

RESEARCH ARTICLE

# Morphological Characters Are Compatible with Mitogenomic Data in Resolving the Phylogeny of Nymphalid Butterflies (Lepidoptera: Papilionoidea: Nymphalidae)

Qing-Hui Shi<sup>1,3</sup>, Xiao-Yan Sun<sup>2</sup>, Yun-Liang Wang<sup>1</sup>, Jia-Sheng Hao<sup>1,2\*</sup>, Qun Yang<sup>2\*</sup>

**1** College of life sciences, Anhui Normal University, Wuhu, China, **2** State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Palaeontology and Stratigraphy, Chinese Academy of Sciences, Nanjing, China, **3** Sanming College, Sanming, China

\* [jshaonigpas@sina.com](mailto:jshaonigpas@sina.com) (JSH); [qunyang@nigpas.ac.cn](mailto:qunyang@nigpas.ac.cn) (QY)



**OPEN ACCESS**

**Citation:** Shi Q-H, Sun X-Y, Wang Y-L, Hao J-S, Yang Q (2015) Morphological Characters Are Compatible with Mitogenomic Data in Resolving the Phylogeny of Nymphalid Butterflies (Lepidoptera: Papilionoidea: Nymphalidae). PLoS ONE 10(4): e0124349. doi:10.1371/journal.pone.0124349

**Academic Editor:** Xiao-Yue Hong, Nanjing Agricultural University, CHINA

**Received:** December 27, 2014

**Accepted:** March 10, 2015

**Published:** April 10, 2015

**Copyright:** © 2015 Shi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data are available from the Genbank (accession numbers KF990128, KF990126, KF906483, KF906485, KF906484, KF990129, KF990124, KF990123, KF990125, KF990127, KM018020).

**Funding:** This work was supported by the National Science Foundation of China (Grant No.41172004, 41472028), 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). The funders had no role in study design, data

## Abstract

Nymphalidae is the largest family of butterflies with their phylogenetic relationships not adequately approached to date. The mitochondrial genomes (mitogenomes) of 11 new nymphalid species were reported and a comparative mitogenomic analysis was conducted together with other 22 available nymphalid mitogenomes. A phylogenetic analysis of the 33 species from all 13 currently recognized nymphalid subfamilies was done based on the mitogenomic data set with three Lycaenidae species as the outgroups. The mitogenome comparison showed that the eleven new mitogenomes were similar with those of other butterflies in gene content and order. The reconstructed phylogenetic trees reveal that the nymphalids are made up of five major clades (the nymphaline, heliconiine, satyrine, danaine and libytheine clades), with sister relationship between subfamilies Cyrestinae and Biblidinae, and most likely between subfamilies Morphinae and Satyrinae. This whole mitogenome-based phylogeny is generally congruent with those of former studies based on nuclear-gene and mitogenomic analyses, but differs considerably from the result of morphological cladistic analysis, such as the basal position of Libytheinae in morpho-phylogeny is not confirmed in molecular studies. However, we found that the mitogenomic phylogeny established herein is compatible with selected morphological characters (including developmental and adult morpho-characters).

## Introduction

The family Nymphalidae (Lepidoptera: Papilionoidea) is the largest group of butterflies with about 7,200 species distributed on all continents except Antarctica [1–3]. They are usually medium to large sized butterflies with striking colors, specialized insect-plant interaction and a deep rooted evolutionary history probably 90 million years ago [4, 5]. Owing to their species richness and ecological diversification, nymphalids have been used as model taxa for a wide

collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

range of evolutionary and ecological studies [6]. However, the phylogeny of Nymphalidae is one of the most controversial issues in insect studies [7–10]. For instance, in Chou's taxonomic system, widely adopted in China, the nymphalid butterflies are split into 10 morphological subfamilies [4], while the most commonly adopted classification of the family, proposed by Ackery et al. [11], Harvey [12] and Müller [13], consists of 13 subfamilies. Recently, although efforts have been focused on the phylogeny of Nymphalidae by using morphological, molecular (mitochondrial and nuclear genes), or combined data, problems still exist, particularly about the phylogenetic status of Biblidinae, Pseudergolinae, Cyrestinae, Morphinae, Libytheinae and Charaxinae [14–21]. Common to previous molecular phylogenetic studies are small number of molecular markers used with conflicting results and often weak supports about their phylogenetic relationships.

The insect mitochondrial genomes (mitogenomes) provide abundant molecular markers for phylogenetic and population genetic studies at various hierarchical levels, due to their unique features, such as maternal inheritance, lack of extensive recombination and moderate nucleotide substitution rates [22, 23].

Up to now, existing information about butterfly mtDNAs, especially the complete genomes of mtDNA are still limited. Only 22 nymphalid mitogenomes so far deposited in GenBank (up to Jun 15, 2014) represent eight of the thirteen nymphalid subfamilies. However, the corresponding sequences of the other major nymphalid groups (e.g., Charaxinae, Pseudergolinae, Cyrestinae, Biblidinae and Morphinae) are still lacking. In this study, we sequenced nine complete and two nearly complete mitogenomes of nymphalids, representing nine subfamilies (*Hypolimnas bolina* for Nymphalinae, *Cethosia bibles* for Heliconiinae, *Polyura nepenthes* for Charaxinae, *Ariadne ariadne* for Biblidinae, *Elymnias hypermnestra*, *Lethe dura* and *Callerebia suroia* for Satyrinae, *Parantica aglea* for Danainae, *Cyrestis thyodamas* for Cyrestinae, *Dichoragia nesimachus* for Pseudergolinae and *Stichophthalma howqua* for Morphinae) and compared them with other available nymphalid mitogenomes. A phylogenetic analysis based on mitogenomic data set, assisted with morphological comparison, is attempted to further resolve the nymphalid phylogeny.

## Materials and Methods

### Ethics Statement

All samples of butterfly species used in this study were collected from the mountains or peri-urban areas of China, where no specific permissions were required for these locations or activities. There are no endangered or protected species and all the all samples are common nymphalid butterfly species which are not included in the "List of Protected Animals in China" and other related lists of animal species protection in the world. Thus the sampling in this study did not violate any law, rule or regulation in China and all around the world, requiring no ethical or institutional approval.

### Samples and DNA extraction

The samples of the eleven newly sequenced nymphalid species were collected in China from 2006 to 2012 (S1 Table). After sample collection, fresh individuals were preserved in 100% ethanol and stored at -20°C until used for genomic DNA extraction. Total genomic DNA was extracted from the thorax muscle of a single individual using the Sangon Animal genome DNA Extraction Kit in accordance with the manufacturer instructions (Shanghai, China).

## Primer design, PCR amplification, and sequencing

Insect universal primers were initially used for the amplification of partial fragments of four genes (*cox1*, *cox3*, *cob* and *rrnS* gene) [24–26], approximately 500–700 bp in length. Primers for amplification of the *nad2*, *nad5* and overlapping long fragments were designed via the sequence alignment of all the other sequenced butterfly mitogenomes available, using ClustalX 2.1 and Primer Premier 5.0 software [27, 28]. All primers were synthesized by the Sangon Biotechnology Co. Ltd., Shanghai, China. Six long fragments (*cox1*—*cox3*, *cox3*—*nad5*, *nad5*—*cob*, *cob*—*rrnS*, *rrnS*—*nad2*, *nad2*—*cox1*) were amplified via long PCR technique with the cycling parameters: 30 cycles of denaturation at 95°C for 50 seconds, annealing at 49–55°C (depending on primer pairs) for 50 seconds, and elongation at 68°C for 150 seconds, and a final extension at 68°C for 10 minutes. The PCR products were visualized by electrophoresis on 1.2% agarose gel, then purified using a 3S Spin PCR Product Purification Kit and sequenced directly for the majority of fragments or by cloning the *rrnS*—*nad2* for some species, with an ABI-377 automatic DNA sequencer. All of the long PCR fragments were sequenced using the primer walking strategy.

## Sequence assembly, gene annotation, and analysis

Sequences from overlapping fragments were initially assembled via the alignment of neighboring fragments using the BioEdit 7.0 [29] and ClustalX 2.1 softwares [27]. Protein-coding genes (PCGs) and ribosomal RNA genes (rRNAs) were identified using ClustalX 2.1 software [27] and the NCBI Internet BLAST search function. Identification of transfer RNA genes (tRNAs) was conducted using software tRNAscan-SE 1.21 (<http://lowelab.ucsc.edu/tRNAscan-SE/>) [30]. Putative tRNAs that could not be found by tRNAscan-SE were identified by alignment with those of the other butterfly species available. The tandem repeats in the A+T-rich region were predicted using the Tandem Repeats Finder online (<http://tandem.bu.edu/trf/trf.html>) [31]. Nucleotide composition and codon usage were calculated by MEGA 5.0 software [32]. All three codon positions of the 13 PCGs concatenated nucleotide dataset were tested independently for substitution saturation, by plotting the number of observed substitutions against the genetic distance estimates using MEGA 5.0 software [32]. The rRNA secondary structures were predicted after the models for *Drosophila* [33], *Apis mellifera* [34] and *Manduca sexta* [35]. Stem-loops were named using both the conventions of Gillespie et al. [34] and Niehuis et al. [36, 37] with the former notation first for each time they are mentioned.

## Phylogenetic analysis

Thirty three nymphalid species with complete or nearly complete mitogenome sequence data, including eleven newly determined in this study, were used in the phylogenetic analyses, representing all currently recognized subfamilies of Nymphalidae (Table 1). Three species, *Coreana raphaelis*, *Spindasis takanonis* and *Protantigius superans* from Lycaenidae, being sister group of Nymphalidae, were selected for outgroup comparisons.

The phylogenetic trees were reconstructed with the neighbor joining (NJ), maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI) methods based on two different nucleotide datasets as follows: D1 (13 PCGs only) and D2 (13 PCGs plus 2 rRNAs). The mitochondrial genes were separately aligned by MUSCLE in MEGA 5.0 software [32]. The PCGs were aligned according to their amino sequence similarity, whereas rRNAs were directly aligned according to sequence similarity using default settings. Then each of individual alignments was concatenated as a combined matrix. However, the scattergrams analysis showed that the substitution for third codon position of 13 PCGs trends toward saturation, potentially obscuring phylogenetic signals. Therefore, all third codon positions were excluded during the

**Table 1. List of the nymphalid species analyzed in this study.**

Subfamily	Tribe	Species	Size (bp)	Genbank accession no.	Resources
Charaxinae	Charaxini	<i>Polyura nepenthes</i>	15,333	KF990128	This study
Calinaginae	-	<i>Calinaga davidis</i>	15,267	HQ658143	Xia et al., 2011
Pseudergolinae	-	<i>Dichorragia nesimachus</i>	14,367	KF990126	This study
Satyriinae	Melanitini	<i>Melanitis Leda</i>	15,122	JF905446	Shi et al., 2013a
	Satyrini	<i>Eumenis autonoe</i>	15,489	GQ868707	Kim et al., 2010
	Satyrini	<i>Callerebia suroia</i>	15,208	KF906483	Shi et al., 2014
	Elymniini	<i>Elymnias hypermnestra</i>	15,167	KF906484	This study
	Elymniini	<i>Lethe dura</i>	15,259	KF906485	This study
Morphinae	-	<i>Stichopthalma howqua</i>	14,020	KF990129	This study
Heliconiinae	Argynnini	<i>Argyreus hyperbius</i>	15,156	JF439070	Wang et al., 2011
	Argynnini	<i>Fabriciana nerippe</i>	15,140	JF504707	Kim et al., 2011a
	Argynnini	<i>Issoria lathonia</i>	15,172	NC_018030	Unpublished
	Heliconiini	<i>Heliconius melpomene</i>	15,325	HE579083	Dasmahapatra et al., 2012
	Acraeini	<i>Acraea issoria</i>	15,245	GQ376195	Hu et al., 2010
	Acraeini	<i>Cethosia biblis</i>	15,211	KF990124	This study
Limenitidinae	Limenitidini	<i>Parathyma sulphitia</i>	15,268	JQ347260	Tian et al., 2012
Apaturinae	-	<i>Apatura ilia</i>	15,242	JF437925	Chen et al., 2012
	-	<i>Apatura metis</i>	15,236	NC_015537	Zhang et al., 2012
	-	<i>Sasakia charonda</i>	15,244	NC_014224	Wang et al., 2012
	-	<i>Sasakia funebris</i>	15,233	JX131328	Wang et al., 2013
	-	<i>Timelaea maculata</i>	15,178	KC572131	Cao et al., 2013
Biblidinae	-	<i>Ariadne ariadne</i>	15,179	KF990123	This study
Cyrestinae	-	<i>Cyrestis thyodamas</i>	15,254	KF990125	This study
Nymphalinae	Kallimini	<i>Kallima inachus</i>	15183	JN857943	Qin et al., 2012
	Melitaeini	<i>Melitaea cinxia</i>	15,170	NC_018029	Unpublished
	Junoniini	<i>Junonia orithya</i>	15,214	KF199862	Shi et al., 2013b
	Junoniini	<i>Hypolimnias bolina</i>	15,260	KF990127	This study
Libytheinae	-	<i>Libythea celtis</i>	15,164	NC_016724	Hao et al., 2013
Danainae	Danaini	<i>Euploea mulciber</i>	15,166	NC_016720	Hao et al., 2013
	Danaini	<i>Danaus plexippus</i>	15,314	NC_021452	Servín and Martínez, 2013
	Danaini	<i>Danaus chrysippus</i>	15,236	KF690637	Gan et al., 2014a
	Danaini	<i>Ideopsis similes</i>	15,200	KJ476729	Gan et al., 2014b
	Danaini	<i>Parantica aglea</i>	15,210	KM018020	This study
outgroup	Theclini	<i>Coreana raphaelis</i>	15,314	NC_007976	Kim et al., 2006
	Theclini	<i>Protantigius superans</i>	15,248	NC_016016	Kim et al., 2011b
	-	<i>Spindasis takanonis</i>	15,349	NC_016018	Kim et al., 2011b

doi:10.1371/journal.pone.0124349.t001

phylogenetic analysis. The NJ analysis was performed in the MEGA5.0 [32], the MP and ML analyses were carried out using the PAUP\*4.0b10 [38], and the BI analysis was conducted using MrBayes 3.1.2 [39]. In the NJ analysis, nucleotide substitution model of the Kimura-2 parameter (K2P) was selected, and the bootstrap proportion values (BPs) with internal branch tests were obtained by 1,000 replicates to estimate the support levels for the nodes in the resultant topologies. The MP tree was reconstructed with branch swapping tree bisection-reconnection (TBR) heuristic search method, and the BPs were obtained after 1000 replicates by using 10 replicates of random stepwise additions of taxa. For ML and BI analyses, the two

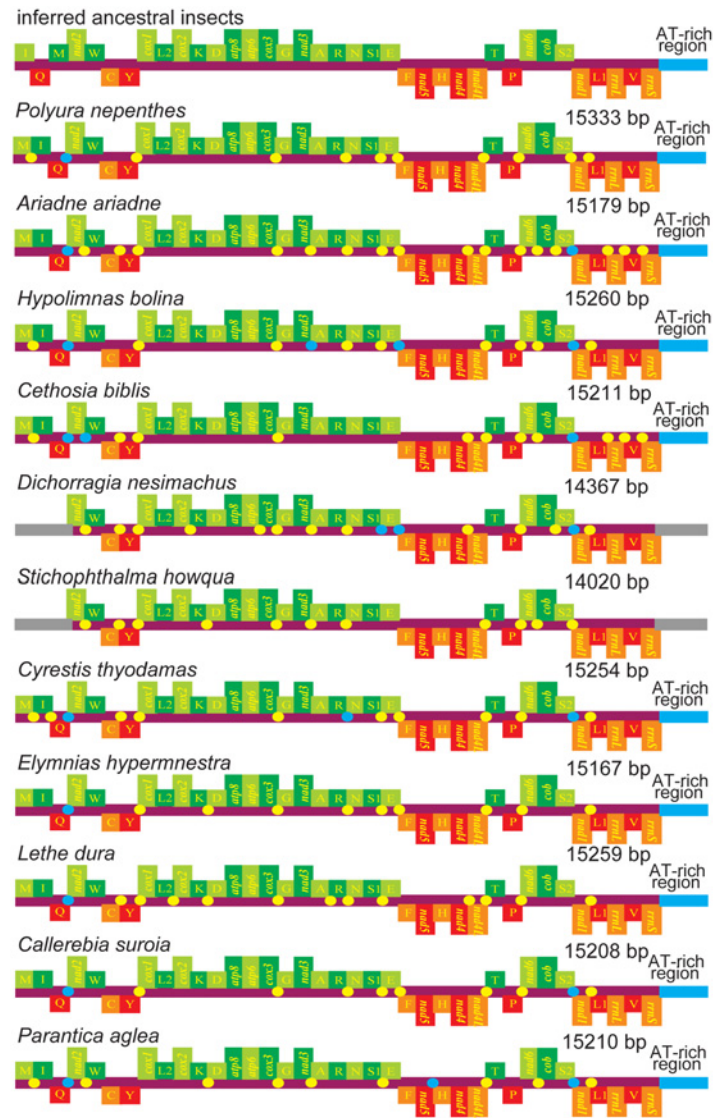
datasets were further partitioned by 4 strategies considering gene region and codon position. For D1 dataset, the partitioning strategies (PSs) were set as (1) 13 PCGs each as a single gene partition (D1-P13), (2) two partitions, including 1st and 2nd codon position of the PCGs (D1-P2). The PSs of D2 dataset were set as (3) 15 gene partitions (D2-P15) and (4) four partitions, including 1st and 2nd codon position of the PCGs and two rRNAs (D2-P4). The optimal substitution model of each partition was determined by Modeltest 3.7 [40], using the corrected Akaike information criterion (AICc). For ML analysis, the general time reversible model with invariant sites and among-site variation (GTR+I+G) was selected as the best fit model for each partition, and the BPs of the tree were also evaluated via the bootstrap test with 1000 iterations. The BI analysis was conducted with the same best fit substitution model used as the one selected for the ML 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 datasets, with each sampling of 100 generations, when the convergence of MCMC chains was achieved ( $<0.01$ ), the first 25% of the sampled trees were discarded as samples of burn-in. The confidence values of the BI tree are presented as the Bayesian posterior probabilities (BPP). All the tree topologies were evaluated using the Approximately Unbiased (AU) [41], Shimodaira-Hasegawa (SH) [42], Kishino-Hasegawa (KH) [43], weighted SH (WSH) [42] and weighted KH (WKH) [43] methods by the Consel v0.2 software [44].

## Results and Discussion

### Genome structure and organization

In this study, nine complete mitogenome sequences from *P. nepenthes* (15,333 bp), *A. ariadne* (15,179 bp), *H. bolina* (15,260 bp), *C. biblis* (15,211 bp), *E. hypermnestra* (15,167 bp), *L. dura* (15,259 bp), *C. suroia* (15,208 bp), *C. thyodamas* (15,254 bp), *P. aglea* (15,210 bp) and two nearly complete mitogenomes from *D. nesimachus* (14,367 bp) and *S. howqua* (14,020 bp) were determined, representing nine subfamilies of Nymphalidae (Tables 1 and 2). All genes identified in the eleven mitochondrial genomes are typical insect mitochondrial genes with normal gene sizes [45]. In all, 37 genes (13 PCGs, 22 tRNAs, 2 rRNAs) and an A+T-rich region were identified in the nine completely sequenced mitogenomes, in which, the gene order was identical to that of the other nymphalid mitogenomes sequenced to date, but different from the gene order in inferred ancestral insects [45] (Fig 1). The regions that we failed to sequence in other two mitogenomes were A+T-rich region and gene cluster *trnM*—*trnI*—*trnQ*, which are usually located in or around *rrnS* and *nad2* (Fig 1), where extremely high A+T content and stable stem-loop structures may have disrupted PCR and sequencing reactions, a common problem in sequencing insect mitochondrial genomes [23, 46].

All analyzed nymphalid mitogenomes are consistently AT biased, with values from 79.1% in *Eumenis autonoe* (subfamily Satyrinae) to 81.9% in *Parathyma sulphitia* (subfamily Limenitidinae), averaging at 80.5% (Table 2). However, these values fall within the range of the A+T content for other Lepidoptera species, from 77.8% in *Ochrogaster lunifer* [47] to 82.7% in *C. raphaelis* [48]. The nucleotide skewness statistics for all nymphalid mitogenomes indicate slight A or T skews with AT-skew values ranging from -0.068 in *C. biblis* (subfamily Heliconiinae) to 0.019 in *H. bolina* (subfamily Nymphalinae), and moderate C skews with GC-skew values ranging from -0.274 in *S. howqua* (subfamily Morphinae) to -0.178 in *P. Sulpitia* (subfamily Limenitidinae) (S1 Fig). A similar trend has been observed in other lepidopterans, the AT-skew of which ranges from -0.047 (*C. raphaelis*) to 0.059 (*Bombyx mori*) and the GC-skew from -0.318 (*O. lunifer*) to -0.158 (*C. raphaelis*) [47, 48].



**Fig 1. Gene arrangement of eleven nymphalid mitochondrial genomes sequenced in this study.** Abbreviations for the genes are as follows: *cox1*, *cox2*, and *cox3* refer to the cytochrome oxidase subunits, *cob* refers to cytochrome b, and *nad1*–*nad6* refer to NADH dehydrogenase components, *rml* and *rmS* refer to ribosomal RNAs. Transfer RNA genes are denoted by one letter symbols according to the IPUC-IUB one-letter amino acid codes. L1, L2, S1, S2 denote tRNA<sup>Leu(CUN)</sup>, tRNA<sup>Leu(UUR)</sup>, tRNA<sup>Ser(AGN)</sup>, tRNA<sup>Ser(UCN)</sup>, respectively. Genes coded on the majority strand are light/dark-green. Genes coded on the minority strand are red or orange. Alternation of colors was applied for distinction. The non-coding regions are presented as cyan/yellow dots. The unknown portions of partial mtDNAs are gray.

doi:10.1371/journal.pone.0124349.g001

### Protein coding genes

Like those of other determined lepidopteran mitogenomes, nine PCGs (*nad2*, *cox1*, *cox2*, *atp8*, *atp6*, *cox3*, *nad3*, *nad6*, and *cob*) of the eleven mitogenomes are coded on the majority strand, while the other four genes (*nad1*, *nad4*, *nad4L*, *nad5*) coded on the minority strand. All PCGs are initiated by typical ATN, with the exception of the *cox1* which uses the unusual CGA(R) as observed in most other sequenced lepidopterans [35, 47, 49–54]; all PCGs are terminated with TAA/TAG, or with truncated codons TA or T which are presumed to be completed via

**Table 2. Characterization of the nymphalid mitogenomes used in this study.**

Species	Whole genome		PCG <sup>a</sup>		<i>rrnL</i>		<i>rrnS</i>		AT-rich region	
	Size (bp)	A+T (%)	Size (bp) <sup>b</sup>	A+T (%)	Size (bp)	A+T (%)	Size (bp)	A+T (%)	Size (bp)	A+T (%)
<i>Apatura ilia</i>	15,242	80.5	11,133	78.9	1,333	84.0	776	84.9	403	92.5
<i>Apatura metis</i>	15,236	80.5	11,109	78.9	1,333	84.5	779	84.8	394	92.9
<i>Sasakia charonda</i>	15,244	79.9	11,108	78.2	1,323	84.4	775	85.0	380	91.8
<i>Sasakia funebris</i>	15,233	81.2	11,181	80.0	1,333	84.6	772	85.5	370	93.0
<i>Timelaea maculate</i>	15,178	81.1	11,145	79.6	1,332	85.0	777	85.9	382	93.2
<i>Ariadne ariadne</i>	15,179	80.1	11,163	78.8	1,320	83.9	799	84.1	319	92.8
<i>Cyrestis thyodamas</i>	15,254	81.1	11,181	79.7	1,328	84.6	780	85.1	380	91.6
<i>Kallima inachus</i>	15,183	80.3	11,163	79.2	1,335	82.7	774	85.1	376	92.0
<i>Melitaea cinxia</i>	15,170	80.0	11,154	78.6	1,336	84.6	772	84.6	338	92.9
<i>Hypolimnas bolina</i>	15,260	79.7	11,154	78.1	1,333	83.3	776	84.9	356	93.3
<i>Junonia orithya</i>	15,214	80.4	11,154	79.2	1,326	82.7	775	84.9	331	94.9
<i>Dichorragia nesimachus</i> <sup>▲</sup>	14,367	81.1	11,001	80.3	1,369	85.0	621	83.3	-	-
<i>Argyreus hyperbius</i>	15,156	80.8	11,154	79.5	1,330	84.5	778	85.2	349	95.4
<i>Acraea issoria</i>	15,245	79.7	11,151	78.0	1,331	83.9	788	83.7	430	96.0
<i>Fabriciana nerippe</i>	15,140	80.9	11,157	79.6	1,321	84.4	773	84.9	329	95.4
<i>Issoria lathonia</i>	15,172	81.1	11,154	79.9	1,319	84.4	771	85.1	361	96.1
<i>Heliconius melpomene</i>	15,328	81.7	11,106	80.3	1,364	85.7	779	85.1	268	95.9
<i>Cethosia biblis</i>	15,211	79.8	11,163	78.4	1,316	83.7	774	85.5	334	93.4
<i>Parathyma sulphitia</i>	15,268	81.9	11,187	80.6	1,319	84.7	779	85.7	349	94.6
<i>Libythea celtis</i>	15,164	81.2	11,166	80.0	1,335	84.7	775	85.4	328	96.3
<i>Eumenis autonoe</i>	15,489	79.1	11,184	76.8	1,335	83.7	775	85.3	678	94.5
<i>Melanitis leda</i>	15,122	79.8	11,163	78.4	1,332	84.0	774	85.0	314	89.5
<i>Elymnias hypermnestra</i>	15,167	80.4	11,160	79.1	1,331	84.7	768	85.0	405	90.6
<i>Lethe dura</i>	15,259	79.3	11,181	77.5	1,341	83.8	775	85.7	409	92.2
<i>Callerebia suroia</i>	15,208	79.5	11,121	77.9	1,345	84.3	777	85.7	393	94.2
<i>Stichophthalma howqua</i> <sup>▲</sup>	14,020	78.5	10,848	77.0	1,332	84.6	600	84.2	-	-
<i>Polyura nepenthes</i>	15,333	80.8	11,214	79.4	1,389	84.9	772	84.5	437	89.5
<i>Calinaga davidis</i>	15,267	80.4	11,211	78.8	1,337	83.8	773	85.9	389	92.0
<i>Euploea mulciber</i>	15,166	81.5	11,139	80.2	1,314	84.6	776	85.3	399	93.5
<i>Danaus plexippus</i>	15,314	81.3	11,148	79.9	1,340	84.4	774	86.1	469	94.5
<i>Danaus chrysippus</i>	15,236	80.3	11,145	78.7	1,314	83.8	779	84.7	418	95.0
<i>Ideopsis similes</i>	15,200	81.6	11,121	80.2	1,315	84.7	776	85.9	404	95.7
<i>Parantica aglea</i>	15,210	79.6	11,121	77.8	1,309	84.1	775	84.9	443	93.2

<sup>▲</sup> Partial mitogenome lacking in the A+T-rich region, *trnM*, *trnI* and *trnQ*, partial *nad2* and *rrnS* sequence. Bar (-) indicates lack of sequence information for the A + T region in the genome.

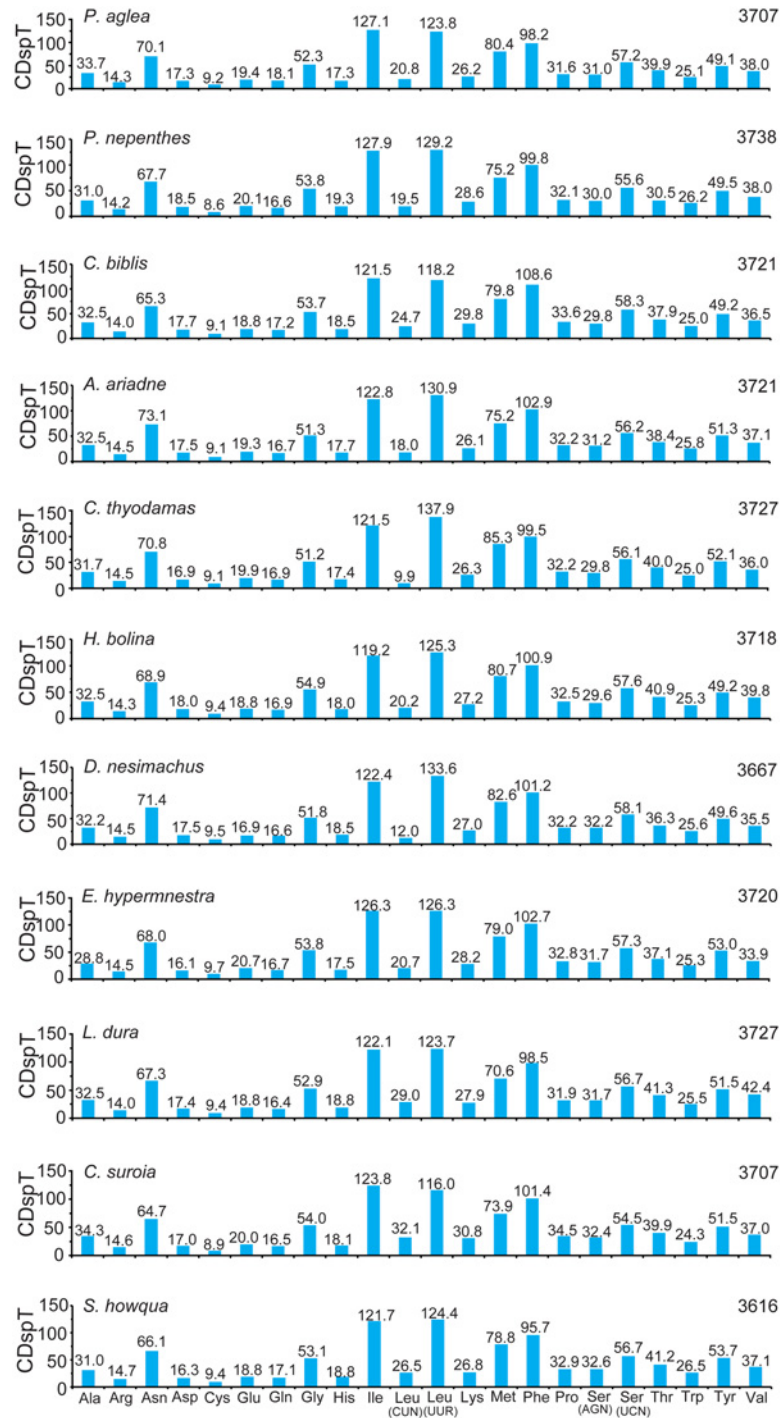
<sup>a</sup> Protein coding genes.

<sup>b</sup> Termination codons were excluded in total size count.

doi:10.1371/journal.pone.0124349.t002

post-transcriptional polyadenylation [55, 56] (S2 Table). The A+T contents of the PCGs, excluding stop codons, were similar to other nymphalid mitogenomes, and fall within the range from 76.8% in *E. autonoe* (subfamily Satyrinae) to 80.6% in *P. sulphitia* (subfamily Limenitidinae), with the average value equal to 79.0%. In all newly sequenced mitogenomes, the third codon positions have a considerably higher A+T content than the first and second positions (S2 Fig) as seen in other sequenced nymphalids.

The abundance of codon families and Relative Synonymous Codon Usage (RSCU) in 13 PCGs were investigated for the eleven newly determined mitogenomes, as shown in Figs 2 and 3. All stop codons, complete and incomplete, were excluded from the analysis to avoid biases due to incomplete stop codons. Total number of non-stop codons (CDs) in nine newly



**Fig 2. Codon distribution in the eleven newly sequenced nymphalid mitogenomes.** Numbers to the left refer to the total number of codons. CDspT, codons per thousand codons. Codon families are provided on the x axis.

doi:10.1371/journal.pone.0124349.g002





**Fig 3. Relative synonymous codon usage (RSCU) in the eleven newly sequenced nymphalid mitogenomes.** Codon families are given on the x axis. Codons that are not present in the genome are indicated in purple.

doi:10.1371/journal.pone.0124349.g003

sequenced complete mitogenomes is very similar to that in other available nymphalid mitogenomes, ranging from 3,695 of *S. charonda* (subfamily Apaturinae) to 3,738 of *Polyura nepenthes* (subfamily Charaxinae) (S3 Fig). Although the numbers of CDs in two nearly complete mitogenomes are less than that in other complete mitogenomes, the codon families exhibit a very

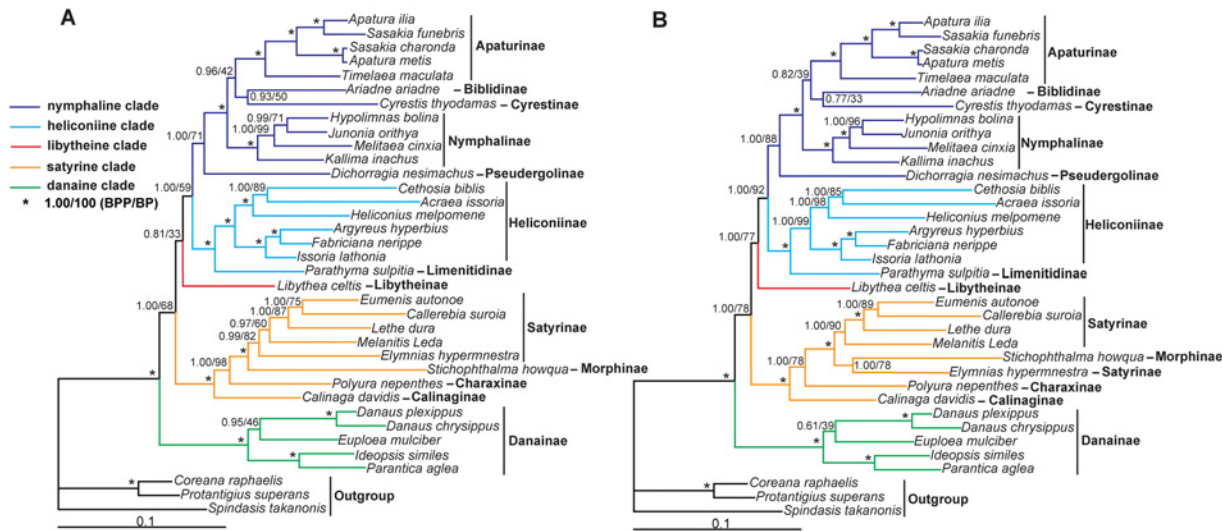
similar behavior among all newly sequenced mitogenomes. The five most frequently used codon families (*Leu* (*UUR*), *Ile*, *Phe*, *Met* and *Asn*), each with at least 60 CDs per thousand CDs, were two-fold degenerate in codon usage and were rich in A and T in every mitogenome. The three codon families with at least 100 CDs per thousand (*Leu* (*UUR*), *Ile* and *Phe*) were found in the *C. biblis* (subfamily Heliconiinae), *A. ariadne* (subfamily Biblidinae), *H. bolina* (subfamily Nymphalinae), *D. nesimachus* (subfamily Pseudergolinae), *E. hypermnestra* and *C. suroia* (subfamily Satyrinae) mitogenomes, whereas only *Leu* (*UUR*) and *Ile* units were found in remaining five newly sequenced mitogenomes. However, the rarest used codon family is Cys, which displayed no more than 10 CDs per thousand in eleven mitogenomes (Fig 2). Additionally, the RSCU values of NNU and NNA are greater than 1 (excluding the NNA for *Leu* (*UUR*)), indicating a strong A+T-bias in their third codon position, whereas the lost codons usually belong to G+C-rich codons (Fig 3). The codon usage bias may be positively correlated with the AT bias of the third codon position for the insect mitogenomes [47, 49, 52, 57–60].

### Transfer RNA and ribosomal RNA genes

All tRNAs of the eleven nymphalid species harbor the typical cloverleaf structures commonly found in insects, except for *trnSI*(*AGN*) whose dihydrouridine (DHU) arm is replaced by a simple loop (maps of secondary structure are not shown here). Two rRNA (*rrnL* and *rrnS*) genes are located between *trnL1* (*CUN*) and *trnV*, and between *trnV* and the A+T-rich region, respectively (Fig 1). Of these two rRNAs, the larger one (*rrnL* gene) has attracted much more attention in the studies of systematic evolution and classification [61]. Here, we first compared the entire secondary structure of all the nymphalid *rrnL* genes for more phylogenetically informative signals. The results showed that all the *rrnL* genes harbored six domains (domain III is absent in Arthropoda) with 49 helices, which is broadly congruent with those proposed for other insects [46, 47]. *D. nesimachus* was shown in S4 Fig as an example with the remaining species not listed. Among this nymphalid species, some of the *rrnL* highly variable regions, such as the stem-loops of H837/D13, 14 in domain II and H2077/G3 in domain V were also shown (S5 and S6 Figs). The structural characteristics of H837/D13, 14 and H2077/G3, suggest that the *A. ariadne* (subfamily Biblidinae) is closer with *C. thyodamas* (subfamily Cyrestinae) than with *D. nesimachus* (subfamily Pseudergolinae), the *S. howqua* (subfamily Morphinae) is most likely to be related to Satyrinae species, and so on.

### Non-coding regions

The non-coding (NC) regions of eleven newly sequenced mitogenomes were illustrated in Fig 1. There were three distinct large NC regions with at least 15 bp in all complete mitogenomes. The first large NC region located between *trnQ* and *nad2*, appeared to be common in the nine mitogenomes (including *P. nepenthes*, *A. ariadne*, *H. bolina*, *C. biblis*, *E. hypermnestra*, *L. dura*, *C. suroia*, *C. thyodamas*, *P. aglea*), and this region was also detected in other nymphalid mitogenomes, varying in length from 40 bp in *E. hypermnestra* (subfamily Satyrinae) to 87 bp in *S. charonda* and *Sasakia funebris* (subfamily Apaturinae). The second large NC region inserted between *trnS2* and *nad1*, was present in *A. ariadne* (subfamily Biblidinae), *H. bolina* (subfamily Nymphalinae), *C. biblis* (subfamily Heliconiinae), *C. thyodamas* (subfamily Cyrestinae), *C. suroia* (subfamily Satyrinae) and *P. aglea* (subfamily Danaina) mitogenomes, whereas this region was absent in *E. hypermnestra* and *L. dura* (subfamily Satyrinae) mitogenomes or showed no more than 15 bp in *P. nepenthes* (subfamily Charaxinae) mitogenome. The third large NC region located between *rrnS* and *trnM*, coincided with the A+T-rich region, also called the CR. The CR regions have no conspicuous macro-repeat units (>50 bp), which are occasionally found in other lepidopteran insects [62–65]. Nonetheless, several conserved structures



**Fig 4. Phylogenetic relationships among 33 nymphalid species.** (A) Bayesian Inference and maximum likelihood phylogram obtained with the D1 dataset, which is divided into 13 partitions (D1-P13). (B) Bayesian Inference and maximum likelihood phylogram obtained with the D2 dataset, which is divided into 15 partitions (D2-P14). Numbers on each node correspond to the posterior probability values of the BI analysis (left) and the ML bootstrap percentage values for 1 000 replicates of ML analysis (right).

doi:10.1371/journal.pone.0124349.g004

characteristic of lepidopterans are observed, such as the poly-T stretch preceded by the ATAGA motif neighboring the *rrnS* gene, and the microsatellite-like elements (TA)<sub>n</sub> (n = 7–9) preceded by the ATTTA motif. Although the CR regions were not sequenced in two incomplete mitogenomes (*D. nesimachus* and *S. howqua*), a few long NC regions (>15 bp in length), including the region located between *trnS2* and *nad1*, were also observed in *D. nesimachus* (subfamily Pseudergolinae) mitogenome, however, the similar region was not found in *S. howqua* (subfamily Morphinae) mitogenome.

### Phylogenetic analysis

The phylogenetic reconstructions conducted in this study using multiple methods have produced generally similar topology for most of the major nymphalid groups. The phylogenetic relationships inferred by BI and ML methods were highly congruent and strongly supported (BPP >0.95 and BP >85% for most nodes), across the partitions for each of the two datasets (Fig 4). The topology tests show that the BI trees from PS1 (BI-P13 tree) (Fig 4A), PS2 (BI-P2 tree) (not shown), PS3 (BI-P15 tree) (Fig 4B) and PS4 (BI-P4 tree) (not shown) are the best trees for the D1 and D2 dataset (Table 3). All the trees reveal that the Nymphalidae include five major clades, namely, the nymphaline clade (including five subfamilies: Apaturinae, Biblidinae, Cyrestinae, Nymphalinae and Pseudergolinae); heliconiine clade (including two subfamilies: Heliconiinae and Limenitidinae); satyrine clade (including four subfamilies: Satyrinae, Morphinae, Charaxinae and Calinaginae); danaine clade (including single subfamily: Danainae); libytheine clade (including one subfamily: Libytheinae). These results are consistent with those obtained from previous molecular studies [5, 14, 66], whereas remarkably different from the results of morphology-based phylogenetic studies, such as the cladistic analysis by Freitas and Brown [15] based on 234 characters of different developmental stages. In order to further evaluate the probable congruence between molecular and morphological evolutions, we carefully examined all the morphological characters from Freitas and Brown [15] and others and tried to map some of them on the mitogenomic phylogenetic trees robustly obtained in this study

**Table 3. Topological tests for two datasets with partitions.**

datasets	trees	statistical tests*					
		obs	AU	KH	SH	WKH	WSH
D1	BI-P13/P2	0.0	0.821	0.000	0.838	0.000	0.838
	ML-P13/P2	0.0	0.454	0.318	0.763	0.318	0.700
	NJ	37.7	0.097	0.085	0.124	0.085	0.138
	MP	38.9	0.068	0.055	0.113	0.055	0.099
D2	BI-P15/P4	0.0	0.629	0.522	0.767	0.522	0.770
	ML-P15/P4	0.0	0.552	0.478	0.651	0.478	0.800
	MP	0.5	0.551	0.477	0.651	0.041	0.525
	NJ	15.4	0.197	0.198	0.278	0.198	0.257

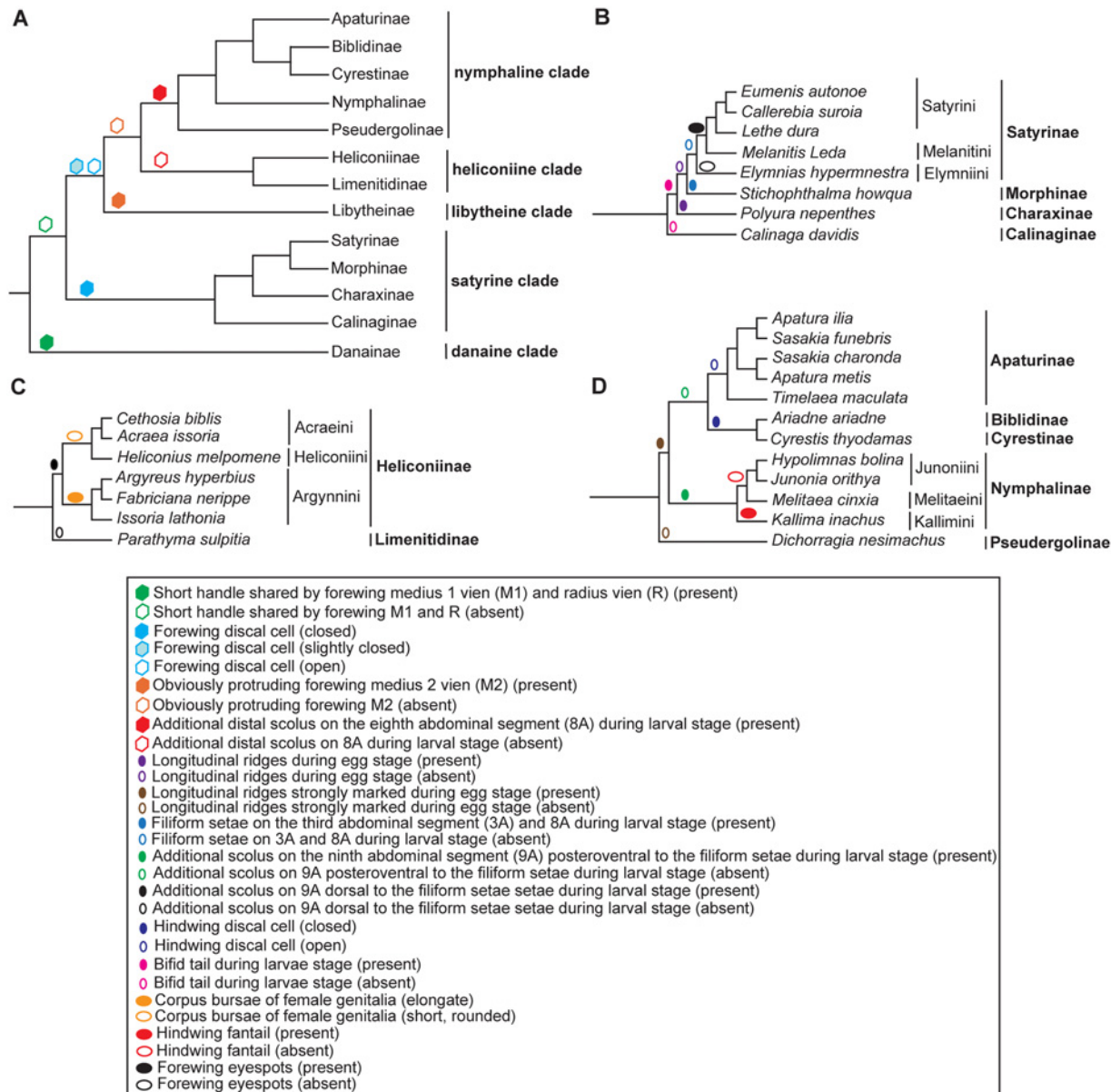
\* obs: the log-likelihood difference  $\delta_x$  to the best tree; AU = Approximately Unbiased; KH = Kishino-Hasegawa; SH = Shimodaira-Hasegawa; WKH = weighted KH; WSH = weighted SH.

doi:10.1371/journal.pone.0124349.t003

(Fig 5), indicating that some of the selected morphological characters (developmental and adult) are somewhat compatible with the molecular phylogeny. For example, the danaine as the basal group in Nymphalidae is supported by a unique morphological plesiomorphy, the connected medius 1 vein (M1) and radius vein (R), whereas the remaining nymphalid groups lack the morphological character (Fig 5A, green hexagon); the satyrine is characterized by a closed discal cell of the forewing, whereas its sister groups (libytheine + heliconiine + nymphaline) have a slightly closed or open discal cell of the forewing (Fig 5A, cyan hexagon).

The danuids (danaine clade), formerly classified as a family under the order Lepidoptera, are now treated as a subfamily of Nymphalidae. Our study indicates that Danainae (represented by *Euploea mulciber*, *Danaus plexippus*, *Danaus chrysippus*, *Ideopsis similes*, *P. aglea*) is sister to the remaining nymphalids including the libytheines with strong supports (BPP = 1.00 in two BI trees and BP = 100% in two ML trees, respectively), same as results from other recent analyses [54, 66–68]. Additionally, the analysis of mitochondrial *rrnL* secondary structures by Hao et al. [69] showed that the *rrnL* morphological characteristics of the danaid species were markedly different from other nymphalid groups. Morphologically, the danuids possess some of their unique features, such as the connected M1 vein and R vein in the forewing, which are markedly distinct from other nymphalid butterflies.

The satyrine clade includes Satyrinae, Morphinae, Charaxinae and Calinaginae, compatible with results of some previous studies based on morphological [15], molecular [10, 14, 70], or combined data [5, 18]. The Calinaginae (represented by *Calinaga davidis*) is the basal group of entire satyrine clade, instead of the sister relationship to Charaxinae as previously stated [14]. The Charaxinae (represented by *P. nepenthes*) is the sister to the grouping of Satyrinae plus Morphinae (Fig 3), consistent with those of multiple-gene based study [20] and most comprehensive taxa-sampling studies [5]. The association of Charaxinae with Satyrinae plus Morphinae is also supported by the morphological features, for example, all groups of satyrine clade, excluding the subfamily Calinaginae, have a bifid tail during their larval stages (Fig 5B). However, the phylogenetic position of Morphinae was not stable among trees from different dataset (Fig 4). Morphologically, the Morphinae has some distinct characteristics, such as the filiform setae on the third abdominal segment (3A) and the eighth abdominal segment (8A) during the larval stage [15] (Fig 5B). In addition, the Amathusiini and Elymniini occupy the same host plant (Arecaceae) during the larval stages, implying a close relationship. It has been proposed that Morphinae (Morphini, Brassolini and Amathusiini) should be grouped within Satyrinae [5, 20, 70–72], though other studies suggested two independent subfamilies [12, 14].



**Fig 5. Morphological character distribution mapped on the mitogenomic tree of Nymphalidae and internal clades (All the morphological characters selected are taken after Freitas and Brown, 2004).** (A) the topology of Nymphalidae clades; (B) the topology of satyrine clade; (C) the topology of heliconiine clade; (D) the topology of nymphaline clade.

doi:10.1371/journal.pone.0124349.g005

The phylogenetic status of libytheines (libytheine clade) within nymphalids has long stood as a controversial issue. They were traditionally treated as a separate family owing to their unusual morphological features such as their exceptionally long labial palpus and fully developed forelegs in female [73–75]. However, some scholars proposed that libytheines should be included in Nymphalidae as a subfamily for the presence of longitudinal ridges on the antenna shared with Nymphalidae [12, 76], sister to all other Nymphalidae because of the lack of apomorphic features such as the simple female foreleg [7, 8, 15]. Evidences from host plant use and geographic distribution together suggest that this group is sister to the remaining Nymphalidae, as a basal subfamily [9, 77, 78], and some recent studies also suggested Libytheinae or the

grouping of Libytheinae plus Danainae was the basal lineage of the Nymphalidae in molecular or molecular plus morphological view [5, 10, 14, 20, 79]. In this study, *L. celtis* (subfamily Libytheinae) was revealed as the sister of the grouping of nymphaline plus heliconiine clades in all phylogenetic trees with strong supports, except the ML-P13 tree (33% BP value at the node) (Fig 4A). This result was concordant with previous mitogenomic phylogenies [67, 68]. Thus, the phylogenetic position of Libytheinae remains uncertain.

The heliconiine clade comprises representatives of the subfamilies Heliconiinae and Limenitidinae. In this study, though our taxon sampling is limited, tribal-level relationships within subfamily Heliconiinae are well supported (Fig 4). The tribe Argynnini (represented by *Argyreus hyperbius*, *Fabriciana nerippe*, *Issoria lathonia*) is sister to the group composed of Acraeini (represented by *Acraea issoria*, *C. biblis*) + Heliconiini (represented by *Heliconius melpomene*) (BPP = 1.00 in two BI trees, BP = 100%, 99% in ML-P13 and ML-P15 trees respectively). This outcome is concordant with those reported earlier based on molecular data [5, 10, 14, 21, 80], but inconsistent with two morphologically-based studies [15, 81], in which Acraeini emerged as sister to the grouping of (Heliconiini + (Vagrantini + Argynnini)).

The final major clade, the nymphaline clade, contains representatives of the subfamilies Apaturinae, Biblidinae, Cyrestinae, Nymphalinae and Pseudergolinae. However, their internal relationships, especially regarding the Pseudergolinae, Biblidinae and Cyrestinae, are still controversial [5, 14, 20, 21]. In this study, the Pseudergolinae (represented by *D. nesimachus*) is placed as sister to other groups of the nymphaline clade with strong or moderate support values (BPP = 1.0 in two BI trees, BP = 71%, 88% in ML-P13 and ML-P15 trees) (Fig 4), in concordance with the result of Wahlberg et al. [5]. The *C. thyodamas* (subfamily Cyrestinae) is the sister of *A. ariadne* (subfamily Biblidinae) with strong or moderate support (BPP = 0.93 in BI-P13 and BPP = 0.77 in BI-P15 tree), but weakly supported in ML-P13 tree (BP = 50%) and ML-P15 tree (BP = 33%) (Fig 4). This relationship is consistent with that by Zhang et al. [21], but contradicting that of Wahlberg et al. [5] and Wahlberg and Wheat [20], which suggested that the Cyrestinae was sister to Nymphalinae. Morphologically, the Cyrestinae and Biblidinae are similar, as they share the short forewing discal cell, which is not found in other nymphalid groups. As for the Apaturinae, in this study, their five representatives (*Apatura ilia*, *Apatura metis*, *S. charonda*, *S. funebris*, *Timelaea maculata*) formed a strongly supported monophyletic group with the Biblidinae plus Cyrestinae as the sister group, albeit with weak support in ML-P13 tree (BP = 42%) and ML-P15 tree (BP = 39%). All of the three subfamilies were grouped with the Nymphalinae (represented by *Kallima inachus*, *Melitaea cinxia*, *H. bolina*, *Junonia orithya*) as its sister taxon, formerly revealed as the Heliconiinae by Freitas and Brown [15], Biblidinae + Apaturinae by Brower [10], Apaturinae by Wahlberg et al. [14] or Cyrestinae by Wahlberg et al. [5]. We found that the internal relationships of the nymphaline clade established in this study are consistent with their morphological pattern; e.g., all groups of nymphaline clade, excluding the basal group (Pseudergolinae), have a strongly marked longitudinal ridges during the egg stage; the Nymphalinae is characterized by an additional scolus on the ninth abdominal segment (9A) posteroventral to the filiform setae during the larval stage, which is absent in its sister groups ((Biblidinae + Cyrestinae) + Apaturinae) [15] (Fig 5D).

## Supporting Information

**S1 Fig. Scatter plot of AT- and GC-skews in the nymphalid subfamilies.** Composition skewness was calculated according to the formulas (AT-skew =  $[A-T] / [A+T]$ ; GC-skew =  $[G-C] / [G+C]$ ). All the species that are represented are listed in Table 1. (TIF)

**S2 Fig. The variation of AT content in different codon position of the concatenated 13 PCGs in the Nymphalidae mitogenomes.** The 1st, 2nd and 3rd represent nucleotide codon position 1, 2 and 3 respectively. All the species that are represented are listed in [Table 1](#). (TIF)

**S3 Fig. Comparison of the total number of the amino acids of the mitogenomic PCGs across the nymphalid species.** All the species that are represented are listed in [Table 1](#). (TIF)

**S4 Fig. Predicted *rrnL* secondary structure in *D. nesimachus* mitogenome.** Roman numerals denote the conserved domain structure. Helices are shaded in grey. Tertiary structures are denoted by boxed bases joined by solid lines. Base-pairing is indicated as follows: Watson-Crick pairs by dashes, wobble GU pairs by red dots and other non-canonical pairs by green dots. (TIF)

**S5 Fig. The secondary structures of H837/D13, 14 in *rrnL* gene from all Nymphalidae species used in this study.** Symbols are as in [S4 Fig](#). (TIF)

**S6 Fig. The secondary structures of H2077/G3 in *rrnL* gene from all Nymphalidae species used in this study.** Symbols are as in [S4 Fig](#). (TIF)

**S1 Table. Materials and their sources in this study.** (PDF)

**S2 Table. Predicted initiation and termination codons distribution of the 13 mitochondrial PCGs among Nymphalidae species in this study.** (PDF)

## Acknowledgments

We thank Prof. Hua-Bin Zhao (College of Life Sciences, Wuhan University, Wuhan) for valuable comments on the manuscript.

## Author Contributions

Conceived and designed the experiments: JSH QY. Performed the experiments: QHS. Analyzed the data: QHS YYS YLW. Contributed reagents/materials/analysis tools: JSH. Wrote the paper: QHS YYS YLW JSH QY.

## References

1. DeVries PJ. The Butterflies of Costa Rica and Their Natural History. Princeton: Princeton University Press; 1987. PMID: [21238350](#)
2. Shields O. World numbers of butterflies. J Lepid Soc. 1989; 43: 178–183.
3. Heppner JB. Faunal regions and the diversity of Lepidoptera. Trop Lepid. 1991; 2: 1–85.
4. Chou I. Classification and Identification of Chinese Butterflies. Zhengzhou, China: Henan Scientific and Technological Publishing House; 1998. (in Chinese).
5. Wahlberg N, Leneveu J, Kodandaramaiah U, Peña C, Nylin S, Freitas AVL, et al. Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. Proc R Soc B. 2009; 276: 4295–4302. doi: [10.1098/rspb.2009.1303](#) PMID: [19793750](#)
6. Boggs CL, Watt WB, Ehrlich PR. Butterflies: evolution and ecology taking flight. Chicago, IL: University of Chicago Press; 2003.

7. Ehrlich PR. The comparative morphology, phylogeny, and higher classification of the butterflies (Lepidoptera: Papilionoidea). *Univ Kansas Sci Bull.* 1958; 39: 305–370.
8. de Jong R, Vane-Wright RI, Ackery PR. The higher classification of butterflies (Lepidoptera): problems and prospects. *Entomol Scand.* 1996; 27: 65–101.
9. Freitas AVL. Nymphalidae (Lepidoptera), filogenia com base em caracteres de imaturos, com experimentos de troca de plantas hospedeiras. Campinas, São Paulo: Universidade Estadual de Campinas; 1999.
10. Brower AVZ. Phylogenetic relationships among the Nymphalidae (Lepidoptera), inferred from partial sequences of the *wingless* gene. *Proc R Soc Lond B.* 2000; 267: 1201–1211. PMID: [10902686](#)
11. Ackery PR, de Jong R, Vane-Wright RI. The butterflies: Hedyloidea, Hesperoidea, and Papilionoidea. In: Kristensen NP, (Ed), *The butterflies: Hedyloidea, Hesperoidea, and Papilionoidea*. Berlin: de Gruyter; 1999. pp. 263–300.
12. Harvey DJ. Higher classification of the Nymphalidae, Appen-dix B. In: Nijhout HF, (Ed), *Higher Classification of the Nymphalidae, Appendix B*. Washington, DC: Smithsonian Institution Press; 1991. pp. 255–273.
13. Müller W. Südamerikanische Nymphalidenraupen: Versuch eines natürlichen Systems der Nymphaliden. *Zool Jahrb.* 1886; 1: 417–678.
14. Wahlberg N, Weingartner E, Nylin S. Towards a better understanding of the higher systematics of Nymphalidae (Lepidoptera: Papilionoidea). *Mol Phylogenet Evol.* 2003; 28: 473–484. PMID: [12927132](#)
15. Freitas AVL, Brown KSJ. Phylogeny of the Nymphalidae (Lepidoptera: Papilionoidea). *Syst Biol.* 2004; 53: 363–383. PMID: [15503668](#)
16. Wahlberg N, Zimmermann M. Pattern of phylogenetic relationships among members of the tribe Melitaeini (Lepidoptera: Nymphalidae) inferred from mtDNA sequences. *Cladistics.* 2000; 16: 347–363.
17. Vandewoestijne S, Baguette M, Brakefield PM, Saccheri IJ. Phylogeography of *Aglaia urticae* (Lepidoptera) based on DNA sequences of the mitochondrial COI gene and control region. *Mol Phylogenet Evol.* 2004; 31: 630–646. PMID: [15062799](#)
18. Wahlberg N, Braby MF, Brower AVZ, De Jong R, Lee MM, Nylin S, et al. Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. *Proc R Soc B.* 2005; 272: 1577–1586. PMID: [16048773](#)
19. Chen N, Zhu GP, Hao JS, Zhang XP, Su CY, Pan HC, et al. Molecular phylogenetic analysis of the main lineages of Nymphalidae (Lepidoptera, Rhopalocera) based on mitochondrial 16S rDNA sequences. *Acta Zool Sin.* 2007; 53: 106–115. (in Chinese)
20. Wahlberg N, Wheat CW. Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Syst Biol.* 2008; 57: 231–242. doi: [10.1080/10635150802033006](#) PMID: [18398768](#)
21. Zhang M, Zhong Y, Cao TW, Geng YP, Zhang Y, Jin K, et al. Phylogenetic relationship and morphological evolution in the subfamily Limenitidinae (Lepidoptera: Nymphalidae). *Prog Nat Sci.* 2008; 18: 1357–1364.
22. Footitt RG, Maw HEL, Von Dohlen CD, Hebert PDN. Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. *Mol Ecol Resour.* 2008; 8: 1189–1201. doi: [10.1111/j.1755-0998.2008.02297.x](#) PMID: [21586006](#)
23. Wei SJ, Min S, Sharkey MJ, Achterberg C, Chen XX. Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to holometabolous insects. *BMC Genomics.* 2010; 11: 371. doi: [10.1186/1471-2164-11-371](#) PMID: [20537196](#)
24. Simon C, Frati F, Bekenbach A, Crespi B, Liu H, Flook P. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am.* 1994; 87: 651–701.
25. Caterino MS, Sperling FAH. *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. *Mol Phylogene Evol.* 1999; 11: 122–137. PMID: [10082616](#)
26. Simmons RB, Weller SJ. Utility and evolution of cytochrome b in insects. *Mol Phylogenet Evol.* 2001; 20: 196–210. PMID: [11476629](#)
27. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The Clustal X windows interface: flexible strategies for multiple sequences alignment aided by quality analysis tools. *Nucl Acids Res.* 1997; 25: 4876–4882. PMID: [9396791](#)
28. Singh VK, Mangalam AK, Dwivedi S, Naik S. Primer premier: program for design of degenerate primers from a protein sequence. *Bio Techniques.* 1998; 24: 318–319.



29. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser.* 1999; 41: 95–98.
30. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl Acids Res.* 1997; 25: 955–964. PMID: [9023104](#)
31. Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucl Acids Res.* 1999; 27: 573–580. PMID: [9862982](#)
32. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 2011; 28: 2731–2739. doi: [10.1093/molbev/msr121](#) PMID: [21546353](#)
33. Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, et al. The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics.* 2002; 3:15.
34. Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Mol Biol.* 2006; 15: 657–686. PMID: [17069639](#)
35. Cameron SL, Whiting MF. 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.* 2008; 408: 112–123. PMID: [18065166](#)
36. Niehuis O, Naumann CM, Misof B. Identification of evolutionary conserved structural elements in the mt SSU rRNA of Zygaenoidea (Lepidoptera): a comparative sequence analysis. *Organisms Divers Evol.* 2006a; 6: 17–32.
37. Niehuis O, Yen SH, Naumann CM, Misof B. Higher phylogeny of zygaenid moths (Insecta: Lepidoptera) inferred from nuclear and mitochondrial sequence data and the evolution of larval cuticular cavities for chemical defence. *Mol Phylogenet Evol.* 2006b; 39: 812–829. PMID: [16483803](#)
38. Swofford DL. PAUP\*. Phylogenetic Analysis Using Parsimony (\* and other methods), v. 4.0b10. Sunderland, MA: Sinauer Associates; 2002.
39. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics.* 2001; 17: 754–755. PMID: [11524383](#)
40. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. *Bioinformatics.* 1998; 14: 817–818. PMID: [9918953](#)
41. Shimodaira H. An approximately unbiased test of phylogenetic tree selection. *Syst Biol.* 2002; 51: 492–508. PMID: [12079646](#)
42. Shimodaira H, Hasegawa M. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol.* 1999; 16: 1114–1116.
43. Kishino H, Hasegawa M. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data and the branching order in Hominoidea. *J Mol Evol.* 1989; 29: 170–179. PMID: [2509717](#)
44. Shimodaira H, Hasegawa M. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics.* 2001; 17: 1246–1247. PMID: [11751242](#)
45. Boore JL. Animal mitochondrial genomes. *Nucl Acids Res.* 1999; 27: 1767–1780. PMID: [10101183](#)
46. Feng X, Liu DF, Wang NX, Zhu CD, Jiang GF. The mitochondrial genome of the butterfly *Papilio xuthus* (Lepidoptera: Papilionidae) and related phylogenetic analyses. *Mol Biol Rep.* 2010; 37: 3877–3888. doi: [10.1007/s11033-010-0044-z](#) PMID: [20213506](#)
47. Salvato P, Simonato M, Battisti A, Negrisola E. The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). *BMC Genomics.* 2008; 9: 331. doi: [10.1186/1471-2164-9-331](#) PMID: [18627592](#)
48. Kim I, Lee EM, Seol KY, Yun EY, Lee YB, Hwang JS, et al. The mitochondrial genome of the Korean hair-streak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). *Insect Mol Biol.* 2006; 15: 217–225. PMID: [16640732](#)
49. Kim MI, Baek JY, Kim MJ, Jeong HC, Kim KG, Bae CH, et al. Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. *Mol Cells.* 2009; 28: 347–363. doi: [10.1007/s10059-009-0129-5](#) PMID: [19823774](#)
50. Jiang ST, Hong GY, Yu M, Li N, Yang Y, Liu YQ, et al. Characterization of the complete mitochondrial genome of the giant silkworm moth, *Eriogyna pyretorum* (Lepidoptera: Saturniidae). *Int J Biol Sci.* 2009; 5: 351–365. PMID: [19471586](#)
51. Kim MJ, Wan X, Kim KG, Hwang JS, Kim I. Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). *Afr J Biotechnol.* 2010; 9: 735–754.

52. Wang XC, Sun XY, Sun QQ, Zhang DX, Hu J, Yang Q, et al. The complete mitochondrial genome of the laced fritillary *Argyreus hyperbius* (Lepidoptera: Nymphalidae). *Zool Res.* 2011; 32: 465–475. doi: [10.3724/SP.J.1141.2011.05465](https://doi.org/10.3724/SP.J.1141.2011.05465) PMID: [22006797](https://pubmed.ncbi.nlm.nih.gov/22006797/)
53. Tian LL, Sun XY, Chen M, Gai YH, Hao JS, Yang Q. Complete mitochondrial genome of the Five-dot Sergeant *Parathyma sulphitia* (Nymphalidae: Limenitidinae) and its phylogenetic implications. *Zool Res.* 2012; 33: 133–143. doi: [10.3724/SP.J.1141.2012.02133](https://doi.org/10.3724/SP.J.1141.2012.02133) PMID: [22467387](https://pubmed.ncbi.nlm.nih.gov/22467387/)
54. Hao JS, Sun ME, Shi QH, Sun XY, Shao LL, Yang Q. Complete mitogenomes of *Euploea mulciber* (Nymphalidae: Danaeinae) and *Libythea celtis* (Nymphalidae: Libytheinae) and their phylogenetic implications. *ISRN Genomics.* 2013; 491636: 1–14.
55. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature.* 1981; 290: 457–465. PMID: [7219534](https://pubmed.ncbi.nlm.nih.gov/7219534/)
56. Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. *Nature.* 1981; 290: 470–474. PMID: [7219536](https://pubmed.ncbi.nlm.nih.gov/7219536/)
57. Liao F, Wang L, Wu S, Li YP, Zhao L, Huang GM, et al. The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *Int J Biol Sci.* 2010; 6: 172–186. PMID: [20376208](https://pubmed.ncbi.nlm.nih.gov/20376208/)
58. Hao JS, Sun QQ, Zhao HB, Sun XY, Gai YH, Yang Q. The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperidae: Pyrginae) and its phylogenetic implication. *Comp Funct Genomics.* 2012; 328049: 1–13. doi: [10.1155/2012/328049](https://doi.org/10.1155/2012/328049) PMID: [22577351](https://pubmed.ncbi.nlm.nih.gov/22577351/)
59. Cao YQ, Ma C, Chen JY, Yang DR. The complete mitochondrial genomes of two ghost moths, *Thitarodes renzhiensis* and *Thitarodes yunnanensis*: the ancestral gene arrangement in Lepidoptera. *BMC Genomics.* 2012; 13: 276. doi: [10.1186/1471-2164-13-276](https://doi.org/10.1186/1471-2164-13-276) PMID: [22726496](https://pubmed.ncbi.nlm.nih.gov/22726496/)
60. Chai HN, Du YZ, Zhai BP. Characterization of the complete mitochondrial genomes of *Cnaphalocrocis medinalis* and *Chilo suppressalis* (Lepidoptera: Pyralidae). *Int J Biol Sci.* 2012; 8: 561–579. doi: [10.7150/ijbs.3540](https://doi.org/10.7150/ijbs.3540) PMID: [22532789](https://pubmed.ncbi.nlm.nih.gov/22532789/)
61. Yukuhiro K, Sezutsu H, Itoh M, Shimizu K, Banno Y. Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkworm, *Bombyx mandarina*, and its close relative, the domesticated silkworm, *Bombyx mori*. *Mol Biol Evol.* 2002; 19: 1385–1389. PMID: [12140251](https://pubmed.ncbi.nlm.nih.gov/12140251/)
62. Liu YQ, Li YP, Pan MH, Dai FY, Zhu XW, Lu C, et al. The complete mitochondrial genome of the Chinese oak silkworm, *Antheraea pernyi* (Lepidoptera: Saturniidae). *Acta Bioch Bioph Sin.* 2008; 40: 693–703.
63. Yang L, Wei ZJ, Hong GY, Jiang ST, Wen LP. The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae). *Mol Biol Rep.* 2009; 36: 1441–1449. doi: [10.1007/s11033-008-9334-0](https://doi.org/10.1007/s11033-008-9334-0) PMID: [18696255](https://pubmed.ncbi.nlm.nih.gov/18696255/)
64. Pan MH, Yu QY, Xia YL, Dai FY, Liu YQ, Lu C, et al. Characterization of mitochondrial genome of Chinese wild mulberry silkworm, *Bombyx mandarina* (Lepidoptera: Bombycidae). *Sci China C (Life Sci).* 2008; 51: 693–701. doi: [10.1007/s11427-008-0097-6](https://doi.org/10.1007/s11427-008-0097-6) PMID: [18677597](https://pubmed.ncbi.nlm.nih.gov/18677597/)
65. Chen GZ, Li WJ, Xu LH, Jiang CL. Advanced in 16S rRNA secondary structure and its application in systematic classification. *J Microbiol.* 2005; 25: 54–57 (in Chinese).
66. Yang X, Xue D, Han H. The complete mitochondrial genome of *Biston panterinaria* (Lepidoptera: Geometridae), with phylogenetic utility of mitochondrial genome in the Lepidoptera. *Gene.* 2013; 515:349–358. doi: [10.1016/j.gene.2012.11.031](https://doi.org/10.1016/j.gene.2012.11.031) PMID: [23220020](https://pubmed.ncbi.nlm.nih.gov/23220020/)
67. Lu HF, Su TJ, Luo AR, Zhu CD, Wu CS. Characterization of the complete mitochondrion genome of diurnal moth *Amata emma* (Butler) (Lepidoptera: Erebidae) and its phylogenetic implications. *PLoS One.* 2013; 8: e72410. doi: [10.1371/journal.pone.0072410](https://doi.org/10.1371/journal.pone.0072410) PMID: [24069145](https://pubmed.ncbi.nlm.nih.gov/24069145/)
68. Wu LW, Lin LH, Lees DC, Hsu YF. Mitogenomic sequences effectively recover relationships within brush-footed butterflies (Lepidoptera: Nymphalidae). *BMC Genomics.* 2014; 15: 1–17. doi: [10.1186/1471-2164-15-1](https://doi.org/10.1186/1471-2164-15-1) PMID: [24382143](https://pubmed.ncbi.nlm.nih.gov/24382143/)
69. Hao JS, Su CY, Zhu GP, Chen N, Wu DX, Zhang XP. The molecular morphologies of mitochondrial 16S rRNA of the main butterfly lineages and their phylogenetic significances. *J Genet Mol Biol.* 2007; 18: 111–113.
70. Peña C, Nylin S, Freitas AVL. Higher level phylogeny of Satyrinae butterflies (Lepidoptera: Nymphalidae) based on DNA sequence data. *Mol Phylogenet Evol.* 2006; 40: 29–49. PMID: [16563805](https://pubmed.ncbi.nlm.nih.gov/16563805/)
71. Peña C, Wahlberg N. Prehistorical climate change increased diversification of a group of butterflies. *Biol Letters.* 2008; 4: 274–278.
72. Marín M, Peña C, Freitas AVL, Wahlberg N, Uribe SI. From the phylogeny of the Satyrinae butterflies to the systematics of Euptychiina (Lepidoptera: Nymphalidae): history, progress and prospects. *Neotrop Entomol.* 2011; 40: 1–13. PMID: [21437476](https://pubmed.ncbi.nlm.nih.gov/21437476/)

73. Pagenstecher A. Libytheidae. *Das Tierreich*. 1901; 14: 1–18.
74. Pagenstecher A. Lepidoptera Rhopalocera: Family Libytheidae. In: Wytzman P, (Ed.), *Fascicle. Bruxelles: Genera Insectorum*; 1902. pp. 1–4.
75. Pagenstecher A. Libytheidae. In: Aurivillius C, Wagner H, (Eds.), *Lepidopterorum Catalogus*. Berlin: W. Junk; 1911. 12 p.
76. Kristensen NP. Remarks on the family-level phylogeny of butterflies (Insecta, Lepidoptera, Rhopalocera). *J Zool Syst Evol Res*. 1976; 14: 25–33.
77. Ackery PR. Systematic and faunistic studies on butterflies. In: Vane-Wright RI, Ackery PR, (Eds.), *Systematic and Faunistic Studies on Butterflies*. Princeton, USA: Princeton University Press; 1984. pp. 9–21.
78. Vane-Wright RI. Evidence and identity in butterfly systematics. In: Boggs CL, Watt WB, and Ehrlich PR, (Eds.), *Butterflies: Ecology and evolution taking flight*. Chicago: University of Chicago Press; 2003. pp. 477–514.
79. Heikkilä M, Kaila L, Mutanen M, Peña C, Wahlberg N. Cretaceous origin and repeated Tertiary diversification of the redefined butterflies. *Proc R Soc B*. 2012; 279: 1093–1099. doi: [10.1098/rspb.2011.1430](https://doi.org/10.1098/rspb.2011.1430) PMID: [21920981](https://pubmed.ncbi.nlm.nih.gov/21920981/)
80. Silva-Brandão KL, Wahlberg N, Francini RB, Azeredo-Espin AML, Brown KS Jr, Paluch M, et al. Phylogenetic relationships of butterflies of the tribe Acraeini (Lepidoptera, Nymphalidae, Heliconiinae) and the evolution of host plant use. *Mol Phylogenet Evol*. 2008; 46: 515–531. doi: [10.1016/j.ympev.2007.11.024](https://doi.org/10.1016/j.ympev.2007.11.024) PMID: [18191590](https://pubmed.ncbi.nlm.nih.gov/18191590/)
81. Penz CM, Peggie D. Phylogenetic relationships among Heliconiinae genera based on morphology (Lepidoptera: Nymphalidae). *Syst Entomol*. 2003; 28: 451–479.