Research Article

Complete Mitogenomes of *Euploea mulciber* (Nymphalidae: Danainae) and *Libythea celtis* (Nymphalidae: Libytheinae) and Their Phylogenetic Implications

Jiasheng Hao,^{1,2} Ming'e Sun,¹ Qinghui Shi,¹ Xiaoyan Sun,² Lili Shao,¹ and Qun Yang²

¹ Laboratory of Molecular Evolution and Biodiversity, College of Life Sciences, Anhui Normal University, Wuhu, Anhui 241000, China

² State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing, Jiangsu 210008, China

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

Received 27 November 2012; Accepted 13 December 2012

Academic Editors: B. Antonio and S. N. Shchelkunov

Copyright © 2013 Jiasheng Hao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The complete mitochondrial genome sequences of the two butterfly species *Euploea mulciber* (Lepidoptera: Nymphalidae: Danainae) and *Libythea celtis* (Lepidoptera: Nymphalidae: Libytheinae) were determined in this study, comprising 15,166 bp and 15,164 bp, respectively. The orientation and the gene order of the two mitogenomes are identical to those of most of the other lepidopteran species. All protein-coding genes of *Euploea mulciber* and *Libythea celtis* mitogenomes start with a typical ATN codon with the exception of COI gene which uses CGA as its initial codon. All tRNA genes possess the typical cloverleaf secondary structure except for tRNA^{Ser} (AGN), which has a simple loop with the absence of the DHU stem. There are short microsatellite-like repeat regions, but no conspicuous macrorepeats scattered throughout the A + T-rich regions. Phylogenetic analysis among the available butterfly species suggests that *Libythea celtis* (Libytheinae) is closely related to *Calinaga davidis* (Calinaginae), indicating that the subfamily Libytheinae may not represent a basal lineage of the Nymphalidae as previously suggested, and that *Euploea mulciber* stands at the base of the nymphalid tree as a sister to all other nymphalids.

1. Introduction

The animal mitochondrial genome (mitogenome) is a circular molecule of 15–20 kb in length. It contains 37 conserved genes including 22 transfer RNA genes (tRNAs), 2 ribosomal RNA genes (rRNAs), and 13 protein-coding genes (PCGs) involved in electron transport and oxidative phosphorylation [1, 2]. It also contains an A + T-rich region which is the largest noncoding area involved in the initiation and regulation of replication and transcription [3]. Mitogenome studies are important for comparative and evolutionary genomics, phylogenetics, molecular evolution, and population genetics due to the genome's unique features, such as maternal inheritance, lack of extensive recombination, and accelerated nucleotide substitution rates [4, 5].

The Lepidoptera (butterflies and moths) is probably the largest insect order, containing over 165,000 described species. The systematics of the lepidopteran higher groups

has long been a matter of contention [6–8]. In China, Chou's taxonomic system [9] is widely adopted, in which the Chinese butterflies are split into 12 families and 32 subfamilies based on their morphological characteristics. However, another taxonomic system proposed by Wahlberg et al. (2005) [10], which is followed herein, has been commonly accepted by most butterfly researchers. In this system, the butterflies are classified into 2 superfamilies (Hesperioidea and Papilionoidea) and 5 families (Hesperiidae, Papilionidae, Pieridae, Lycaenidae, and Nymphalidae). Thus, the taxonomic levels and the phylogenetic relationships within and among butterfly lineages of some butterfly groups, such as the danainids and libytheines, remain controversial issues among different researchers.

Euploea mulciber is a member of the subfamily Danainae (formerly treated as a family). This medium-sized butterfly species is commonly called “striped blue crow” due to the male's bright blue shot forewing and the female's striped

hindwing with narrow white streaks; its main host plants, *Ficus* and *Nerium*, are poisonous, mimicked by other butterfly species such as *Elymnias malelas*, *Hypolimnas anomala*, and *Chilasa paradoxa*. *E. mulciber* is distributed in India, Nepal, Bhutan, Burma, Malaysia, and China [9].

Libythea celtis is a species of the subfamily Libytheinae (formerly treated as a family). This species is recorded in Southern Europe, Northern Africa, and Asia, and their larvae generally feed on leaves of *Celtis* (Celtidaceae). The Libytheinae include only about 10 species in the world, so far as known, 3 of which are in China. They can be easily recognized by their unusually long labial palpus, although there are some other nymphalids that have an equally long or longer labial palpus [8]; the libytheine threadlike labial palpus looks like the petiole of a dead leaf, offering camouflage when the butterfly is taking a rest [11].

Currently, data of lepidopteran mitochondrial genomes are quite limited [12–19]. Among these, 20 complete and 2 nearly complete mitochondrial sequences of butterflies have been determined. In this study, the complete mitogenomes of the two representative nymphalid species mentioned previously were determined and compared with other butterflies available, in order to provide more information for the phylogenetic studies of lepidopterans and further clarify their taxonomic status within the family Nymphalidae at the mitogenomic level.

2. Materials and Methods

2.1. Specimen Collection. Adult individuals of *Euploea mulciber* and *Libythea celtis* were collected in Dali, Yunnan province, and Huangshan, Anhui province, China, respectively. The fresh material was preserved in 100% ethanol and stored at -20°C before the DNA extraction.

2.2. DNA Extraction, PCR Amplification, and Sequencing. The whole genomic DNA was extracted from thoracic muscle tissue with the DNeasy Blood & Tissue kit (Qiagen) after Hao et al. (2005) [20]. Some universal PCR primers for short fragment amplifications of srRNA, COI, CytB, ND1, and COII genes were synthesized after Simon et al. (1994) [21], Caterino and Sperling (1999) [22], and Simmons and Weller (2001) [23]. The remaining primers for the two species were designed by the multiple sequence alignments of the complete mitogenomes of all the lepidopterans available, using ClustalX1.8 [24] and Primer Premier 5.0 softwares [25]. The entire mitogenomes of *Euploea mulciber* and *Libythea celtis* were amplified in five fragments (COI-COIII, COIII-ND4, ND4-CytB, and CytB-lrRNA, lrRNA-COI) and ten fragments (COI-COII, COII-COIII, COIII-ND5, ND5-ND4, ND4-CytB, CytB-ND1, ND1-lrRNA, lrRNA-srRNA, srRNA-ND2, and ND2-COI) using long PCR technique, which was performed using TaKaRa LA Taq polymerase with the following cycling parameters: 95°C for 5 minutes, 30 cycles of 95°C for 50 seconds, $47\text{--}51^{\circ}\text{C}$ for 50 seconds, 68°C for 2 minutes and 30 seconds, and a final extension step of 68°C for 10 minutes. The PCR products were detected via electrophoresis in 1.2% agarose gel, purified using the 3S

Spin PCR Product Purification Kit and sequenced directly with ABI-377 automatic DNA sequencer. All the amplified products were sequenced directly except the ND2-COI of *Libythea celtis*, which was sequenced after cloning. For each long PCR product, the full double-stranded sequence was determined by primer walking.

2.3. Sequence Analysis and Annotation. Raw sequence files were proofread and assembled with BioEdit version 7.0 [26]. PCGs and rRNAs were identified by sequence comparison using ClustalX1.8 software and the NCBI Internet BLAST search function [27]. tRNA gene analysis was conducted using tRNAscan-SE software v.1.21 [28], and the putative tRNAs, which were not found by tRNAscan-SE, were confirmed by sequence comparison of *Euploea mulciber* and *Libythea celtis* with other lepidopterans. Nucleotide composition and codon usage were calculated in MEGA 4.0 software [29]. The *Euploea mulciber* and *Libythea celtis* mitogenome sequence data have been deposited into the GenBank database under the accession numbers HQ378507 and HQ378508, respectively.

2.4. Phylogenetic Analysis. The multiple alignments of the concatenated 13 PCG nucleotide sequences of the 24 butterfly mitogenomes, including two newly determined and 22 available from GenBank (Table 1), were done by the ClustalX1.8 software and then manually proofread. The genetic distances were calculated based on selected sequence evolutionary models; meanwhile, the neighbor joining (NJ) [30] and Bayesian inference (BI) [31] phylogenetic trees were reconstructed based on the PCG sequences using one moth species, *Lymantria dispar* (GenBank accession no. FJ617240) as the outgroup. In the NJ analysis, the Kimura 2 parameter (K2P) evolutionary model [32] was selected, and the bootstrap values with internal branch tests were obtained by 1,000 replicates to estimate the support levels for the nodes in the resultant topologies using the MEGA 4.0 software. The Bayesian analyses were performed based on the PCG nucleotide sequences using MrBayes 3.1.2 [33, 34] under the model GTR+I+G chosen by the modeltest 3.06 software [35], and the Markov's chains were run simultaneously for 1,000,000 generations sampled every 100 generations; the first 2,500 trees were discarded as burn-in samples, and the remaining trees were used to generate a majority rule consensus tree, in which the percentage of samples recovering a particular clade represents its support measured as posterior probabilities.

3. Results and Discussion

3.1. Genome Organization and Structure. The mitogenomes of *Euploea mulciber* and *Libythea celtis* both consist of 13 PCGs, two ribosomal RNA genes for the small and large subunits (srRNA and lrRNA), and 22 tRNAs, and an A + T-rich region (see Tables 1 and 2 and Figure 1). Like many other insect mitogenomes, its major strand code for more genes (9PCGs and 14 tRNAs), whereas less genes were coded in the minor strand (4 PCGs, 8 tRNAs, and 2 rRNA genes).

TABLE 1: The available rhopaloceran mitogenomes (two new and 22 obtained from GenBank) and their characteristics.

Family/subfamily	Species	Whole genome				PGC ^b				1rRNA				srRNA				A + T-rich region	Gen of Bank accession no.	References
		Size	A + T		No. codons ^a	A + T	Size	A + T		Size	A + T	Size	A + T		Size	A + T				
			(bp)	(%)				(%)	(bp)				(%)	(bp)			(%)			
Papilionidae/Papilioninae	<i>Troides aeacus</i>	15,263	80.2	3745	79.1	1234	83.3	784	85.0	419	89.8	EU625344	Unpublished	[63]						
	<i>Papilio xuthus</i>	13,964	80.0	3583	78.9	1327	83.6	598	84.5	—	—	EF176224	[49]							
	<i>Aeghania maraho</i>	16,094	80.5	3718	79.2	1333	83.7	779	85.5	1270	95.0	FJ810212	[64]							
	<i>Teinopalpus aureus</i>	15,242	79.8	3720	78.3	1320	82.4	781	85.6	395	93.1	HM563681	[13]							
	<i>Parnassius bremeri</i>	15,389	81.3	3723	80.2	1344	83.9	773	85.1	504	93.6	FJ871125	[18]							
Papilionidae/Parnassiinae	<i>Sericius montela</i>	15,242	80.8	3691	80.2	1338	83.6	759	84.6	408	94.1	HQ259122	Unpublished	[19]						
	<i>Luehdorfia chinensis</i>	13,860	80.5	3551	79.6	1342	84.0	460	81.5	—	—	EU622524	[14]							
Hesperiidae/Pyrginae	<i>Ctenoptilum vasava</i>	15,468	80.5	3698	78.9	1343	84.1	774	86.4	429	88.1	JF713818	[39]							
Pieridae/Pierinae	<i>Artogeia melete</i>	15,140	79.8	3715	78.4	1319	83.4	777	85.5	351	89.2	NC_010568	[12]							
	<i>Pieris rapae</i>	15,157	79.7	3722	78.3	1320	84.0	764	85.0	393	91.6	HM156697	[65]							
Lycaenidae/Theclinae	<i>Coreana raphaelis</i>	15,314	82.7	3708	81.5	1330	85.3	777	85.8	375	94.1	DQ102703	[50]							
	<i>Protantigius superans</i>	15,248	81.7	3712	80.3	1331	85.1	739	85.6	361	93.6	HQ184265	[66]							
Lycaenidae/Aphnaeinae	<i>Spindasis takanonis</i>	15,349	82.4	3719	81.0	1333	85.6	777	84.7	371	94.6	HQ184266	[67]							
	<i>Apatura ilia</i>	15,242	80.5	3711	78.9	1333	84.0	776	84.9	403	92.5	JF439725	[68]							
Nymphalidae/Apaturinae	<i>Sasakia charonda</i>	15,244	79.9	3695	78.2	1323	84.4	775	85.0	380	91.8	NC_014224	Unpublished	[69]						
	<i>Sasakia charonda</i> *	15,222	79.9	3586	78.3	1311	84.2	775	85.0	380	91.8	NC_014223	Unpublished	[66]						
Nymphalidae/Nymphalinae	<i>Kallima inachus</i>	15,183	80.3	3721	79.2	1335	82.7	774	85.1	376	92.0	JN857943	[37]							
Nymphalidae/Heliconiinae	<i>Argyreus hyperbius</i>	15,156	80.8	3718	79.5	1330	84.5	778	85.2	349	95.4	JF439070	[15]							
	<i>Acraea issoria</i>	15,245	79.7	3717	78.0	1331	83.9	788	83.7	430	96.0	GQ376195	[17]							
Nymphalidae/Limenitidinae	<i>Parathyma sulphita</i>	15,268	81.9	3729	80.6	1319	84.7	779	85.7	349	94.6	JQ347260	[67]							
Nymphalidae/Calinaginae	<i>Calinaga davidis</i>	15,267	80.4	3737	78.9	1337	83.8	773	85.9	389	92.0	NC_015480	[68]							
Nymphalidae/Danainae	<i>Eumenis autonoe</i>	15,489	79.1	3728	76.8	1335	83.7	775	85.3	678	94.5	GQ868707	This study	This study						
Nymphalidae/Libytheinae	<i>Libythea celtis</i>	15,164	81.2	3722	80.0	1335	84.7	775	85.4	328	96.3	HQ378508	[69]							
Nymphalidae/Danainae	<i>Euploea mulciber</i>	15,166	81.5	3705	80.2	1314	84.6	776	85.3	399	93.5	HQ37507	This study	[69]						
Lymantriidae	<i>Lymantria dispar</i>	15,569	79.9	3742	77.8	1351	84.2	799	85.2	435	96.1	NC_012893								

^aTermination codons were excluded in total codon count. ^bProtein coding genes. The bar (—) indicates lack of sequence information on the A + T-rich region in the genome. **Sasakia charonda kuriyamaensis*.

TABLE 2: Summary of the mitogenomes of *Euploea mulciber* and *Libythea celtis*.

Gene	Direction		Nucleotide no.		Size (bp)		Intergenic nucleotides		Anticodon		Start codon		Stop codon	
Species	1*	2*	1	2	1	2	1	2	1	2	1	2	1	2
tRNA ^{Met}	F	F	1-68	1-68	68	68			32-34 CAT	32-34 CAT				
tRNA ^{Leu(AUR)}	F	F	78-140	67-131	63	65	9	-2	106-108 GAT	95-97 GAT				
tRNA ^{Gln}	R	R	138-206	129-197	69	69	-3	-3	174-176 TTG	165-167 TTG				
ND2	F	F	253-1251	250-1263	999	1014	46	52			ATT	ATC	TAA	TAA
tRNA ^{Trp}	F	F	1250-1317	1265-1331	68	67	-2	1	1281-1283 TCA	1296-1298 TCA				
tRNA ^{Cys}	R	R	1310-1372	1324-1386	63	63	-8	-8	1340-1342 GCA	1355-1357 GCA				
tRNA ^{Tyr}	R	R	1374-1439	1387-1451	66	65	1	0	1406-1408 GTA	1418-1420 GTA				
COI	F	F	1442-2972	1458-2988	1531	1531	2	6			CGA	CGA	T-tRNA ^{Leu(UUR)}	T-tRNA ^{Leu(UUR)}
tRNA ^{Leu(UUR)}	F	F	2973-3039	2989-3055	67	67	0	0	3003-3005 TAA	3019-3021 TAA				
COII	F	F	3040-3712	3056-3737	673	682	0	0			ATG	ATG	T-tRNA ^{Lys}	T-tRNA ^{Lys}
tRNA ^{Lys}	F	F	3713-3783	3738-3808	71	71	0	0	3743-3745 CTT	3768-3770 CTT				
tRNA ^{Asp}	F	F	3783-3848	3808-3873	66	66	-1	-1	3813-3815 GTC	3838-3840 GTC				
ATP8	F	F	3849-4010	3874-4041	162	168	0	0			ATT	ATT	TAA	TAA
ATP6	F	F	4004-4681	4035-4712	678	678	-7	-7			ATG	ATG	TAA	TAA
COIII	F	F	4681-5469	4712-5500	789	789	-1	-1			ATG	ATG	TAA	TAA
tRNA ^{Gly}	F	F	5476-5541	5503-5569	66	67	6	2	5506-5508 TCC	5533-5535 TCC				
ND3	F	F	5542-5895	5570-5923	354	354	0	0			ATT	ATT	TAG	TAA
tRNA ^{Ala}	F	F	5894-5959	5928-5996	66	69	-2	4	5925-5927 TGC	5961-5963 TGC				
tRNA ^{Arg}	F	F	5959-6024	5997-6059	66	63	-1	0	5987-5989 TCG	6024-6026 TCG				
tRNA ^{Asn}	F	F	6026-6091	6065-6130	66	66	1	5	6057-6059 GTT	6096-6098 GTT				
tRNA ^{Ser(AGN)}	F	F	6090-6149	6129-6190	60	62	-2	-2	6110-6112 TCT	6150-6152 GCT				
tRNA ^{Glu}	F	F	6158-6223	6192-6259	66	68	8	1	6188-6190 TTC	6224-6226 TTC				
tRNA ^{Phe}	R	R	6222-6286	6258-6325	65	68	-2	-2	6254-6256 GAA	6290-6292 GAA				
ND5	R	R	6287-8003	6326-8057	1717	1732	0	0			ATT	ATT	T-tRNA ^{Phe}	T-tRNA ^{Phe}
tRNA ^{His}	R	R	8019-8083	8058-8122	65	65	15	0	8051-8053 GTG	8090-8092 GTG				
ND4	R	R	8084-9422	8123-9461	1339	1339	0	0			ATG	ATG	T-tRNA ^{His}	T-tRNA ^{His}
ND4L	R	R	9422-9709	9464-9754	288	291	-1	2			ATG	ATG	TAA	TAG
tRNA ^{Thr}	F	F	9712-9775	9757-9821	64	65	2	2	9742-9744 TGT	9788-9790 TGT				
tRNA ^{Pro}	R	R	9776-9839	9822-9886	64	65	0	0	9808-9810 TGG	9854-9856 TGG				
ND6	F	F	9842-10366	9889-10419	525	531	2	2			ATA	ATT	TAA	TAA
CytB	F	F	10371-11,519	10419-11,567	1149	1149	4	-1			ATG	ATG	TAA	TAA
tRNA ^{Ser(UCN)}	F	F	11,518-11,582	11,570-11,636	65	67	-2	2	11547-11549 TGA	11600-11602 TGA				

TABLE 2: Continued.

Gene	Direction		Nucleotide no.		Size (bp)		Intergenic nucleotides		Anticodon		Start codon		Stop codon	
	1*	2*	1	2	1	2	1	2	1	2	1	2	1	2
Species														
ND1	R	R	11,603–12,544	11,654–12,592	942	939	20	17						
tRNA ^{Leu(CUN)}	R	R	12,546–12,612	12,594–12,660	67	67	1	1	12581–12583 TAG	12629–12631 TAG				
lrRNA	R	R	12,613–13,926	12,661–13,995	1314	1335	0	0						
tRNA ^{Val}	R	R	13,927–13,991	13,996–14,061	65	66	0	0	13961–13963 TAC	14029–14031 TAC				
srRNA	R	R	13,992–14,767	14,062–14,836	776	775	0	0						
D-loop	F	F	14,768–15,166	14,837–15,164	399	328	0	0						

1* indicates the species of *Euploea mulciber* and 2* indicates the species of *Libythea celtis*.

FIGURE 1: Map of the mitogenomes of *Euploea mulciber* and *Libythea celtis*. The abbreviations for the genes are as follows: ATP6, ATP8, ATP9: synthase subunits 6 and 8 genes; CytB: cytochrome oxidase b gene; COI–COIII: cytochrome oxidase c subunit 1–3 genes; ND1–6, ND4L, ND4H: dehydrogenase subunits 1–6 and 4L. Gene names are the standard abbreviations used in this paper; tRNA genes are denoted as one-letter symbol according to the IUPAC-IUB single letter amino acid codes. Gene names that are not underlined indicate the direction of transcription clockwise, while the underlining indicates counterclockwise.

TABLE 3: Nucleotide composition at each codon position of the concatenated 13 PCGs in the Papilionoidea mitogenome.

Species	1st codon position				2nd codon position				3rd codon position				Overall			
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
<i>Papilio xuthus</i>	36.9	37.2	10.5	15.4	21.9	48.9	16.2	13.0	40.6	48.9	6.5	3.9	33.1	45.0	11.1	10.8
<i>Agehana maraho</i>	36.9	37.2	10.5	15.4	21.9	48.9	16.2	13.0	40.6	48.9	6.5	3.9	33.1	45.0	11.1	10.8
<i>Troides aeacus</i>	36.6	37.0	10.3	16.1	22.5	48.4	15.9	13.2	40.5	52.2	4.9	2.4	33.2	45.8	10.4	10.6
<i>Sericinus montela</i>	37.5	37.8	9.6	15.1	22.0	48.9	16.2	13.0	43.1	50.1	4.3	2.5	34.2	45.6	10.0	10.2
<i>Teinopalpus aureus</i>	37.0	37.0	10.4	15.6	21.9	48.2	16.6	13.3	40.6	50.2	5.5	3.7	33.1	45.1	10.8	10.9
<i>Luehdorfia chinensis</i>	36.6	37.4	10.1	15.9	22.2	48.5	16.2	13.1	42.7	51.3	3.7	2.3	33.8	45.7	10.0	10.4
<i>Parnassius bremeri</i>	37.0	37.9	9.4	15.7	22.2	48.4	16.4	13.0	44.1	51.0	3.3	1.7	34.4	45.8	9.7	10.2
<i>Coreana raphaelis</i>	38.3	38.1	8.8	14.9	22.0	49.0	15.8	13.2	45.8	51.2	2.0	1.0	35.4	46.1	8.8	9.7
<i>Spindasis takanonis</i>	38.6	38.2	9.2	14.1	22.5	49.4	15.2	12.9	43.6	51.0	2.8	2.7	34.9	46.2	9.1	9.9
<i>Protantigius superans</i>	37.7	37.7	9.2	15.3	22.0	48.8	15.9	13.2	44.0	50.6	3.2	2.2	34.6	45.7	9.4	10.2
<i>Pieris rapae</i>	36.0	37.6	9.9	16.5	22.1	48.0	16.6	13.4	42.7	48.2	5.3	3.8	33.6	44.6	10.6	11.2
<i>Artogeia melete</i>	36.7	36.8	10.3	16.2	21.9	47.9	16.8	13.4	42.3	49.8	4.6	3.4	33.6	44.8	10.6	11.0
<i>Apatura ilia</i>	37.2	37.6	9.6	15.6	22.0	48.8	15.9	13.2	41.7	49.7	5.3	3.3	33.6	45.4	10.3	10.7
<i>Sasakia charonda</i>	36.7	37.6	9.8	15.9	22.3	48.3	16.1	13.3	40.9	48.5	6.1	4.4	33.3	44.8	10.7	11.2
<i>Sasakia charonda*</i>	36.8	37.7	9.7	15.8	22.2	48.5	16.1	13.2	40.6	48.8	6.0	4.6	33.2	45.0	10.6	11.2
<i>Kallima inachus</i>	36.4	37.2	10.1	16.3	21.8	48.5	16.4	13.3	42.5	51.2	3.5	2.7	33.6	45.6	10.0	10.8
<i>Argyreus hyperbius</i>	37.4	37.1	9.9	15.6	22.6	48.2	6.1	13.1	41.5	51.5	4.6	2.4	33.8	45.6	10.2	10.4
<i>Acraea issoria</i>	36.6	36.7	10.7	15.9	23.0	47.8	16.2	13.1	39.7	50.1	6.4	3.7	33.1	44.9	11.1	10.9
<i>Parathyma sulpitia</i>	36.9	37.9	9.8	15.4	22.3	48.2	16.4	13.1	42.0	54.7	2.1	1.2	33.7	46.9	9.4	9.9
<i>Calinaga davidis</i>	36.3	37.7	10.4	15.7	22.2	48.4	16.4	13.1	40.7	51.5	5.1	2.7	33.0	45.8	10.6	10.5
<i>Eumenis autonoe</i>	35.9	36.3	11.3	16.4	21.5	48.4	16.4	13.7	39.7	48.8	7.1	4.4	32.3	44.5	11.6	11.5
<i>Euploea mulciber</i>	37.4	37.4	9.5	15.7	21.9	48.5	16.5	13.1	44.1	51.3	3.0	1.6	34.5	45.7	9.7	10.1
<i>Libythea celtis</i>	36.6	37.6	9.6	16.2	22.0	48.2	16.5	13.3	44.1	51.4	2.6	1.9	34.2	45.7	9.6	10.4
<i>Ctenoptilum vasava</i>	36.1	38.1	10.0	15.7	22.3	48.0	16.7	13.0	41.0	51.2	4.9	2.9	33.1	45.8	10.5	10.5

Start codons and stop codons were excluded in the count. **Sasakia charonda kuryamaensis*.

TABLE 4: Codon usage of the protein-coding genes of the *Libythea celtis* and *Euploea mulciber* mitogenomes.

Codon	N (RSCU)
UUU(F)	360 (1.88)/346 (1.89)
UUC(F)	23 (0.12)/21 (0.11)
UUA(L)	489 (5.37)/483 (5.35)
UUG(L)	5 (0.05)/8 (0.09)
CUU(L)	35 (0.38)/34 (0.38)
CUC(L)	1 (0.01)/3 (0.03)
CUA(L)	16 (0.18)/14 (0.15)
CUG(L)	0 (0)/0 (0)
AUU(I)	425 (1.95)/436 (1.91)
AUC(I)	11 (0.05)/21 (0.09)
AUA(M)	276 (1.94)/287 (1.93)
AUG(M)	9 (0.06)/10 (0.07)
GUU(V)	69 (1.99)/63 (1.98)
GUC(V)	2 (0.06)/1 (0.03)
GUA(V)	65 (1.87)/62 (1.95)
GUG(V)	3 (0.09)/1 (0.03)
UCU(S)	120 (2.94)/97 (2.45)
UCC(S)	7 (0.17)/8 (0.2)
UCA(S)	83 (2.03)/94 (2.37)
UCG(S)	1 (0.02)/5 (0.13)
CCU(P)	73 (2.39)/77 (2.52)
CCC(P)	9 (0.3)/10 (0.33)
CCA(P)	39 (1.28)/35 (1.15)
CCG(P)	1 (0.03)/0 (0)
ACU(T)	84 (2.15)/92 (2.36)
ACC(T)	2 (0.05)/7 (0.18)
ACA(T)	70 (1.79)/53 (1.36)
ACG(T)	0 (0)/4 (0.1)
GCU(A)	70 (2.28)/75 (2.36)
GCC(A)	6 (0.2)/1 (0.03)
GCA(A)	46 (1.5)/49 (1.54)
GCG(A)	1 (0.03)/2 (0.06)
UAU(Y)	173 (1.9)/18 (1.94)
UAC(Y)	9 (0.1)/6 (0.06)
UAA(*)	0 (0)/0 (0)
UAG(*)	0 (0)/0 (0)
CAU(H)	59 (1.79)/60 (1.79)
CAC(H)	7 (0.21)/7 (0.21)
CAA(Q)	61 (1.91)/56 (1.9)
CAG(Q)	3 (0.09)/3 (0.1)
AAU(N)	250 (1.91)/24 (1.91)
AAC(N)	12 (0.09)/11 (0.09)
AAA(K)	93 (1.81)/99 (1.96)
AAG(K)	10 (0.19)/2 (0.04)
GAU(D)	69 (1.97)/58 (1.9)
GAC(D)	1 (0.03)/3 (0.1)
GAA(E)	60 (1.76)/68 (1.89)
GAG(E)	8 (0.24)/4 (0.11)
UGU(C)	26 (1.73)/28 (1.65)
UGC(C)	4 (0.27)/6 (0.35)

TABLE 4: Continued.

Codon	N (RSCU)
UGA(W)	93 (1.96)/92 (1.98)
UGG(W)	2 (0.04)/1 (0.02)
CGU(R)	10 (0.77)/14 (1.1)
CGC(R)	1 (0.08)/1 (0.08)
CGA(R)	39 (3)/36 (2.82)
CGG(R)	2 (0.15)/0 (0)
AGU(S)	32 (0.78)/29 (0.73)
AGC(S)	3 (0.07)/2 (0.05)
AGA(S)	81 (1.98)/81 (2.04)
AGG(S)	0 (0)/1 (0.03)
GGU(G)	50 (1)/55 (1.13)
GGC(G)	0 (0)/2 (0.04)
GGA(G)	126 (2.52)/120 (2.47)
GGG(G)	24 (0.48)/17 (0.35)

N: frequency of codon used; RSCU: relative synonymous codon usage; * stop codon.

Start codons and stop codons were excluded in total codon counts.

The mitogenome size of 15,166 bp and 15,164 bp determined here is close to most lepidopteran mitogenomes although slightly shorter than those of most completely sequenced lepidopteran insects (Table 1). The *Euploea mulciber* mitochondrial genes are interleaved with a total of 117 bp intergenic spacer sequences, which are spread over 13 regions, ranging from 1 to 46 bp, and only three spacers span longer than 10 bp, located between tRNA^{Gln} and ND2 (46 bp), between ND5 and tRNA^{His} (15 bp), and between tRNA^{Ser} (UCN) and ND1 (20 bp). Additionally, a total of 32 bp overlapped nucleotides are spread over 12 regions, and every region is shorter than 10 bp in length. The *Libythea celtis* mitochondrial genes are interleaved with a total of 97 bp intergenic spacer sequences, which are also spread over 13 regions, ranging from 1 to 52 bp, among which only two spacers span longer than 10 bp, located between tRNA^{Gln} and ND2 (52 bp) and between tRNA^{Ser} (UCN) and ND1 (17 bp). Besides, a total of 27 bp overlapped nucleotides are spread over 9 regions, and every region is shorter than 10 bp in size. The nucleotide compositions of A + T in *Euploea mulciber* and *Libythea celtis* mitogenomes are 81.5% and 81.2%, respectively, showing strong A + T bias. Correspondingly, the AT contents of the two butterfly species are up to 80.2% and 80.0% for the PCGs, 84.6% and 84.7% for the lrRNA gene, 85.3% and 85.4% for the srRNA gene, and 93.5% and 96.3% for the control region. This case is similar to those observed in other rhopaloceran mitogenomes, ranging from 79.7% in *Pieris rapae* and *Acraea issoria* to 82.7% in *Coreana raphaelis* (Tables 1 and 2).

3.2. Protein-Coding Genes. The total codon numbers of the 13 PCGs for *Euploea mulciber* and *Libythea celtis* are 3722 and 3705, respectively (Table 1). The nucleotide composition for each codon position and the codon usage of PCGs are shown in Tables 3 and 4. For codon usage, some codons have much more numbers, such as the TTA(L), ATT(I), TTT(F), and ATA(M), while some codons, such as CGT(L), AGG(S), and

TABLE 5: Nucleotide composition and skewness in the rhopaloceran mitogenomes.

Species	Major strand PCGs			Minor strand PCGs			Whole PCGs			Whole genome		
	A + T	AT skew	GC skew	A + T	AT skew	GC skew	A + T	AT skew	GC skew	A + T	AT skew	GC skew
	%			%			%			%		
<i>Papilio xuthus</i>	77.1	-0.114	-0.206	79.8	-0.210	0.334	78.1	-0.152	-0.014	80.0	-0.011	-0.216
<i>Agehana maraho</i>	77.1	-0.114	-0.206	79.8	-0.210	0.334	78.1	-0.152	-0.014	80.6	0.006	-0.262
<i>Teinopalpus aureus</i>	77.2	-0.125	-0.172	80.0	-0.196	0.322	78.3	-0.153	0.002	79.8	-0.005	-0.238
<i>Troides aeacus</i>	77.7	-0.166	-0.168	81.1	-0.150	0.332	79.1	-0.160	0.010	80.2	-0.040	-0.230
<i>Parnassius bremeri</i>	78.9	-0.126	-0.112	82.2	-0.166	0.281	80.2	-0.142	0.024	81.3	-0.011	-0.194
<i>Sericinus montela</i>	78.5	-0.124	-0.146	81.8	-0.172	0.301	79.8	-0.143	0.009	81.0	-0.008	-0.223
<i>Luehdorfia chinensis</i>	78.0	-0.129	-0.134	81.9	-0.179	0.296	79.6	-0.150	0.019	80.5	-0.005	-0.205
<i>Ctenoptilum vasava</i>	77.2	-0.151	-0.145	81.6	-0.174	0.285	78.9	-0.160	0.000	80.5	-0.029	-0.205
<i>Pieris rapae</i>	77.2	-0.095	-0.150	79.8	-0.212	0.358	78.2	-0.141	0.030	79.7	0.014	-0.232
<i>Artogeia melete</i>	77.7	-0.098	-0.152	79.7	-0.213	0.324	78.4	-0.143	0.020	79.8	0.012	-0.222
<i>Coreana raphaelis</i>	80.3	-0.143	-0.074	83.4	-0.115	0.270	81.5	-0.132	0.044	82.7	-0.047	-0.158
<i>Protantigius superans</i>	79.3	-0.137	-0.087	81.9	-0.140	0.276	80.3	-0.138	0.041	81.7	-0.036	-0.174
<i>Spindasis takanonis</i>	80.1	-0.104	-0.117	82.7	-0.193	0.338	81.1	-0.139	0.044	82.4	0.004	-0.214
<i>Apatura ilia</i>	77.9	-0.124	-0.140	80.7	-0.187	0.314	79.0	-0.149	0.021	80.4	-0.011	-0.207
<i>Sasakia charonda</i>	76.9	-0.118	-0.152	80.0	-0.194	0.330	78.1	-0.147	0.023	79.9	-0.006	-0.219
<i>Sasakia charonda*</i>	77.2	-0.121	-0.142	79.7	-0.196	0.324	78.2	-0.151	0.028	79.9	-0.007	-0.218
<i>Kallima inachus</i>	78.1	-0.136	-0.125	81.1	-0.177	0.336	79.2	-0.152	0.036	80.3	-0.013	-0.212
<i>Parathyma sulpitia</i>	79.2	-0.172	-0.100	83.1	-0.154	0.266	80.6	-0.164	0.026	81.9	-0.048	-0.178
<i>Eumenis autonoë</i>	75.4	-0.135	-0.187	79.3	-0.193	0.337	76.8	-0.159	-0.004	79.1	-0.016	-0.244
<i>Calinaga davidis</i>	77.5	-0.164	-0.147	81.1	-0.159	0.270	78.8	-0.162	-0.005	80.4	-0.045	-0.200
<i>Acraea issoria</i>	76.7	-0.142	-0.176	80.1	-0.164	0.307	78.0	-0.146	-0.009	79.7	-0.024	-0.238
<i>Argyreus hyperbius</i>	77.8	-0.136	-0.153	82.0	-0.166	0.322	79.4	-0.149	0.010	80.8	-0.025	-0.219
<i>Libythea celtis</i>	78.8	-0.124	-0.094	81.8	-0.174	0.297	79.9	-0.144	0.040	81.2	-0.017	-0.181
<i>Euploea mulciber</i>	79.0	-0.142	-0.124	82.1	-0.140	0.307	80.2	-0.140	0.020	81.5	-0.038	-0.211

Start codons and stop codons were excluded in the count.

The skewness of whole PCGs and whole genome were calculated from major strands. **Sasakia charonda kuriyamaensis*.

GGC(G) are not used. In addition, the synonymous codon bias, namely, the relative synonymous codon usage (RSCU) are detected, and their bias patterns roughly correspond to those of codon usage, for example, the four most frequently used codons (TTA, ATT, TTT, and ATA) account for 10.14% of all the codons. The AT and GC nucleotide skewness values for the PCGs as well as the whole genomes are listed in Table 5, and the results showed that they are slightly negative and positive for the whole PCGs and their minor strand of the two nymphalid species, respectively, indicating that more T and G are used than A and C; however, both values are slightly negative for the major strands, indicating that more T and C are used in these coding regions. In general, these codon constitutional and biased patterns are significantly congruent with those of other lepidopteran species determined to date [13, 18, 19, 36–38].

Except for COI, 12 out of 13 PCGs in *Euploea mulciber* and *Libythea celtis* use standard ATN start codons: ATA for ND6; ATT for ND2, ATP8, ND3, and ND5; ATG for COII, ATP6, COIII, ND4, ND4L, CytB, and ND1 in *Euploea mulciber*; ATC for ND2; ATT for ATP8, ND3, ND5, and ND6; and ATG for

COII, ATP6, COIII, ND4, ND4L, CytB, and ND1 in *Libythea celtis*.

In Figure 2, we show the alignment of the initiation site for the COI genes of 24 completely sequenced rhopaloceran mitogenomes. All the mitogenomes harbor a few canonical ATN initiators detected in the start region of the COI gene or in the immediately neighboring sequence of the precedent tRNA^{Tyr}. However, in eleven of them (*Artogeia melete*, *Coreana raphaelis*, *Argyreus hyperbius*, *Teinopalpus aureus*, *Papilio maraho*, *Luehdorfia chinensis*, *Parnassius bremeri*, *Protantigius superans*, *Troides aeacus*, *Libythea celtis*, and *Euploea mulciber*), a TAG (stop codon) is present at the beginning region of the COI gene, while in the other five mitogenomes (*Papilio xuthus*, *Sericinus montela*, *Ctenoptilum vasava*, and two *Sasakia* subspecies), the proposed ATN initiators are present within the region coding for tRNA^{Tyr}. Thus, the ATN initiators should not be considered as the start codon for the COI gene. According to these criteria, the first nonoverlapping codon in the COI gene is the CGA designating arginine existing in a highly conserved region across rhopaloceran insects [13].

		← tRNA ^{Tyr}	COI→						
<i>Papilio xuthus</i>	ataaattacaattatcgctatatattcagccattttat	acg	CGA	AAA	TGA	CTT	RKWL		
<i>Troides aeacus</i>	aataaattacaattatcgcttaaaactcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Agehana maraho</i>	aataaattacaattatcgcttatattcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Teinopalpus aureus</i>	ataaattacaattatcgcttaaaattcagccattttat	tag	CGA	AAA	TGA	TTT	RKWF		
<i>Parnassius bremeri</i>	aataaattacaattatcgcttatactcagccattttat	tag	CGA	AAA	TGA	CTA	RKWL		
<i>Sericinus montela</i>	taaattacaattatcgcttaataactcagccattttat	aag	CGA	AAA	TGA	CTT	MKRKWL		
<i>Luehdorfia chinensis</i>	tacaattatcgcttataattcagccattttattatt	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Ctenoptilum vasava</i>	ataaattacaattatcgcttatactcagccattttat	cta	GCG	AAA	AAA	TTT	MLAKKF		
<i>Artogeia melete</i>	aataaattacaattatcgcttataactcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Pieris rapae</i>	aaatttacaattatcgcttatacctcagccattttat	aag	CGA	AAA	TGA	CTT	RKWL		
<i>Coreana raphaelis</i>	ataaattacaattatcgcttatactcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Protantigius superans</i>	ataaattacaattatcgcttatacctcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Spindasis takanonis</i>	aataaattacaattatcgcttataactcagccattttat	tgc	CGA	AAA	TGA	CTT	RKWL		
<i>Apatura ilia</i>	aaatttacaattatcgcttataactcagccattttat	aag	CGA	AAA	TGA	TTA	RKWL		
<i>Sasakia charonda</i>	aaatttacaattatcgcttataactcagccattttat	aag	CGA	AAA	TGA	TTA	MTQPFYYKRKWL		
<i>Sasakia charonda*</i>	aaatttacaattatcgcttataactcagccattttat	aag	CGA	AAA	TGA	TTA	MTQPFYYKRKWL		
<i>Kallima inachus</i>	taaattacaattatcgcttatacctcagccattttat	aag	CGA	AAA	TGA	CTT	RKWL		
<i>Argyreus hyperbius</i>	ataaattacaattatcgcttaataactcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Acræa issoria</i>	aataaattacaattatcgcttataactcagccattttat	tgc	CGA	AAA	TGA	CTT	MRKWL		
<i>Parathyma sulphitia</i>	taaattacaattatcgcttataactcagccattttat	aag	CGA	AAA	TGA	CTT	RKWL		
<i>Calinaga davidis</i>	aataaattacaattatcgcttatacctcagccattttat	tgc	CGA	AAA	TGA	CTT	RKWL		
<i>Eumenis autonoe</i>	tttacaattatcgcttatacctcagccattttactttt	aag	CGA	AAA	TGA	CTT	RKWL		
<i>Libythea celtis</i>	aatttacaattatcgcttatacctcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Euploea mulciber</i>	aataaattacaattatcgcttataactcagccattttat	tag	CGA	AAA	TGA	CTT	RKWL		
<i>Lymantria dispar</i>	ttacaattaccgcttaaaactcagccattttat	aag	CGA	AAA	TGA	TTA	RKWL		

FIGURE 2: Alignment of the initiation context of the COI genes of 25 lepidopteran insect species, including that of *Libythea celtis* and *Euploea mulciber*. The first several codons for COI and their amino acids are shown on the right-hand side of the figure. Wave-lined nucleotides indicate the adjacent partial sequence of tRNA^{Tyr}. Arrows indicate the transcriptional direction. Boxed nucleotides indicate the currently proposed translation initiators for the COI gene of lepidopteran insects. The start codons for *Libythea celtis* and *Euploea mulciber* are designated as CGA. * *Sasakia charonda kuriyamaensis*.

Nine of the PCGs in *Euploea mulciber* and *Libythea celtis* harbor the complete termination codon TAA or TAG, that is, eight protein-coding sequences stop with TAA, one with TAG. The other four are with the incomplete termination codon T. This phenomenon of partial termination codons (i.e., T or TA) is observed in all sequenced lepidopteran insects and has been interpreted in terms of posttranscriptional polyadenylation, by which “A” residue(s) are added to create TAA terminator [13].

3.3. Ribosomal RNAs and Transfer RNAs. The 18S rRNA genes of the *Euploea mulciber* and *Libythea celtis* mitogenomes are 1314 and 1335 bp in length, respectively. As has been observed in other insects, this gene is located between tRNA^{Leu} (CUN) and tRNA^{Val} genes [12, 13, 15, 16, 19, 37, 39–45]. In the same way, the 16S rRNA genes of the *Euploea mulciber* and *Libythea celtis* mitogenomes are 776 and 775 bp in length, respectively, and located between tRNA^{Val} and A + T-rich region. Both of the two rRNA genes are encoded by the minor strand, and the lengths of them are well within the size range of other insects

determined. The AT contents of 18S rRNA and 16S rRNA genes are 84.6% and 85.3% in *Euploea mulciber* and 84.7% and 85.4% in *Libythea celtis*, respectively, and these values are also well within the range of the sequenced insects.

Like other insects, both *Euploea mulciber* and *Libythea celtis* harbor 22 tRNAs, among which both Leucine and Serine have two tRNAs, and the left tRNAs for each of 18 amino acid residues are all one in number (Figure 1). In this study, all the tRNAs of *Euploea mulciber* and *Libythea celtis* evidence the typical cloverleaf structures with the exception of tRNA^{Ser} (AGN), whose dihydrouridine (DHU) arm forms a simple loop (maps of secondary structure are not shown here). The case of tRNA^{Ser} (AGN) with the simple loop in the DHU arm has also been detected in most of insect species including the lepidopterans determined till now [12–19, 36, 37, 39, 42, 45–48].

3.4. A + T-Rich Region. The A + T-rich region is known for the initiation of replication in vertebrates and invertebrates and thus also called as the control region. The

FIGURE 3: Alignment of the initiation site for the A + T-rich regions of 23 completely sequenced lepidopteran mitogenomes. The shaded nucleotides indicate the poly-T stretch, and the boxed nucleotides indicate the conserved motif, ATAGA. **Sasakia charonda kuriyamaensis*.

Euploea mulciber and *Libythea celtis* A + T-rich regions located between srRNA and tRNA^{Met} were 399 and 328 bp in length, respectively (see Figure 1 and Table 2). Though the A + T-rich region of *Libythea celtis* is the shortest, its A + T content is the highest (96.3%) in all of the completely sequenced rhopaloceran mitogenomes known to date. In general, this region harbors the highest A + T content in the whole regions of the insect mitogenomes, and this feature is well presented in those of the two nymphalid species in this study. However, the regions of these two species harbor no large sequence repeats detected frequently in other insect groups, including some of the lepidopterans. Nonetheless, there are several short sequence repeats scattered throughout the whole region, for example, a poly T closed to the srRNA, a TA repeat in the middle of the A + T-rich region and a poly A closed to tRNA^{Met}.

Alignment of the initiation sites for the A + T-rich regions of 22 completely sequenced rhopaloceran mitogenomes is shown in Figure 3, and the results showed that all the A + T-rich regions harbor a poly-T stretch preceded by the ATAGA motif. The *Euploea mulciber* and *Libythea celtis* harbored a 21 bp and 19 bp long poly-T stretch, respectively. To the best of our knowledge, the poly-T stretch is quite conserved in all sequenced rhopaloceran insects, ranging from 12 bp in *Agehana maraho* [49] to 21 bp in *Euploea mulciber* of this study and *Apatura ilia* [50] (Figure 3), and it has been suggested as a possible recognition site for the replication initiation of the mtDNA minor strand owing to the facts that the insect mitochondrial A + T-rich region has been revealed to harbor the replication originating sites for both strands in *Drosophila* [40, 41], and its neighboring region has been identified as the position of the minor-strand replication origin in *Bombyx mori* [51].

A microsatellite-like elements (AT)₃(TA)₆ and (TA)₁₂ preceded by the ATTTA motif are presented in the middle of the *Euploea mulciber* and *Libythea celtis* A + T-rich region, respectively. Similar cases are commonly found in all other sequenced lepidopteran species [12–15, 17, 19, 36–39, 42–44, 46–48]. Additionally, a 12 bp poly A interrupted by T and a 7 bp poly A interrupted by G presented immediately the upstream of the tRNA^{Met}. This poly A element at the end

of the A + T-rich region and the short repeating sequences are probably universal features in lepidopterans, and their unknown functions await further in-depth studies.

3.5. Phylogenetic Analysis. Thirteen concatenated PCGs nucleotide sequences of complete rhopaloceran mitogenomes were used to analyze the phylogenetic relationships among the butterfly groups in this study. The resultant 14-taxon (including outgroup) data set is 11,459 characters after removal of gaps, among which 5402 are variable and 3759 are parsimony-informative. The results of neighbor-joining (NJ) and Bayesian inference (BI) analyses show that the topologies of the NJ and BI trees shown in Figure 4 are identical only with minor differences in their node support values. The trees indicates the following relationship: (Papilionidae + (Pieridae + (Nymphalidae + Lycaenidae))), which is congruent with that obtained by Walhberg et al. (2005): (Hesperiidae + (Papilionidae + (Pieridae + (Nymphalidae + (Lycaenidae + Riodinidae)))) [10], though the data from the Hesperiidae and Riodinidae is not available here. Additionally, within the Nymphalidae, *Euploea mulciber* stands at the base of the nymphalid tree, as a sister to all the remaining groups of the family, and *Libythea celtis* (Libytheinae) is the nearest sister to the *Calinaga davidis* (Calinaginae).

The danaids, formerly classified as familial taxa under the family Nymphalidae, have now been relegated to the subfamily Danainae. They possess unique features, such as the brush-like modified forelegs, poisonousness to their potential predators, which are markedly distinct with other nymphalid butterflies, and have been proposed as one of the basal groups in the family Nymphalidae by Walhberg et al. (2003) based on molecular evidences [52]. Morphologically, Freitas and Brown (2004) also revealed that the danaid clade (including Danaini, Tellervini, and Ithomiini) appeared as the basal group of the Nymphalidae excluding Libytheinae based on the immature features [53]. Hao et al. (2007) reconstructed the mitochondrial 16S rRNA secondary structures of the main butterfly lineages, and the results showed that the morphological characteristics of the danaid species *Euploea core* and *Parantica aglea* were markedly different from other nymphalid groups [54]. In this study, both NJ and BI trees

FIGURE 4: NJ tree (a) and BI tree (b) of the main lineages of butterflies based on nucleotide sequences of the 13 protein-coding genes; numbers at the nodes: bootstrap values in (a) and posterior probabilities in (b).

indicate that *Euploea mulciber* is sister to all other nymphalid species, standing at the base of the whole nymphalid tree. Therefore, we conclude that the danaid butterflies are the most primitive group of Nymphalidae.

The Libytheinae is one of the most morphologically unusual groups of butterflies, and their relationship with other butterfly groups remains controversial. Traditionally, they are considered as the basal group of Nymphalidae [16, 55–57]. Some scholars even considered them as a separate family owing to their unique morphological characters such as the markedly prolonged snout and well-developed foreleg in females [58, 59]. In addition, some phylogenetic analyses based on morphological, molecular, or combined

data also proposed that Libytheinae is the basal lineage of the Nymphalidae [10, 60–62]. However, this study reveals that *Libythea celtis* is closely related to *Calinaga davidis* (Calinaginae) with a weak bootstrap support value (65%) in NJ tree, whereas, stands as a sister to the species grouping of Limenitidinae, Heliconiinae, Nymphalinae, and Apaturinae with a high posterior probability value (0.97) in BI tree (Figure 4). Thus, its phylogenetic position within Nymphalidae awaits further clarification in the future, on the grounds of more taxa samplings as well as more realistic analysis methods. In conclusion, it is obvious that the Danainae is the basal branch of the Nymphalidae, rather than the Libytheinae.

Acknowledgments

This work was supported by the National Science Foundation of China (Grant no. 41172004), the CAS/SAFEA International Partnership Program for Creative Research Teams, Chinese Academy of Sciences (Grant no. KZCX22YW2JC104), the Provincial Key Projects of Natural Science Foundation, Colleges of Anhui Province (Grant no. KJ2010A142), and the Opening Funds from the State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences (Grant no. 104143).

References

- [1] D. R. Wolstenholme, "Animal mitochondrial DNA: structure and evolution," *International Review of Cytology*, vol. 141, pp. 173–216, 1992.
- [2] J. L. Boore, "Animal mitochondrial genomes," *Nucleic Acids Research*, vol. 27, no. 8, pp. 1767–1780, 1999.
- [3] D. A. Clayton, "Transcription and replication of animal mitochondrial DNAs," *International Review of Cytology*, vol. 141, pp. 217–232, 1992.
- [4] W. M. Brown, "Evolution of animal mitochondrial DNA," in *Evolution of Genes and Proteins*, M. Nei and P. K. Koehn, Eds., pp. 62–88, Sinauer, Sunderland, UK, 1983.
- [5] J. C. Avise, *Molecular Markers, Natural History and Evolution*, Chapman & Hall, New York, NY, USA, 1994.
- [6] P. R. Ackery, "Systematic and faunistic studies on butterflies," in *The Biology of Butterflies*, R. I. Vane-Wright and P. R. Ackery, Eds., pp. 9–21, Academic Press, London, UK, 1984.
- [7] P. R. Ackery, R. de Jong, and R. I. Vane-Wright, "The butterflies: hedyloidea, Hesperioidea and Papilionoidea," in *Lepidoptera: Moths and Butterflies. 1. Evolution, Systematics and Biogeography, Handbook of Zoology*, N. P. Kristensen, Ed., vol. 4, pp. 264–300, De Gruyter, Berlin, Germany, 1999.
- [8] D. J. Harvey, "Higher classification of the Nymphalidae, Appendix B," in *The Development and Evolution of Butterfly Wing Patterns*, H. F. Nijhout, Ed., pp. 255–273, Smithsonian Institution Press, Washington, DC, USA, 1991.
- [9] I. O. Chou, *Classification and Identification of Chinese Butterflies*, Henan Scientific and Technological Publishing House, Zhengzhou, China, 1998.
- [10] N. Wahlberg, M. F. Braby, A. V. Z. Brower et al., "Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers," *Proceedings of the Royal Society B*, vol. 272, no. 1572, pp. 1577–1586, 2005.
- [11] A. Y. Kawahara, "Behavioral observations of *Libytheana carinenta* Cramer," *News of the Lepidopterists' Society*, vol. 45, pp. 107–108, 2003.
- [12] I. Kim, E. M. Lee, K. Y. Seol et al., "The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae)," *Insect Molecular Biology*, vol. 15, no. 2, pp. 217–225, 2006.
- [13] M. I. Kim, J. Y. Baek, M. J. Kim et al., "Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects," *Molecules and Cells*, vol. 28, no. 4, pp. 347–363, 2009.
- [14] G. Hong, S. Jiang, M. Yu et al., "The complete nucleotide sequence of the mitochondrial genome of the cabbage butterfly, *Artogeia melete* (Lepidoptera: Pieridae)," *Acta Biochimica et Biophysica Sinica*, vol. 41, no. 6, pp. 446–455, 2009.
- [15] J. Hu, D. Zhang, J. Hao, D. Huang, S. Cameron, and C. Zhu, "The complete mitochondrial genome of the yellow coaster, *Acraea issoria* (Lepidoptera: Nymphalidae: Heliconiinae: Acraeini): sequence, gene organization and a unique tRNA translocation event," *Molecular Biology Reports*, vol. 37, no. 7, pp. 3431–3438, 2010.
- [16] Y. J. Gong, B. C. Shi, Z. J. Kang, F. Zhang, and S. J. Wei, "The complete mitochondrial genome of the oriental fruit moth *Grapholita molesta* (Busck) (Lepidoptera: Tortricidae)," *Molecular Biology Reports*, pp. 1–8, 2011.
- [17] L. L. Tian, X. Y. Sun, M. Chen, Y. H. Gai, J. S. Hao, and Q. Yang, "Complete mitochondrial genome of the five-dot sergeant *Parathyma sulphitia* (Nymphalidae: Limenitidinae) and its phylogenetic implications," *Zoology Research*, vol. 33, no. 2, pp. 133–143, 2012.
- [18] L. W. Ji, J. S. Hao, Y. Wang, D. Y. Huang, J. L. Zhao, and C. D. Zhu, "The complete mitochondrial genome of the dragon swallowtail, *Sericinus montela* Gray (Lepidoptera: Papilionidae) and its phylogenetic implication," *Acta Entomologica Sinica*, vol. 55, no. 1, pp. 91–100, 2012.
- [19] J. S. Hao, Q. Q. Sun, H. B. Zhao, X. Y. Sun, Y. H. Gai, and Q. Yang, "The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperidae: Pyrginae) and its phylogenetic implication," *Comparative and Functional Genomics*, vol. 2012, Article ID 328049, 13 pages, 2012.
- [20] J. S. Hao, C. X. Li, X. Y. Sun, and Q. Yang, "Phylogeny and divergence time estimation of cheilostome bryozoans based on mitochondrial 16S rRNA sequences," *Chinese Science Bulletin*, vol. 12, pp. 1205–1211, 2005.
- [21] C. Simon, F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook, "Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers," *Annals of the Entomological Society of America*, vol. 87, no. 6, pp. 651–701, 1994.
- [22] M. S. Caterino and F. A. H. Sperling, "PapilioPhylogeny based on mitochondrial cytochrome oxidase I and II genes," *Molecular Phylogenetics and Evolution*, vol. 11, no. 1, pp. 122–137, 1999.
- [23] R. B. Simmons and S. J. Weller, "Utility and evolution of cytochrome b in insects," *Molecular Phylogenetics and Evolution*, vol. 20, no. 2, pp. 196–210, 2001.
- [24] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins, "The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools," *Nucleic Acids Research*, vol. 25, no. 24, pp. 4876–4882, 1997.
- [25] V. K. Singh, A. K. Mangalam, S. Dwivedi, and S. Naik, "Primer premier: program for design of degenerate primers from a protein sequence," *BioTechniques*, vol. 24, no. 2, pp. 318–319, 1998.
- [26] T. A. Hall, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT," *Nucleic Acids Symposium Series*, no. 41, pp. 95–98, 1999.
- [27] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.
- [28] T. M. Lowe and S. R. Eddy, "tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence," *Nucleic Acids Research*, vol. 25, no. 5, pp. 955–964, 1997.

- [29] S. Kumar, K. Tamura, and M. Nei, "MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment," *Briefings in Bioinformatics*, vol. 5, no. 2, pp. 150–163, 2004.
- [30] N. Saitou and M. Nei, "The neighbor-joining method: a new method for reconstructing phylogenetic trees," *Molecular Biology and Evolution*, vol. 4, no. 4, pp. 406–425, 1987.
- [31] Z. Yang and B. Rannala, "Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method," *Molecular Biology and Evolution*, vol. 14, no. 7, pp. 717–724, 1997.
- [32] M. Kimura, "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences," *Journal of Molecular Evolution*, vol. 16, no. 2, pp. 111–120, 1980.
- [33] J. P. Huelsenbeck and F. Ronquist, "MRBAYES: bayesian inference of phylogenetic trees," *Bioinformatics*, vol. 17, no. 8, pp. 754–755, 2001.
- [34] F. Ronquist and J. P. Huelsenbeck, "MrBayes 3: bayesian phylogenetic inference under mixed models," *Bioinformatics*, vol. 19, no. 12, pp. 1572–1574, 2003.
- [35] D. Posada and K. A. Crandall, "MODELTEST: testing the model of DNA substitution," *Bioinformatics*, vol. 14, no. 9, pp. 817–818, 1998.
- [36] P. Salvato, M. Simonato, A. Battisti, and E. Negrisola, "The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae)," *BMC Genomics*, vol. 9, p. 331, 2008.
- [37] X. C. Wang, X. Y. Sun, Q. Q. Sun et al., "The complete mitochondrial genome of the laced fritillary *Argyreus hyperbius* (Lepidoptera: Nymphalidae)," *Zoology Research*, vol. 32, no. 5, pp. 465–475, 2011.
- [38] H. N. Chai, Y. Z. Du, and B. P. Zhai, "Characterization of the complete mitochondrial genomes of *Cnaphalocrocis medinalis* and *Chilo suppressalis* (Lepidoptera: Pyralidae)," *International Journal of Biological Sciences*, vol. 8, no. 4, pp. 561–579, 2012.
- [39] Z. H. Mao, J. S. Hao, G. P. Zhu, J. Hu, M. M. Si, and C. D. Zhu, "Sequencing and analysis of the complete mitochondrial genome of *Pieris rapae* Linnaeus (Lepidoptera: Pieridae)," *Acta Entomologica Sinica*, vol. 53, no. 11, pp. 1295–1304, 2010.
- [40] C. R. Fauron and D. R. Wolstenholme, "Extensive diversity among *Drosophila* species with respect to nucleotide sequences within the adenine + thymine-rich region of mitochondrial DNA molecules," *Nucleic Acids Research*, vol. 8, no. 11, pp. 2439–2452, 1980.
- [41] D. O. Clary and D. R. Wolstenholme, "Drosophila mitochondrial DNA: conserved sequences in the A+T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA," *Journal of Molecular Evolution*, vol. 25, no. 2, pp. 116–125, 1987.
- [42] K. Yukuhiro, H. Sezutsu, M. Itoh, K. Shimizu, and Y. Banno, "Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkworm, *Bombyx mandarina*, and its close relative, the domesticated silkworm, *Bombyx mori*," *Molecular Biology and Evolution*, vol. 19, no. 8, pp. 1385–1389, 2002.
- [43] E. S. Lee, K. S. Shin, M. S. Kim, H. Park, S. Cho, and C. B. Kim, "The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae)," *Gene*, vol. 373, no. 1–2, pp. 52–57, 2006.
- [44] S. R. Kim, M. I. Kim, M. Y. Hong et al., "The complete mitogenome sequence of the Japanese oak silkworm, *Antheraea yamamai* (Lepidoptera: Saturniidae)," *Molecular Biology Reports*, vol. 36, no. 7, pp. 1871–1880, 2009.
- [45] L. Yang, Z. J. Wei, G. Y. Hong, S. T. Jiang, and L. P. Wen, "The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae)," *Molecular Biology Reports*, vol. 36, no. 6, pp. 1441–1449, 2009.
- [46] S. L. Cameron and M. F. Whiting, "The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths," *Gene*, vol. 408, no. 1–2, pp. 112–123, 2008.
- [47] M. Y. Hong, E. M. Lee, Y. H. Jo et al., "Complete nucleotide sequence and organization of the mitogenome of the silk moth *Caligula boisduvalii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects," *Gene*, vol. 413, no. 1–2, pp. 49–57, 2008.
- [48] Y. Liu, Y. Li, M. Pan et al., "The complete mitochondrial genome of the Chinese oak silkworm, *Antheraea pernyi* (Lepidoptera: Saturniidae)," *Acta Biochimica et Biophysica Sinica*, vol. 40, no. 8, pp. 693–703, 2008.
- [49] L. W. Wu, D. C. Lees, S. H. Yen, C. C. Lu, and Y. F. Hsu, "complete mitochondrial genome of the near-threatened swallowtail, *Agehana maraho* (Lepidoptera: Papilionidae): evaluating sequence variability and suitable markers for conservation genetic studies," *Entomological News*, vol. 121, no. 3, pp. 267–280, 2010.
- [50] M. Chen, L. L. Tian, Q. H. Shi, T. W. Cao, and J. S. Hao, "Complete mitogenome of the lesser purple emperor *Apatura ilia* (Lepidoptera: Nymphalidae: Apaturinae) and comparison with other nymphalid butterflies," *Zoology Research*, vol. 33, no. 2, pp. 191–201, 2012.
- [51] S. Saito, K. Tamuea, and T. Aotsuka, "Replication origin of mitochondrial DNA in insects," *Genetics*, vol. 171, pp. 433–448, 2005.
- [52] N. Wahlberg, E. Weingartner, and S. Nylin, "Towards a better understanding of the higher systematics of Nymphalidae (Lepidoptera: Papilionoidea)," *Molecular Phylogenetics and Evolution*, vol. 28, no. 3, pp. 473–484, 2003.
- [53] A. V. L. Freitas and K. S. Brown, "Phylogeny of the Nymphalidae (Lepidoptera)," *Systematic Biology*, vol. 53, no. 3, pp. 363–383, 2004.
- [54] J. S. Hao, C. Y. Su, G. P. Zhu, N. Chen, D. X. Wu, and X. P. Zhang, "The molecular morphologies of mitochondrial 16S rRNA of the main butterfly lineages and their phylogenetic significances," *Journal of Genetics and Molecular Biology*, vol. 18, no. 2, pp. 111–113, 2007.
- [55] N. P. Kristensen, "Remarks on the family-level phylogeny of butterflies (Insecta, Lepidoptera, Rhopalocera)," *Zoological Systematics and Evolutionary Research*, vol. 14, no. 1, pp. 25–33, 1976.
- [56] R. De Jong, R. I. Vane-Wright, and P. R. Ackery, "The higher classification of butterflies (Lepidoptera): problems and prospects," *Insect Systematics and Evolution*, vol. 27, no. 1, pp. 65–101, 1996.
- [57] R. I. Vane-Wright, "Evidence and identity in butterfly systematics," in *Butterflies: Ecology and Evolution Taking Flight*, C. L. Boggs, W. B. Watt, and P. R. Ehrlich, Eds., pp. 477–513, University of Chicago Press, Chicago, 2003.
- [58] A. Pagenstecher, "Lepidoptera rhopalocera: family libytheidae," in *Fascicle*, P. Wytsman, Ed., vol. 5, pp. 1–4, Genera Insectorum, Bruxelles, Belgium, 1902.

- [59] A. Pagenstecher, "Libytheidae," in *Lepidopterorum Catalogus*, C. Aurivillius and H. Wagner, Eds., R. Friedländer und sohn, Berlin, Germany, 1911.
- [60] A. V. Z. Brower, "Phylogenetic relationships among the Nymphalidae (Lepidoptera) inferred from partial sequences of the wingless gene," *Proceedings of the Royal Society B*, vol. 267, no. 1449, pp. 1201–1211, 2000.
- [61] P. R. Ehrlich, "The comparative morphology, phylogeny, and higher classification of the butterflies (Lepidoptera: Papilionoidea)," *University of Kansas Science Bulletin*, vol. 39, pp. 305–370, 1958.
- [62] A. Y. Kawahara, "Phylogeny of snout butterflies (Lepidoptera: Nymphalidae: Libytheinae): combining evidence from the morphology of extant, fossil, and recently extinct taxa," *Cladistics*, vol. 25, no. 3, pp. 263–278, 2009.
- [63] X. Feng, D. F. Liu, N. X. Wang, C. D. Zhu, and G. F. Jiang, "The mitochondrial genome of the butterfly *Papilio xuthus* (Lepidoptera: Papilionidae) and related phylogenetic analyses," *Molecular Biology Reports*, vol. 37, no. 8, pp. 3877–3888, 2010.
- [64] F. Qin, G. F. Jiang, and S. Y. Zhou, "Complete mitochondrial genome of the *Teinopalpus aureus guangxiensis* (Lepidoptera: Papilionidae) and related phylogenetic analyses," *Mitochondrial DNA*, vol. 23, no. 2, pp. 123–125, 2012.
- [65] M. J. Kim, R. A. Kang, H. C. Jeong, K. G. Kim, and I. Kim, "Reconstructing intraordinal relationships in Lepidoptera using mitochondrial genome data with the description of two newly sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera: Lycaenidae)," *Molecular Phylogenetics and Evolution*, vol. 61, no. 2, pp. 436–445, 2011.
- [66] X. M. Qin, Q. X. Guan, D. L. Zeng, F. Qin, and H. M. Li, "Complete mitochondrial genome of *Kallima inachus* (Lepidoptera: Nymphalidae: Nymphalinae): comparison of *K. inachus* and *Argynnis hyperbius*," *Mitochondrial DNA*, vol. 23, no. 4, pp. 318–320, 2012.
- [67] J. Xia, J. Hu, G. P. Zhu, C. D. Zhu, and J. S. Hao, "Complete mitochondrial DNA sequence of the *Calinaga davidis*," *Acta Entomologica Sinica*, vol. 54, no. 5, pp. 555–565, 2011.
- [68] M. J. Kim, X. Wan, K. G. Kim, J. S. Hwang, and I. Kim, "Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae)," *African Journal of Biotechnology*, vol. 9, no. 5, pp. 735–754, 2010.
- [69] Y. J. Zhu, R. Fang, G. L. Zhou, J. Ye, and J. P. Yi, "The complete sequence determination and analysis of *Lymantria dispar* genome," *Plant Quarantine*, vol. 24, no. 4, pp. 6–11, 2010.