the insect mitogenome sequencing is relatively easy to perform now, and more than 500 insect mitogenomes have been determined, to date (Kim et al., 2014). However, among the huge mitogenomic data, only 14 sequences from Pieridae species have been reported.

The Pieridae is one of the largest butterfly families that contains about 83 genera and over 1100 species widely distributed around the world, especially in tropical Africa and Asia (Robbins, 1982; Braby, 2005). This family is traditionally divided into four subfamilies (Pseudopontiinae, Dismorphiinae, Coliadinae, Pierinae) (Braby et al., 2006; Wahlberg et al., 2014; Ding and Zhang, 2016). The subfamily Pseudopontiinae contains one genus *Pseudopontia*, probably only five species, and is distributed in some local areas of tropical Africa; the subfamily Dismorphiinae contains seven genera covering about sixty species, and is mostly distributed in the Neotropical areas. However, the subfamily Pierinae contains nearly sixty genera and over 900 species, and the subfamily Coliadinae contains about 18 genera covering about 220 species, both of the these subfamilies having worldwide distribution (Braby et al., 2006).

Previous studies on the systematics of Pieridae generated controversial results regarding the higher phylogeny of Pieridae (Ehrlich, 1958; Ehrlich and Ehrlich, 1967; Venables, 1993; Pollock et al., 1998; Braby et al., 2006; Wahlberg et al., 2014). Although the Pseudopontiinae and Dismorphiinae were suggested as sister groups, being reciprocally monophyletic in some studies (Ehrlich, 1958; Braby et al., 2006), the former subfamily was considered as intermediate between the Dismorphoiinae and Pierinae *s.l.* (including the Coliadinae, formerly the tribe Coliadini) by Clench (1955). The Coliadinae was shown as a strongly supported monophyletic group based on three mitochondrial gene sequence data

(Pollock et al., 1998), whereas it was shown as a paraphyletic group in morphological view (Venables, 1993); the monophyly of Pierinae was also generally assumed by researchers based on both molecular and morphological evidences (Ehrlich and Ehrlich, 1967; Braby et al., 2006; Wahlberg et al., 2014); the above two subfamilies were recovered as sister subfamilies in the majority of previous studies. Additionally, the relationships between Pseudopontiinae and Dismorphiinae, as well as between Pierinae and Coliadinae were recovered in most of the previous studies, whereas it was supported that the Pseudopontiinae was paraphyletic to Dismorphiinae (Venables, 1993). Furthermore, within the Pierinae, the relationship of tribe Pierini with Anthocharidini was also a controversial issue; these were considered to be paraphyletic or sisters by different research groups (Ehrlich, 1958; Braby et al., 2006; Wahlberg et al., 2014; Ding and Zhang., 2016). For example, Braby et al. (2006) found that the tribe Anthocharidini was nonmonophyletic, and thus suggested that *Hebomoia* should be removed from the Anthocharidini. However, Wahlberg et al. (2014) found that *Hebomoia* was sister to the rest of the *Anthocharis* clade, suggesting that it should be included in the Anthocharidini.

The study of the evolutionary history of the butterflies is still at an initial stage in that the fossil record is relatively scarce (Braby, 2005; Braby et al., 2006). To date, only a few of undoubted Pieridae fossils have been discovered, namely *Stolopsyche libytheoidea*, *Oligodonta florissantensis*, and *Coliates proserpina* (putatively the close relatives of genera *Pieris*, *Leodonta*, and *Aporia*, respectively) with their deposits preserved in the late Eocene [37.2-33.9 million years ago (mya)] or in the early Oligocene (33.5-30 mya) (Scudder, 1875, 1889; Brown, 1976). Based on the coevolutionary relationship between insects and their host plants, the Pieridae should have a relatively longer evolutionary history because their host plants (the Fabales, Brassicales, and Santalales) date about 60 mya in the Palaeocene age or even older than 90 mya in the late Cretaceous (Braby et al., 2006; Braby and Trueman, 2006). With respect to molecular dating, the Pierinae was estimated to diverge from Coliadinae at about more than 85 mya, based on nuclear *EF-1a* gene sequence data (Wheat et al., 2007).

In this study, we determined the complete mitogenomic sequence of three *Aporia* species (*Aporia bieti*, *Aporia hippia*, and *Aporia martineti*) and one species of their close relatives, namely *Mesapia peloria* for the first time. They were identified and verified by detailed morphological features described in the Monograph of Chinese Butterflies edited by Chou (1999).We also compared the genomic structure and composition, such as gene content, tRNA secondary structure, rRNA secondary structure, and gene order, with other available pierid butterfly species. Moreover, for the first time, the Pieridae phylogeny comprising the three main subfamilies was reconstructed based on sequence data of thirteen protein-coding genes (PCG). Additionally, the divergence time of these groups was estimated using a relaxed-clock method, aimed at presenting a timescale for the origin and diversification of the Pieridae. The results will preliminarily provide new evidence for the evolutionary history of Pieridae and shed new light on the molecular dating timescale of insects based on the mitogenomic data.

MATERIAL AND METHODS

Sample collection

Adult individuals of *A. martineti*, *A. hippia*, and *A. bieti* were collected from the Diebu county and Hezuo city of Gansu Province, China, in July 2006, while *M. peloria* was gathered from the Qilian Mountains, Qinghai Province, China, in July 2009. After sample collection,

the fresh tissues were immediately preserved in 100% ethanol and stored at -20°C until used for total genomic DNA extraction.

DNA extraction and PCR amplification

The total genomic DNA was extracted from the thoracic muscle of an adult individual by the improved glass powder method used in our laboratory (Hao et al., 2005). The universal short primers for amplification of the partial *COI*, *ND4*, and *Cytb* fragments were synthesized as described by Simon et al. (1994). The primers for the implication of long and the remaining short segments were designed after sequence alignment of the relevant butterfly mitochondrial genome sequences via the Clustal X1.8 (Thompson et al., 1997) and the Primer Premier 5.0 (Singh et al., 1998) software. All the primers were synthesized by Sangon Biotechnology Co. Ltd., Shanghai, China. Long PCR amplification were performed under the following conditions: an initial denaturation at 95°C for 5 min; 30 cycles of denaturation at 95°C for 50 s, annealing at 50°-55°C (depending on the primer pairs) for 50 s, extension at 68°C for 150 s, and a final extension at 68°C for 10 min. The PCR products were detected by 1.5% agarose gel electrophoresis, purified using a 3S Spin PCR Product Purification Kit (Sangon Biotechnology Co. Ltd., Shanghai) and sequenced directly with an ABI-377 automatic DNA sequencer. For each long PCR product, the complete, double-stranded sequence was determined by primer walking.

Sequence analysis

The raw sequences files were edited and assembled in BioEdit version 7.0 (Hall, 1999). The concatenated amino acid sequences of the 13 PCGs, rRNA genes, and A+T-rich regions were determined by alignment of the sequences with the homologous regions of other Pierinae mitogenome sequences using Clustal X1.8 software (Thompson et al., 1997). The tRNA genes were identified using the tRNAscan-SE 1.21 (Lowe and Eddy, 1997), and the putative tRNAs that could not be scanned were confirmed by sequence comparison between the determined and other available pierid butterflies. The tRNA and rRNA secondary structures were predicted by tRNAscan-SE 1.21 (Lowe and Eddy, 1997) or through comparison with the known homologous regions in the other related butterfly species (Cameron and Whiting, 2008; Wang et al., 2015). The nucleotide sequences of the PCGs were translated based on the invertebrate mtDNA genetic code. The nucleotide composition was calculated by using MEGA6.0 (Tamura et al., 2013) and the tandem repeats in the A+T-rich regions were predicted by the Tandem Repeats Finder (Benson, 1999). The four mitogenome sequences were deposited in the GenBank database under the accession Nos. KX495165 (*A. bieti*), KX495166 (*A. hippia*), KX495167 (*A. martineti*), and KX495168 (*M. peloria*).

Phylogenetic analysis

The ingroup taxa for the phylogenetic analyses include 18 available pierid butterfly species representing three subfamilies with available mitogenomes (four newly determined in this study, one previously determined in our lab, and 13 extracted from GenBank). Two Hesperiidae butterfly species, namely *Ochlodes venata* (GenBank accession No. NC018048) and *Celaenorrhinus maculosus* (GenBank accession No. NC022853) were selected as outgroups because of their relatively close relationships with the Pieridae. The maximum

likelihood (ML) and Bayesian inference (BI) phylogenetic trees were reconstructed based on the PCG 12 (matrix of the codon positions 1 and 2 of PCGs) data using the PAUP* version 4.0b8 (Swofford, 2002) and the MrBayes Version 3.1.2 (Huelsenbeck and Ronquist, 2001) software, respectively. According to the Akaike information criterion, the GTR+I+G model was selected as optimal for the analysis based on Modeltest 3.7 software (Posada and Crandall, 1998). In the ML analysis, the trees were searched through the TBR branch swapping method using a neighbor-joining tree as the starting tree; the nodal support values were assessed by bootstrap resampling calculated using 500 replicates. In the Bayesian analysis, 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 the datasets and each set was sampled every 100 generations with a burn-in of 25%, and when the average standard deviation of the split frequencies was less than 0.01, the chain convergence (stationarity) was considered to have been reached. The confidence values of the BI tree were presented as the Bayesian posterior probabilities in percentages.

Divergence time estimations

Based on the PCG12 data matrix, divergence time estimations were conducted with the Bayesian relaxed molecular clock method (Sanderson et al., 2004) using the program BEAST Version 1.7.5 (Drummond et al., 2012), which uses MCMC approximation to estimate the joint posterior probability of a tree topology. In the analysis, the minimum age for crown-group of the genus *Pieris* was constrained to be 37.2-33.9 mya based on the closely related fossil *S. libytheoidea* (Scudder, 1889), and that of the genus *Aporia* was constrained to be 33.5-30 mya based on its closely related fossil *C. proserpina* (Scudder, 1875). Both the fossils mentioned above were also used in the previous studies of molecular dating (Braby et al., 2006). In addition, the maximum age of the clade Pierinae that diverged with Coliadinae was set to be 90-85 mya based on the molecular dating for the origin of Brassicales, the hostplants for the Pierinae. The maximum age of the clade Pieridae was set to be 94-89 mya based on the estimate for the origin of their host plants, Fabales (Wikström et al., 2001). Every 1000 chain was sampled with the 12,500 burn-in value set under the Tree Annotator Version 1.7.5 (Drummond et al., 2012), discarding the aged samples before stationarity. The phylogenetic tree was viewed and edited by using Fig Tree Version 1.4.0 (Rambaut and Drummond, 2012).

RESULTS

General mitogenomic features

The four complete mitogenome sequences of *A. bieti*, *A. hippia*, *A. martineti*, and *M. peloria*, were 15,147, 15,154, 15,192, and 15,160 bp, respectively. Their sizes fell within the genome size range reported for other pierid butterflies. Each genome contained 13 PCGs (*COI1-COIII3*, *ATP6*, *ATP8*, *ND1-6*, *ND4L*, *Cytb*), 22 putative tRNA genes, 2 rRNA genes (*lrRNA* and *srRNA*), and a non-coding control area called A+T-rich region (Figure 1). The mitogenomes were similar, in terms of gene order and arrangement, to those of many other butterflies. Four of the 13 PCGs (*ND5*, *ND4*, *ND4L*, *ND1*), 8 tRNAs (*trnQ*, *trnC*, *trnY*, *trnF*, *trnH*, *trnP*, *trnL*, *trnV*), and 2 rRNAs (*rrnL* and *rrnS*) coded with the minority-strand, whereas the rest 23 genes coded with the majority-strand. Like other butterfly mitogenomes, the three

tRNAs located between the A+T-rich region and *ND2*, were also arranged in the pattern of $tRNA^{Met}$/$tRNA^{Ile}$/$tRNA^{Gln}$, which was different from that of $tRNA^{Ile}$/$tRNA^{Gln}$/$tRNA^{Met}$ found in the ancestral insects (Figure 1).

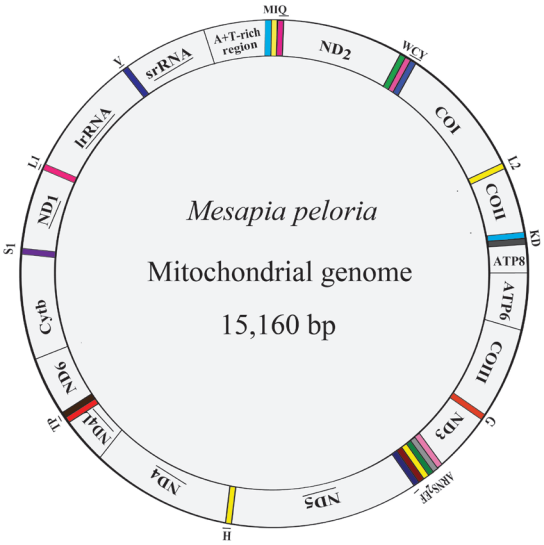

Figure 1. Circular map of *Mesapia peloria* mitochondrial genome. *COI*, *COII*, and *COIII* refer to the cytochrome oxidase subunits, *CytB* refers to cytochrome B, *ATP6* and *ATP8* refer to subunits 6 and 8 of F0 ATPase, and *ND1-6* refer to the components of NADH dehydrogenase. The tRNAs are denoted by the color blocks and indicated by the IUPAC-IUB single letter amino acid codes, whereas L1, L2, S1, and S2 denote $tRNA^{Leu}$*(CUN)*, $tRNA^{Leu}$*(UUR)*, $tRNA^{Ser}$*(AGN)*, and $tRNA^{Ser}$*(UCN)*, respectively. Gene names that are not underlined indicates transcription on the majority strand whereas underlines indicate transcription on the minority strand.

Like the other pierid species, the nucleotide composition of the four mitogenomes was significantly AT-biased, ranging from 79.5% in *Aporia hippia* to 81.1% in *Mesapia peloria* (Table 1). The AT content of different genes or regions of the mitogenomes also varied significantly; for example, the AT-rich regions were 94.7, 94.7, 93.1, and 92.8%, and the PCGs were 79.7, 79.1, 78.6, and 77.9% (exclusive of the termination codon) respectively (Table 1).

PCGs

The PCGs in the four species contained a total of 3715, 3717, 3718, and 3717 codons, respectively, and the sizes fell within the range for the other pierid butterfly species reported earlier (Table 1). For the four species, twelve of the 13 PCGs started with the typical ATN codons, whereas the *COI* began with CGA (Table 2). This CGA codon of *COI* is highly conserved throughout all the lepidopteran groups including all the butterfly groups and may be a synapomorphic character of the Lepidoptera (Kim et al., 2009). For example, *A. bieti ATP8* starts with ATA, whereas the others use ATT or ATC as their start codons; the *A. hippia* and *A. bieti ND6* start with ATT, whereas *M. peloria* and *A. martineti ND6* start with the codon ATC (Table 2). For the stop codon of the four species, the *ND3* uses TAG, the *COI*, *COII*, *ND4*, and *ND5* use single nucleotide T, whereas the remaining eight of the 13 PCGs use the standard TAA (Table 2).

Table 1. Summarized mitogenomic characteristics of the 20 butterfly species investigated in this study.

Species	Whole genome		PCG		16S rRNA		12S rRNA		A+T-rich region	
	Size (bp)	A+T (%)	Size (bp)	A+T (%)	Size (bp)	A+T (%)	Size (bp)	A+T (%)	Size (bp)	A+T (%)
Pieris canidia	15,153	79.7	3,712	78.2	1,323	83.9	767	84.2	418	91.2
Pieris rapae	15157	79.8	3,720	78.2	1,330	84.0	764	85.0	393	91.6
Pieris melete	15,140	79.8	3,714	78.5	1,319	83.4	777	86.9	351	88.0
Anthocharis bambusarum	15,180	80.2	3,721	78.9	1,348	84.0	780	84.1	318	90.3
Hebomoia glaucippe	15,701	79.9	3,715	78.0	1,323	83.7	777	85.2	899	92.2
Delias hyparete	15,186	79.8	3,703	78.4	1,311	80.7	786	78.2	377	92.0
Aporia hippia	**15,154**	**79.5**	**3,715**	**77.9**	**1,326**	**84.2**	**771**	**84.6**	**382**	**94.7**
Aporia martineti	**15,192**	**80.6**	**3,717**	**79.1**	**1,326**	**84.6**	**774**	**84.8**	**418**	**93.1**
Aporia bieti	**15,147**	**80.0**	**3,718**	**78.6**	**1,328**	**84.4**	**774**	**85.2**	**362**	**92.8**
Mesapia peloria	**15,160**	**81.1**	**3,717**	**79.7**	**1,323**	**85.0**	**772**	**85.2**	**378**	**94.7**
Aporia crataegi	15,140	81.3	3,708	79.9	1,326	85.4	779	85.5	354	95.2
Aporia intercostata	15,144	80.4	3,717	79.0	1,322	84.9	771	82.7	368	95.9
Prioneris clemanthe	15,115	80.7	3,713	79.3	1,342	84.9	752	84.6	341	95.3
Colias erate	15,184	81.3	3,719	80.0	1,327	84.7	767	85.2	364	95.1
Catopsilia Pomona	15,142	81.3	3,724	80.0	1,332	85.2	779	85.1	313	97.1
Gonepteryx rhamni	15,203	80.3	3,714	78.8	1,329	84.4	779	84.9	371	94.9
Gonepteryx mahaguru	15,221	80.9	3,712	79.4	1,334	84.5	778	85.7	385	95.1
Leptidea morsei	15,122	80.2	3,713	79.2	1,337	84.3	764	83.2	356	89.3
Ochlodes venata	15,622	82.0	3,715	80.2	1,334	84.7	770	85.6	587	94.0
Celaenorrhinus maculosus	15,282	79.9	3,722	78.3	1,413	83.7	776	84.9	331	92.4

The four new sequences listed in bold cells were determined in this study. The others were downloaded directly from GenBank.

Table 2. Initiation and termination codons for the 13 protein-coding genes (PCGs) of the 20 butterfly species used in this study.

Species	Predicted initiation and termination codons												
	ATP6	ATP8	COI	COII	COIII	Cytb	ND1	ND2	ND3	ND4	ND4L	ND5	ND6
Pieris melete	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATA/TAA	ATA/TAA	ATT/TAA	ATT/TAA
Gonepteryx mahaguru	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAA	ATG/TA	ATG/TAA	ATT/T	ATT/TAA
Pieris canidia	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATG/TAA	ATG/TAA	ATT/TAG	ATT/TAA
Hebomoia glaucippe	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATA/T	ATG/TAA	ATT/T	ATT/TAA
Pieris rapae	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATG/TAA	ATG/TAA	ATT/TAG	ATT/TAA
Delias hyparete	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/TAA	ATT/TAG	ATG/TAA	ATA/TAA	ATT/T	ATT/TAA
Aporia crataegi	ATG/TAA	ATC/TAA	ATT/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/TA	ATG/T	ATT/T	ATC/TAA
Catopsilia Pomona	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATA/T	ATG/TAA	ATT/T	ATC/TAA
Leptidea morsei	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATA/TAA	ATA/T	ATG/TAA	ATT/T	ATT/TAA
Gonepteryx rhamni	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATA/TAA	ATA/TAA	ATG/TAA	ATT/TAA	ATT/TAA	ATA/TA	ATG/TAA	ATT/T	ATT/TAA
Aporia intercostata	ATG/TAA	ATC/TAA	ATG/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/TAA	ATC/TAA
Anthocharis bambusarum	ATG/TAA	ATT/TAA	ATT/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/T	ATT/TAG	ATG/T	ATG/TAA	ATA/TAA	ATC/TAA
Colias erate	ATG/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATA/T	ATG/TAA	ATT/T	ATC/TAA
Prioneris clemanthe	ATG/TAA	ATC/TAA	CGA/T	ATG/T	ATA/TAA	ATA/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/T	ATG/TAA	ATT/T	ATT/TAA
Aporia martineti	**ATG/TAA**	**ATT/TAA**	**CGA/T**	**ATG/T**	**ATG/TAA**	**ATG/TAA**	**ATG/TAA**	**ATT/TAA**	**ATT/TAG**	**ATG/T**	**ATG/TAA**	**ATT/T**	**ATC/TAA**
Aporia bieti	**ATG/TAA**	**ATA/TAA**	**CGA/T**	**ATG/T**	**ATG/TAA**	**ATG/TAA**	**ATG/TAA**	**ATT/TAA**	**ATT/TAG**	**ATG/T**	**ATG/TAA**	**ATT/T**	**ATT/TAA**
Aporia hippia	**ATG/TAA**	**ATC/TAA**	**CGA/T**	**ATG/T**	**ATG/TAA**	**ATA/TAA**	**ATG/TAA**	**ATT/TAA**	**ATT/TAG**	**ATG/T**	**ATG/TAA**	**ATT/T**	**ATT/TAA**
Mesapia peloria	**ATG/TAA**	**ATC/TAA**	**CGA/T**	**ATG/T**	**ATG/TAA**	**ATG/TAA**	**ATG/TAA**	**ATT/TAA**	**ATT/TAG**	**ATG/T**	**ATG/TAA**	**ATT/T**	**ATC/TAA**
Celaenorrhins maculosus	ATG/TAA	ATT/TAA	ATT/T	ATG/T	ATG/TAA	ATG/TAA	ATA/TAA	ATT/T	ATT/TAG	ATG/T	ATG/TAA	ATT/T	ATT/TAA
Ochlodes venata	ATT/TAA	ATT/TAA	CGA/T	ATG/T	ATG/TAA	ATG/TAA	ATG/TAA	ATT/TAA	ATT/TAG	ATG/TAA	ATG/TAA	ATT/TAA	ATA/TAA

The four new sequences listed in bold cells were determined in this study. The others were downloaded directly from GenBank.

For the four species, the AT contents of 13 PCGs were 77.9, 78.6, 79.1, and 79.7% respectively, and these values were also similar to those detected in other pierid butterflies, which ranged from 78% in *Hebomoia glaucippe* to 80% in *Catopsilia pomona* (Table 1). When the first, second, and third codon positions were considered separately, the highest AT contents were the third positions for four pierid butterflies. In addition, the lowest C and G contents were also the third positions (Table 3).

tRNAs and rRNAs

All the four pierids contained 22 tRNA genes, among which 14 were encoded by the heavy chain and the other eight were encoded by the light chain. All the tRNAs can form the typical clover leaf secondary structure except the *tRNA*^{ser} (AGN), which lacked the dihydrouridine loop, as universally found in other butterfly mitogenomes (Zhang et al., 2013; Hao et al., 2014). All the tRNA genes had base pair mismatches, such as G·U, U·G, U·U, A·C, and A·A (Figure 2); these unusual base mismatches, which lead to the weak bond, could be corrected by RNA editing (Lavrov et al., 2000).

All the four mitogenomes contained two ribosomal RNA genes, namely the 16S rRNA and the 12S rRNA, which were located between *tRNA*$^{Leu(CUN)}$ and *tRNA*^{Val}, and between *tRNA*^{Val} and the A+T-rich region, respectively. The 16S rRNAs were 1326, 1326, 1328, and 1323 bp in size, with 84.2, 84.6, 84.4, and 85.0% A+T content; the 12S rRNAs were 771, 774, 774, and 772 bp in size, with 84.6, 84.8, 85.2, and 85.2% A+T content, respectively. These cases were similar to those reported for other insect mitogenomes (Table 1). The 12S rRNA and 16S rRNA secondary structures of the four species were predicted based on the comparison with the structures of other related lepidopteran rRNAs; the predicted topologies were generally in congruence with those reported from other lepidopterans (Cameron and Whiting, 2008; Sun et al., 2012; Shi et al., 2015). The structure of *A. hippia* 12S rRNA containing six domains (labeled I, II, III, IV, V, and VI) with 49 helices is presented, herein (Figure 3).

A+T-rich region

The A+T-rich regions of the four mitogenomes were all located between the *srRNA* and *tRNA*^{Met} and their lengths were 382, 418, 362, and 378 bp, respectively (Table 1). The A+T contents in these regions in the four mitogenomes were 94.7, 93.1, 92.8, and 94.7%, respectively (Figure 1 and Table 1). Ranging from the smallest *Pieris melete* (88.0%) to the largest *C. pomona* (97.1%), all the mitogenomes showed a significantly strong A+T bias, as observed in other insect groups (Li et al., 2015). The A+T-rich regions of all the four mitogenomes contained a poly-T stretch upstream of the characteristic motif ATAGA. The microsatellite-like elements, such as $(TA)_n$ (N = 6-10), which are frequently preceded by the ATTTA motif, are commonly present in the lepidopteran mitochondrial genomes (Figure 4).

Phylogenetic analysis

In the phylogenetic analyses, one data matrix PCG12 (owing to the substitution saturation of the third position) was used. There were 7372 sites in the PCG12 matrix in total. Both the phylogenetic trees (ML and BI) generated the same topologies only with a few different supporting values, showing that the Pieridae family constituted three well supported clades (subfamilies Dismorphiinae, Coliadinae, and Pierinae), and their relationships were [Dismorphiinae + (Coliadinae + Pierinae)] (Figure 5).

Table 3. Nucleotide compositions of the different regions or sites in the *Aporia bieti* (*Ab*), *Aporia hippia* (*Ah*), *Aporia martineti* (*Am*), and *Mesapia peloria* (*Mp*) mitogemomes.

Region (site)	A (%)				C (%)				G (%)				T (%)				A+T (%)			
	Ab	*Ah*	*Am*	*Mp*	*Ab*	*Ah*	*Am*	*Mp*	*Ab*	*Ah*	*Am*	*Mp*	*Ab*	*Ah*	*Am*	*Mp*	*Ab*	*Ah*	*Am*	*Mp*
Whole genome	39.6	39.8	39.7	39.8	12.3	12.9	12.0	11.6	7.6	7.5	7.4	7.3	40.4	39.7	40.9	41.4	80.0	79.5	80.6	81.1
Protein-coding genes	33.5	33.3	33.6	33.7	10.7	11.1	10.3	9.9	10.7	11.0	10.5	10.4	45.1	44.6	45.5	46.0	78.6	77.9	79.1	79.7
1st codon positions	37.2	37.0	36.9	37.4	10.0	10.4	10.1	9.5	15.2	15.5	15.2	14.9	38.0	37.0	38.0	38.0	75.2	74.0	74.9	75.4
2st codon positions	22.2	22.0	22.2	21.8	16.2	16.2	16.1	16	13.2	13.3	13.3	13.5	48.0	49.0	48.0	49.0	70.2	71.0	70.2	70.8
3st codon positions	41.1	40.9	41.8	41.7	6.0	6.7	4.8	4.1	3.7	4.1	3.0	2.9	49.0	48.0	50.0	51.0	90.1	88.9	91.8	92.7
12S rRNA genes	39.3	44.7	46.0	46.2	9.9	4.8	4.9	5.1	4.9	10.5	10.3	9.7	45.9	39.9	38.8	39.0	85.2	84.6	84.8	85.2
16S rRNA genes	44.0	41.6	43.8	43.8	5.3	11.1	4.9	4.9	10.4	4.8	10.5	10.1	40.4	42.6	40.8	41.2	84.4	84.2	84.6	85.0
A+T-rich region	47.2	45.5	46.9	46.0	3.3	3.7	2.4	3.2	3.9	1.6	4.5	2.1	45.6	49.2	46.2	48.7	92.8	94.7	93.1	94.7

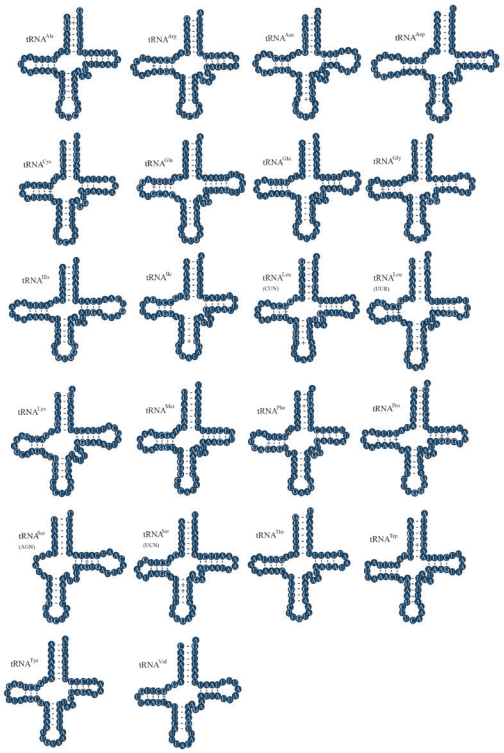

Figure 2. Predicted cloverleaf secondary structure of the *Aporia hippia* 22 tRNA genes.

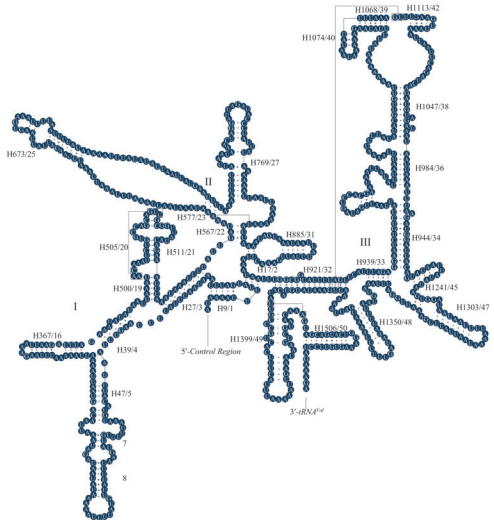

Figure 3. Predicted secondary structure of the *Aporia hippia* 12S rRNA. The conserved domains are denoted by Roman numerals and the numbered helices are after Cameron and Whiting (2008) and Wang et al. (2015). The tertiary structures are denoted by boxed bases joined by solid lines. The Watson-Crick pairs are joined by dashes; other interactions are joined by plus sign.

Aporia bieti

srRNA-14785ATTAATAAAAAATTTCACATAGATTTTTTTTTTTTTTTTTATATTAAAAATTTACATTTATAAAAAATATTTTAAT

The orgin of minority strand replation

AATTTCTTTCTTTCTTTTTTATAATATGTTTATAAATATAAATTTATAATAAACGATTAATAATAAGTTTAAATAAGAATAATTATA

AATATAATATTAATTTTTTCAATAATTTATTATATATATATATATTATAATTATATAAATATTTAATATATTATTAATTTTAATAATAAAT

A micro satellite (TA)$_7$

TAAATATTTAATATATATGAATATATATGTGAATTTATGTGTATAAGAAATTCTATACCAATTACAAAAATTTTAATATATAATACAT

ATGTACATATAATAAAAA 15,147-tRNAMet

A poly-A motif

Aporia hippia

srRNA-14772CTATTTATAAAAAATTTTACATAGATTTTTTTTTTTTTATATTAAAAATTTACATTTATAATAAATATTTTAATAATT

The orgin of minority strand replation

TTATTCTTTCTTTTCATAATATGTTTATAAATATAAATTTATAATAAACGATTAATAATAAGTTTAAATAAAAATAATTATAAATATA

ATATTAATTTTTTAATAATTTATTATATTATATTTATATATATATATTAAATATTTAATTTATTATTAATTTTCATAATAAATTAAATA

A micro satellite (TA)$_6$

TTTAATATATATATATAATATAATATAATAAATTATTAAACCAACTTCAATAATTTTTTCATATAATATATAATTCAATATAAGGAAT

A micro satellite (TA)$_6$

TTTTAATTTAAATAATTTTTATAATATATAATAATA 15,154-tRNAMet

Aporia martineti

srRNA-14774ATTTATAAAAAATTTTAAATAGATTTTTTTTTTTTTTTATATTAAAAATTTATATTTATAATAAATATTTTAATAATTT

The orgin of minority strand replation

TATTCTTTCTTTTTCATAATATTATTATAAATATAAATTTATAATAAACGATCAATAATAAGTTTAAATAAGAATAATTATAAATATA

ATATTAATTTTTTAATAATTTATTATATTATATATATATATTATATAATTATATAAATATTTAATATATTAATAATTTTTATAATATATTAAATA

A micro satellite (TA)$_7$

TTTAATATGCATATATATATATATATACGCACACATAAATACATATATACATTTAGTGTATAAAAAGTTTTATGCCAATTTCAAAAA

A micro satellite (TA)$_8$

TTTTAATATATAATATATAATACATAATTCATAAAAAAATTCAATACTTATTATATATATATATATATATATA 15,192-tRNAMet

(TA)$_{10}$

Mesapia peloria

srRNA-14782ATTAATAAAAAATTTCATATAGATTTTTTTTTTTTTTTATATTAAAAATTTACATTTATAATAAATATTTTAATAATT

The orgin of minority strand replation

TTACTCTTTTTTTTTCATAATATGTTTATAAATATAAATTCATAATAAACGGTCAATAATAAGTTTAAATAAAAATAATTATAAATA

TAATATTAATTTTTTAATAATTTATTATATATATATATTATAATTATTTAAATATTTAATTTATTATTAATTTTTATAATAAATTAAATAT

A micro satellite (TA)$_6$

TTAATATATATATATATATATAAACAATTTAGTGTATAAAAAGTTATATACCAATTTCAAAAATTTTAATATATAATATATAATAATT

A micro satellite (TA)$_9$

AATTTATAAATTTAATATATATATATATATATA15,160-tRNAMet

A micro satellite (TA)$_9$

Figure 4. A+T-rich regions of the four pierid butterfly species investigated in this study.

These phylogenetic results are supported by numerous previous studies based on morphological characters and molecular sequence data (Ehrlich, 1958; Braby et al., 2006).

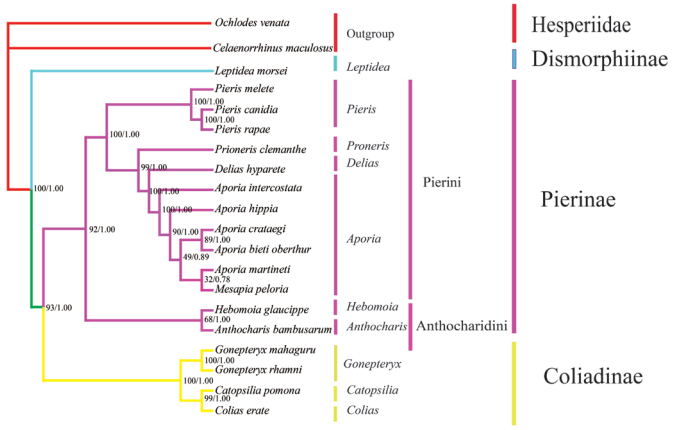

Figure 5. Maximum likelihood and Bayesian inference phylogenetic trees of the 18 pierid species investigated in this study based on the concatenated nucleotide sequences of 13 protein-coding gene sequence data (the value of bootstrap support and posterior probability is shown on each node of the tree).

The subfamily Dismorphiinae only represented by *Leptidea morsei* in this study is a clade which is sister to the grouping of the other two subfamilies; the Coliadinae and Pierinae are all monophyletic and sister to each other with strong supporting values (bootstrap: 93% ML, 1.00 BI) (Figure 5). In addition, the Pseudopontiinae, traditionally treated as a unique subfamily covering only a few species was not included in this study; however, its close relationship with the Dismorphiinae was repeatedly confirmed by numerous previous studies based on morphological and molecular evidences (Ehrlich, 1958; Braby et al., 2006; Wahlberg et al., 2014), and it was even regarded as a subordinate taxa (Braby et al., 2006).

The Pierinae contains three well supported subclades: the first is the *Pieris*, the second is the grouping of *Aporia* + *Delias* + *Prioneris*, and the third is *Hebomoia* + *Anthocharis* (Figure 5). This result is generally in agreement with those of the previous molecular and morphological studies (Ehrlich, 1958; Braby et al., 2006; Wahlberg et al., 2014; Ding and Zhang, 2016); however, it is inconsistent with the results reported by Braby et al. (2006) regarding the position of *Hebomoia*, which stands basal on the phylogenetic tree of the subfamily Pierinae, and paraphyletic to *Anthocharis*. Within the tribe Pierini, both ML and BI analyses showed the *M. peloria* was clustered within the *Aporia* species, suggesting that the genus *M.* should probably be reclassified into *Aporia* as a subordinate taxon, an opinion held by Braby et al. (2006), Qiao et al. (2014), and Ding and Zhang (2016).

The Coliadinae contains two well supported subclades, namely *Gonepteryx* and *Colias* + *Catopsilia*, in accordance with previous molecular studies (Braby et al., 2006; Wahlberg et al., 2014; Ding and Zhang., 2016), suggesting that the traditional higher classification of Coliadinae should be robust.

Divergence time estimations

As of date, only a few molecular dating studies concerning the Pieridae have been

conducted (Braby et al., 2006; Wheat et al., 2007). The present analysis represents the first divergence time estimation on Pieridae with Bayesian relaxed molecular clock method by using the mitogenomic PCG data. Our molecular dating results regarding the higher Pieridae groups are shown in Figure 6. The Pieridae began to diverge at about 92.45 mya in the middle or late Cretaceous Period (95%CI = 89.78-95.13 mya); the divergence between Coliadinae and Pierinae was dated at about 86.34 mya (95%CI = 83.59-88.83 mya) in the late Cretaceous, and the two subfamilies began to diverge at about 76.46 mya (95%CI = 66.76-83.85 mya) and 51.05 mya (95%CI = 31.8-72.68 mya), respectively. These results are generally consistent with the corresponding time estimates obtained in previous studies based on the elongation factor *1α*, wingless, *COI*, and *28S* rRNA data (Braby et al., 2006; Braby and Trueman, 2006; Wheat et al., 2007). It is noted that the divergences of the main pierid groups during the Cretaceous time corresponds to a world greenhouse period when angiosperms and many modern animal groups also began to diversify and active tectonic movements causing large scale geographic isolations began to happen.

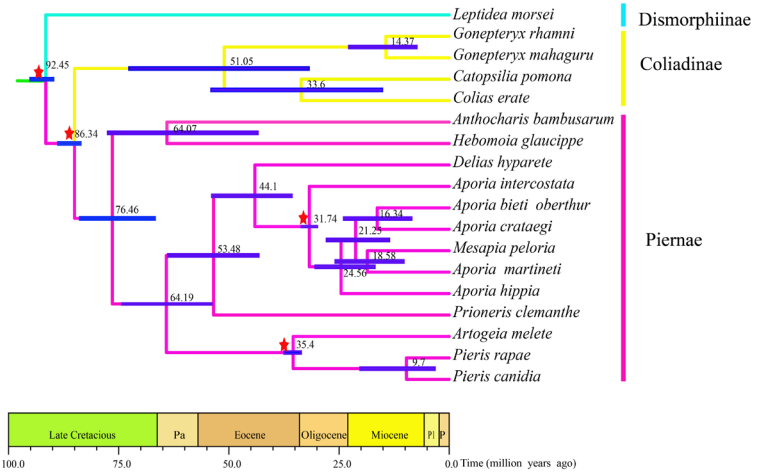

Figure 6. Evolutionary time-scale of the main Pieridae groups estimated by relaxed molecular clock method based on the 13 mitogenomic PCG sequence data (fossil or hostplant calibration points are shown by asterisks and the confidence intervals are shown by blue-colored bars on the phylogram). P: the Pleistocene series, Pl: the Pliocene series, Pa: the Paleocene series.

DISCUSSION

Our study provides the first molecular phylogenetic and molecular dating analyses of the Pieridae based on the mitochondrial genome sequence data. The results obtained in the present study were mostly consistent with those reported in some of the previous studies (Braby et al., 2006; Wahlberg et al., 2014; Ding and Zhang, 2016). However, these results were somewhat incongruent with those reported by Braby et al. (2006) in which the genus *Hebomoia* was not shown within the tribe Anthocharidini. Furthermore, our results indicated that in all the phylogenetic trees, the genus *Mesapia* was clustered within the genus *Aporia*, and thus it should be reduced to the subgenus rank considering their remarkably smaller body size compared to those of the other *Aporia* species. This opinion was also proposed by Qiao et al. (2014) and Ding and Zhang (2016).

The phylogenetic analyses based on the mitochondrial genome data are usually limited by the whole sequencing of the total genome and accordingly the sufficient taxa sampling for specific target group. Additionally, the underlying computational methods, such as the selection of the evolutionary model, test of substitution saturation, etc., are relatively more complex than those of a single or few gene sequence data. In this study, the preliminary phylogenetic and dating analyses of Pieridae butterflies were conducted using the relatively smaller number of mitogenome sequence data from the present study as well as the data available in the GenBank. Comprehensive analyses based on more whole mitogenomic sequence data from this group should be undertaken in future studies.

ACKNOWLEDGMENTS

We express our sincere thanks to Dr. Dunyuan Huang and Mr. Bo He (Jiangxi Environmental Engineering College, Ganzhou, China) for some specimen collection. Research supported by the National Science Foundation of China (Grant #41172004) and the Opening Funds from the State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences (Grant #113110).

REFERENCES

Benson G (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27: 573-580. http://dx.doi.org/10.1093/nar/27.2.573

Braby MF (2005). Provisional checklist of genera of the Pieridae (Lepidoptera: Papilionoidea). *Zootaxa* 832: 1-16.

Braby MF and Trueman JWH (2006). Evolution of larval host plant associations and adaptive radiation in pierid butterflies. *J. Evol. Biol.* 19: 1677-1690. http://dx.doi.org/10.1111/j.1420-9101.2006.01109.x

Braby MF, Vila R and Pierce NE (2006). Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. *Zool. J. Linn. Soc.* 147: 239-275. http://dx.doi.org/10.1111/j.1096-3642.2006.00218.x

Brown FM (1976). *Oligodonta ñorissantensis*, gen. n., sp. nov. (Lepidoptera: Pieridae). *Bull. Allyn Museum* 37: 1-4.

Cameron SL (2014). Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu. Rev. Entomol.* 59: 95-117. http://dx.doi.org/10.1146/annurev-ento-011613-162007

Cameron SL and Whiting MF (2008). The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene* 408: 112-123. http://dx.doi.org/10.1016/j.gene.2007.10.023

Chou I (1999). Monograph of Chinese Butterflies, First Volume, Revised edn. Henan Scientific and Technological Publishing House, Zhengzhou, 408.

Clayton DA (1992). Transcription and replication of animal mitochondrial DNAs. *Int. Rev. Cytol.* 141: 217-232. http://dx.doi.org/10.1016/S0074-7696(08)62067-7

Clench HK (1955). Revised classiðcation of the butterñy family Lycaenidae and its allies. *Ann. Carnegie Mus.* 33: 261-274.

Ding C and Zhang Y (2016). Phylogenetic relationships of Pieridae (Lepidoptera: Papilionoidea) in China based on seven gene fragments. *Entomol. Sci.* http://dx.doi.org/10.1111/ens.12214.

Drummond AJ, Suchard MA, Xie D and Rambaut A (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29: 1969-1973. http://dx.doi.org/10.1093/molbev/mss075

Ehrlich PR (1958). The comparative morphology, phylogeny and higher classification of the butterflies (Lepidoptera: Papilionoidea). *Univ. Kans. Sci. Bull.* 39: 305-364.

Ehrlich PR and Ehrlich AH (1967). The phenetic relationships of the butterflies. I. Adult taxonomy and the non-specificity hypothesis. *Syst. Biol.* 16: 301-317.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser* 41: 95-98.

Hao JJ, Hao JS, Sun XY, Zhang LL, et al. (2014). The complete mitochondrial genomes of the Fenton's wood white, *Leptidea morsei*, and the lemon emigrant, *Catopsilia pomona*. *J. Insect Sci.* 14: 130. http://dx.doi.org/10.1093/jis/14.1.130

Hao JS, Li CX, Sun XY and Yang Q (2005). The phylogeny and divergence time estimation of cheilostome bryozoans based on mitochondrial *16S rRNA* sequences. *Chin. Sci. Bull.* 12: 1205-1211. http://dx.doi.org/10.1007/BF03183694

Huelsenbeck JP and Ronquist F (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754-755. http://dx.doi.org/10.1093/bioinformatics/17.8.754

Kim MI, Baek JY, Kim MJ, Jeong HC, et al. (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. *Mol. Cells* 28: 347-363. http://dx.doi.org/10.1007/s10059-009-0129-5

Kim MJ, Wang AR, Park JS and Kim I (2014). Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperiidae) and phylogenetic reconstruction of Lepidoptera. *Gene* 549: 97-112. http://dx.doi.org/10.1016/j.gene.2014.07.052

Lavrov DV, Brown WM and Boore JL (2000). A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus. Proc. Natl. Acad. Sci. USA* 97: 13738-13742. http://dx.doi.org/10.1073/pnas.250402997

Li H, Shao R, Song N, Song F, et al. (2015). Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. *Sci. Rep.* 5: 8527. http://dx.doi.org/10.1038/srep08527

Lowe TM and Eddy SR (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955-964. http://dx.doi.org/10.1093/nar/25.5.0955

Pollock DD, Watt WB, Rashbrook VK and Iyengar EV (1998). Molecular phylogeny for *Colias* butterñies and their relatives (Lepidoptera: Pieridae). *Ann. Entomol. Soc. Am.* 91: 524-531. http://dx.doi.org/10.1093/aesa/91.5.524

Posada D and Crandall KA (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818. http://dx.doi.org/10.1093/bioinformatics/14.9.817

Qiao F, Chen ZN, Chen Z, Xie HC, et al. (2014). Phylogenetic relationships among seven Pieridae genera based on partial sequences of the *COI* and *Cytb* genes. *Chin. J. Appl. Entomol* 51: 127-138.

Rambaut A and Drummond AJ (2012). FigTree version 1.4. URL http://tree. bio. ed. ac. uk/software/figtree/.

Robbins RK (1982). How many butterñy species? *News Lepid. Soc* 3: 40-41.

Sanderson MJ, Thorne JL, Wikström N and Bremer K (2004). Molecular evidence on plant divergence times. *Am. J. Bot.* 91: 1656-1665. http://dx.doi.org/10.3732/ajb.91.10.1656

Scudder SH (1875). Fossil butterñies. *Mem. Am. Assoc. Adv. Sci.* 1: 1-99.

Scudder SH (1889). The fossil butterñies of Florissant. *US Geological Survey, 8th Annual Report Part I*, 433-474.

Shi QH, Sun XY, Wang YL, Hao JS, et al. (2015). Morphological characters are compatible with mitogenomic data in resolving the phylogeny of nymphalid butterflies (lepidoptera: papilionoidea: nymphalidae). *PLoS One* 10: e0124349. http://dx.doi.org/10.1371/journal.pone.0124349

Simon C, Frati F, Beckenbach A, Crespi B, et al. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* 87: 651-701. http://dx.doi.org/10.1093/aesa/87.6.651

Singh VK, Mangalam AK, Dwivedi S and Naik S (1998). Primer premier: program for design of degenerate primers from a protein sequence. *Biotechniques* 24: 318-319.

Sun QQ, Sun XY, Wang XC, Gai YH, et al. (2012). Complete sequence of the mitochondrial genome of the Japanese buff-tip moth, *Phalera flavescens* (Lepidoptera: Notodontidae). *Genet. Mol. Res.* 11: 4213-4225. http://dx.doi.org/10.4238/2012.September.10.2

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

Tamura K, Stecher G, Peterson D, Filipski A, et al. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729. http://dx.doi.org/10.1093/molbev/mst197

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882. http://dx.doi.org/10.1093/nar/25.24.4876

Venables BAB (1993). Phylogeny of the white and sulphur butterñies (Pieridae). Dissertation project proposal by Dr. Robbins RK, National Museum of Natural History, Smithsonian Institution.

Wahlberg N, Rota J, Braby MF, Pierce NE, et al. (2014). Revised systematics and higher classification of pierid butterflies (Lepidoptera: Pieridae) based on molecular data. *Zool. Scr.* 43: 641-650. http://dx.doi.org/10.1111/zsc.12075

Wikström N, Savolainen V and Chase MW (2001). Evolution of the angiosperms: calibrating the family tree. *Proc. Biol. Sci.* 268: 2211-2220. http://dx.doi.org/10.1098/rspb.2001.1782

Wang Y, Chen YH, Xia CC, Xia XQ, et al. (2015). The complete mitochondrial genome of the Common Red Apollo, *Parnassius epaphus* (Lepidoptera: Papilionidae: Parnassiinae). *J. Asia Pac. Entomol.* 18: 239-248. http://dx.doi.org/10.1016/j.aspen.2015.02.002

Wheat CW, Vogel H, Wittstock U, Braby MF, et al. (2007). The genetic basis of a plant-insect coevolutionary key innovation. *Proc. Natl. Acad. Sci. USA* 104: 20427-20431. http://dx.doi.org/10.1073/pnas.0706229104

Zhang LL, Hao JS, Huang DY, Sun XY, et al. (2013). Complete mitochondrial genomes of the Bright Sunbeam *Curetis bulis* and the Small Copper *Lycaena phlaeas* (Lepidoptera: Lycaenidae) and their phylogenetic implications. *Genet. Mol. Res.* 12: 4434-4445. http://dx.doi.org/10.4238/2013.October.10.9