Front. Biol. 2015, 10(5): 458–472
DOI 10.1007/s11515-015-1371-1

RESEARCH ARTICLE

The complete mitogenome of *Lamproptera curia* (Lepidoptera: Papilionidae) and phylogenetic analyses of Lepidoptera

Xin-Min Qin (✉), Qing-Xin Guan, Hui-Min Li, Yu Zhang, Yu-Ji Liu, Dan-Ni Guo

Guangxi Key Laboratory of Rare and Endangered Animal Ecology, College of Life Science, Guangxi Normal University, Guilin 541004, China

Abstract **The complete mitochondrial genome sequence of *Lamproptera curia* was determined in the present study. Our findings showed that the mtDNA of *L. curia* had a typical organization of insect mitochrondrial DNA – being 15277 base pairs in length, it contained 13 protein-coding genes (PCGs), 2 rRNA genes, 22 tRNA genes, and a control region (CR). The newly determined sequence was used for phylogenetic analyses, together with those of 45 species of Lepidoptera published elsewhere, including sequences of three species of Diptera as outgroups. The phylogenetic trees were constructed using the concatenated amino acid and nucleotide sequences of the 13 protein-coding genes (PCGs) based on the maximum likelihood (ML) and Bayesian inference (BI) methods. Both BI and ML trees revealed a similar topology structure: (((((Bombycoidea + Geometroidea) + Noctuoide) + Pyraloidea) + (Papilionoidea + Hesperioidea)) + Tortricoidea). Furthermore, the phylogenetic analyses demonstrated that each of the 16 families belonged to a monophyletic group respectively. The results of molecular phylogeny from the present study were congruent with traditional classification based on morphology but failed to demonstrate the monophyly of Hesperiidae.**

Keywords *Lamproptera curia*, phylogeny, mitochondrial genome

Introduction

Insect mitochondrial DNA (mtDNA) is a circular DNA molecule 14–20 kb in size with 13 protein-coding genes (PCGs), two rRNA genes, 22 tRNA genes, and one A + T-rich region which contains the initiation sites for transcription and replication (Clayton, 1992; Wolstenholme, 1992; Boore, 1999). In recent years, the mitochondrial genome are popularly used in studies on phylogenetics, comparative and evolutionary genomics, population genetics, and molecular evolution. The advantages of the mitochondrial material over those of the nuclear DNA include its stability in maternal inheritance, limited recombination and lower rates of nucleotide substitution.

Lepidoptera is the 2nd largest order of insects, containing 47 superfamilies, 126 families and 250 subfamilies (Kristensen and Skalski, 1999). Thus far, complete mitochondrial genome (mitogenome) sequences have been determined in more than 60 species of Lepidoptera. Although many studies have focused on the phylogeny of Lepidoptera, there are still many relationships within the order that remain to be elucidated (Lavrov et al., 2000). Among the unanswered questions include the superfamilial relationships, particularly those within the Macrolepidoptera (Bombycoidea, Geometroidea, Noctuoidea, Papilionoidea, Lasiocampoidea, Mimallonoidea, Axioidea, Calliduloidea, Hedyloidea, Hesperoidea, and Drepanoidea). Minet (1991) and Nielsen (1989) proposed a sister relationship between Geometroidea and Papilionoidea, but the relationship of this group to other macrolepidopteran superfamilies remains uncertain, resulting in a trichotomy among the most speciose macrolepidopteran superfamilies. Some studies with comprehensive selection of lepidopteran species and multiple genes representing a substantial length of sequence information demonstrated several unconventional relationships, including one that was incompatible to the monophyletic Macrolepidoptera (Regier et al., 2009; Mutanen et al., 2010). More recently, Kim et al. (2011) formulated a refined hypothesis of a superfamilial relationship (((((Bombycoidea + Geometroidea) + Noctuoidea) + Pyraloidea) + Papilionoidea) + Tortricoidea), signifying the lack of support for the traditionally defined

Received June 7, 2015; accepted August 12, 2015
Correspondence: Xin-Min Qin
E-mail: xmqin@mailbox.gxnu.edu.cn

Macrolepidoptera.

Another major area of controversy is regarding the familial relationships within the true butterflies. Butterflies are commonly recognized to comprise somewhere between four and 14 families (Kristensen, 1976; Smart, 1989).The relationships (((Lycaenidae + Nymphalidae) + Pieridae) + Papilionidae) were recently revisited by studies utilizing a combination of morphological and molecular characters (Weller and Pashely, 1995; Wahlberg et al., 2005) or molecular data alone (Kim et al., 2010). However, more recent and extensive molecular phylogenetic studies of the Lepidoptera revealed a non-monophyletic Papilionoidea, when Hesperiidae was grouped within another macrolepidopteran family – Hedylidae within the superfamily (Regier et al., 2009; Mutanen et al., 2010).

In the present study, we aimed at determining the complete mitogenomic sequences of one additional papilionids *Lamproptera curia* and elucidating the organization of the genome. Our focus was on utilizing the concatenated amino acid and nucleotide sequences of 13 protein-coding genes (PCGs) for phyolgenetic analyses. The newly obtained sequences of *L. curia*, together with those corresponding sequences generated in previous studies from 45 other species of Lepidoptera were utilized in order to determine relationships within the order and to test relationships within Papilionoidea.

Materials and methods

Specimen collection and DNA extraction

An adult specimen of *Lamproptera curia* was collected from Lingui, Guangxi, China, in August 2012; preserved in 100% ethanol and stored at 4°C until DNA extraction. Total genomic DNA was isolated from the muscles of thorax or leg using a routine phenol/chloroform method (Zhou et al., 2007).

Primer design, PCR, and sequencing:The primers used for the amplification of complete mitogenomes in this study were based on sequences listed in Table 1. A few the exact numbers of universal PCR primers for short fragment amplifications of the *cox1*, *cox2,nd5* and *cytb* genes were synthesized based on sequences from a previous study (Simon et al., 1994). The remaining primers were designed based on the sequence alignment of the available complete lepidopteran mitogenomes using Primer Premier 5.0 software (Singh et al., 1998). The entire mitogenome of *L. curia* was amplified in six fragments (*cox1-cox3*, *cox3-nad5*, *nad5-nad4*, *nad4-cob*, *cob-rrnL*, *rrnL-cox1*) using long-PCR techniques with TaKaRa LATaq polymerase under the following cycling conditions: initial denaturation for five minutes at 95°C, followed by 30 cycles of 95°C for 50 s, 45 – 50°C for 50 s, 68°C for 2 min and 30 s; and a final extension step of 68°C for 10 min. The PCR products were visualized by electrophoresis on 1.0% agarose gel, then purified using a PCR purification kit from QIAGEN (QIAGEN, Germany) and sequenced directly with an ABI 3730 DNA Analyzer using a BigDye chemistry kit (Applied Biosystems, Inc., Carlsbad, CA, USA), in which the same PCR primers were used. All PCR products were sequenced from both strands. The resultant mitogenome sequence data were deposited into the GenBank database under the accession number KJ141168.

Sequence analysis and annotation

Sequence annotation was performed using the blast tools in NCBI website (http://blast.ncbi.nlm.nih.gov/Blast) and DNAStar package (DNAStar Inc. Madison, USA). The tRNA genes and their secondary structure were predicted using tRNAscan-SE software v.1.21(Lowe and Eddy, 1997). The PCGs and rRNAs were confirmed by sequence comparison with ClustalX1.8 software and NCBI BLAST search function (Altschul et al., 1990). Nucleotide composition and codon usage were calculated with DAMBE software (Xia and Xie, 2001).

Table 1 List of primers used to amplify and sequence the mitogenome of *L. curius*

Primer 1	Sequence (5′→3′)	Primer 1	Sequence (5′→3′)	Fragment
1A	GCTAAATTAAGCTTTTGGGTTCA	1B	CCCGGTAAAATTAAAATATAAACTTC	ND2-COI
2A	GGTCAACAAATCATAAAGATATTG	2B	TAAACTTCAGGGTGACCAAAAAAT	COI
3A	GGATCACCTGATATAGCATTCCC	3B	GAGACCATTACTTGCTTTCAGTCATC	COI-COII
4A	GAAATTTGTGGAGCTAATCATAG	4B	TCAACAAAATGTCAATATCA	COII-COIII
5A	GGTTTACGATGAGGAATAATTT	5B	TTACAATGAAAATGTAATG	COIII-ND3
6A	CATTACATTTTCATTGTAA	6B	TTCTGCTTTGGTTCATTCT	ND3-ND5
7A	GGTTTACGATGAGGAATAATTT	7B	TTAGGTTGAGATGGGTTAGG	COIII-ND5
8A	GCTAATTATGAATTTGATT	8B	GATACTCTTCATCATATA	ND5
9A	ATAATAACTCCAGCACAT	9B	GCTTATTCTTCAGTTGCTCA	ND5-ND4
10A	GAAGGAGGAGCTGCTATATTAG	10B	CCTCAAAATGATATTTGACCTC	ND4- Cytb
11A	TACGTTTTACCATGAGGTCAAATATC	11B	ACTTCTTTTCTTATGTTTTCAAAAC	Cytb
12A	CCGACCTGTTGAAGATCCTTAT	12B	TCAGATCAAGATGCCGATT	Cytb-12S
13A	AGGGTATCTAATCCTAGTTT	13B	TGGGGTATGAACCCAAAAGC	12S-ND2

Phylogenetic analysis

To reconstruct the phylogenetic relationship among lepidopteran insects, the complete mitogenomes of 46 Lepidoptera species(including that of *L.curia*)were obtained from the GenBank database (Table 2). These mitogenomes were divided into 7 Lepidopteran superfamilies within the Lepidopteran order. The mitogenomes of three diptera species were used as outgroups. The concatenated nucleotide and amino acid sequence of 13 protein-coding genes were used for constructing phylogenetic trees.

The phylogenetic trees were constructed using maximum likelihood (ML) (Abascal et al., 2007) and Bayesian inference (BI) (Yang and Rannala, 1997) methods. The ML analyses were conducted using PHYML (Guindon et al., 2005) under the following conditions: the proportion of invariable sites as "estimated," number of substitution rate categories as four, gamma distribution parameter as "estimated," and the starting tree as a BIONJ distance-based tree. The confidence values of the ML tree were evaluated via the bootstrap test with 500 iterations. The Bayesian analyses were performed using MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) with the partitioned strategy. The best fitting substitution model was selected as in the ML analysis. The MCMC analyses (with random starting trees) were run with one cold and three heated chains simultaneously for 1 000 000 generations sampled every 100 generations. Bayesianposterior probabilities were calculated from the sample points after the MCMC algorithm started to converge.

Results and discussion

Genome organization

The complete mtDNA sequence of *L. curia* was 15277 bp in

Table 2 list of the complete mitogenome in the phylogenetic analyses

Superfamily/Family		Species	GenBank accession No.
Papilionoidea			
	Nymphalidae	*Kallima inachus*	JN857943
		Acraea issoria	NC_013604
		Apatura ilia	NC_016062
		Apatura metis	NC_015537
		Argynnis hyperbius	NC_015988
		Calinaga davidis	NC_015480
		Hipparchia autonoe	NC_014587
		Sasakia charonda	NC_014224
		Libythea celtis	NC_016724
		Euploea mulciber	NC_016720
		Fabriciana nerippe	NC_016419
	Papilionidae	*Lamproptera curius*	KJ141168
		Papilio maraho	NC_014055
		Parnassius bremeri	NC_014053
		Teinopalpus aureus	NC_014398
		Troides aeacus	EU625344
	Lycaenidae	*Coreana raphaelis*	NC_007976
		Protantigius superans	NC_016016
		Spindasis takanonis	NC_016018
	Pieridae	*Pieris melete*	NC_010568
		Pieris rapae	NC_015895
Hesperioidea			
	Hesperiidae	*Ctenoptilum vasava*	NC_016704
Bombycoidea			
	Saturniidae	*Antheraea pernyi*	NC_004622
		Antheraea yamamai	NC_012739
		Eriogyna pyretorum	NC_012727
		Saturnia boisduvalii	NC_010613
		Samia cynthia ricini	JN215366
	Bombycidae	*Bombyx mandarina*	NC_003395
		Bombyx mori	NC_002355
	Sphingidae	*Manduca sexta*	NC_010266

(*Continued*)

Superfamily/Family		Species	GenBank accession No.
Noctuoide			
	Noctuidae	*Helicoverpa armigera*	NC_014668
		Sesamia inferens	NC_015835
	Notodontidae	*Phalera flavescens*	NC_016067
		Ochrogaster lunifer	NC_011128
	Arctiidae	*Hyphantria cunea*	NC_014058
	Lymantriidae	*Lymantria dispar*	NC_012893
	Pyraloidea		
Crambidae		*Chilo suppressalis*	HQ860290
		Cnaphalocrocis medinalis	NC_015985
		Diatraea saccharalis	NC_013274
		Ostrinia furnacalis	NC_003368
		Ostrinia nubilalis	NC_003367
	Pyralidae	*Corcyra cephalonica*	NC_016866
Tortricoidea			
	Tortricidae	*Adoxophyes honmai*	NC_008141
		Grapholita molesta	NC_014806
		Spilonota lechriaspis	NC_014294
Geometroidea			
	Geometridae	*Phthonandria atrilineata*	NC_010522
Culicoidea			
	Culicidae	*Anopheles gambiae*	NC_002084
Tephritoidea			
	Tephritidae	*Bactrocera oleae*	NC_005333
Ephydroidea			
	Drosophilidae	*Drosophila yakuba*	NC_001322

length (Table 3) and it consisted of 2 rRNAs, 22 tRNAs, 13 PCGs and one major non-coding A + T-rich region. As shown in Fig. 1 and in consistent with the case in many insect mitogenomes, the major strand of the DNA coded for a higher number of genes (9 PCGs and 14 tRNAs), whereas the minor strand coded a lesser number (4 PCGs, 8 tRNAs and 2 rRNA genes).

All PCGs in the *L. curia* mitogenome were initiated by typical ATN codons (seven with ATG, four with ATT, one with ATA), except the *cox1* gene which was tentatively designated by the CGA codon (Table1). Generally, the trinucleotide TTG was assumed to be the *cox1* start codon for some invertebrate taxa including insect species, such as *Pyrocoelia rufa* (Bae et al., 2004), *Caligula boisdnvalii* (Hong et al., 2008), and *Acraea issoria* (Hu et al., 2010).

Among the stop codons of 13 protein-coding genes, three kinds of codon were found in *L. curia* TAA (*ND2, ATPase8, ATPase6, COIII, ND4L, ND6, Cytb*); TAG (*ND1,ND3*); and incomplete stop codon T (*COI, COII, ND4,ND5*). Incomplete termination codons are frequently observed in most insect mitogenomes and, in fact, all the sequenced mitogenomes of lepidopteran insects to date contained such codons (Kim et al., 2009). This incomplete codon is often activated through post-transcriptional polyadenylation, in which two A residues are added to create the TAA terminator (Anderson et al., 1981; Ojala et al., 1981).

tRNA and rRNA genes

The 22 tRNAs varied from 61 [$tRNA^{Ser}$(AGN)] to 69 bp ($tRNA^{Met}$, $tRNA^{Ile}$, $tRNA^{Gln}$, $tRNA^{Lys}$) in size, and presented typical clover-leaf structure, with the unique exception of $tRNA^{Ser}$ (AGN), which lacked the dihydrouridine DHU) stem (Fig. 2). The *L. curia* tRNAs harbored a total of 24 pair mismatches in their stems, including six pairs in the DHU stems, three pairs in the amino acid acceptor stems, eight pairs in the TΨC stems and seven pairs in the anticodon stems, respectively. Among these 24 mismatches, 16 were G-U pairs which formed a weak bond in the secondary structure, the remaining eight were atypical pairings: one mismatch in the tRNAHis (C-U), one mismatch in the $tRNA^{Lys}$ (C-U), one mismatch in the $tRNA^{Ser}$(AGY) (U-C), one mismatch in the $tRNA^{Ile}$(U-U), 2 in the $tRNA^{Phe}$ (one U-A and one C-G), and 2 in the $tRNA^{Ser}$ (UCN) (2U-U) (Fig. 2). The number of mismatches in the *L. curia* tRNAs found by the present study was well within the range reported from previous studies for other lepidopteran insect tRNAs (Liu et al., 2008; Jiang et al., 2009; Kim et al., 2010). These tRNAs mismatches can be

Table 3 Summary of the mitogenome of *L. curius*

Gene	Direction	Location	Size (bp)	Spacer(+) Overlap(−)	Anti- codon	Start codon	Stop codon
$tRNA^{Met}$	F	1–69	69	+1	CAT		
$tRNA^{Ile}$	F	69–137	69	−3	GAT		
$tRNA^{Gln}$	R	141–209	69	−39	TTG		
ND2	F	249–1262	1014	+2		ATT	TAA
$tRNA^{Trp}$	F	1261–1325	65	+8	TCA		
$tRNA^{Cys}$	R	1318–1383	66		GCA		
$tRNA^{Tyr}$	R	1384–1450	67	−2	GTA		
COI	F	1453–2983	1531			CGA	T–
$tRNA^{Leu}$(UUR)	F	2984–3050	67		TAA		
COII	F	3051–3726	676			ATG	T–
$tRNA^{Lys}$	F	3727–3795	69	+1	CTT		
$tRNA^{Asp}$	F	3795–3859	65		GTC		
ATPase8	F	3860–4021	162	+7		ATT	TAA
ATPase6	F	4015–4692	678	+1		ATG	TAA
COIII	F	4692–5483	792	−3		ATG	TAA
$tRNA^{Gly}$	F	5487–5552	66	+3	TCC		
ND3	F	5550–5906	357			ATA	TAG
$tRNA^{Ala}$	F	5907–5970	64	+1	TGC		
$tRNA^{Arg}$	F	5970–6032	63	−1	TCG		
$tRNA^{Asn}$	F	6034–6099	66	−1	GTT		
$tRNA^{Ser}$(AGY)	F	6101–6160	60		TCT		
$tRNA^{Glu}$	F	6161–6227	67	+2	TTC		
$tRNA^{Phe}$	R	6226–6291	66	+6	GAA		
ND5	R	6286–8008	1723	−18		ATT	T–
$tRNA^{His}$	R	8027–8093	67		GTG		
ND4	R	8094–9432	1339			ATG	T–
ND4L	R	9433–9723	291	−2		ATG	TAA
$tRNA^{Thr}$	F	9726–9789	64		TGT		
$tRNA^{Pro}$	R	9790–9854	65	−2	TGG		
ND6	F	9857–10387	531	+1		ATT	TAA
Cytb	F	10387–11535	1149	+2		ATG	TAA
$tRNA^{Ser}$(UCN)	F	11534–11598	65	−16	TGA		
ND1	R	11615–12553	939	−1		ATG	TAG
$tRNA^{Leu}$(CUN)	R	12555–12622	68		TAG		
16SrRNA	R	12623–13956	1334				
$tRNA^{Val}$	R	13957–14022	66		TAC		
12SrRNA	R	14023–14807	785				
A + T-rich region		14808–15277	469				

corrected through RNA editing mechanisms, which are well known for arthropod mtDNA (Lavrov et al., 2000).

As in all other insect mitogenome sequences, two rRNA genes (*rrnL* and *rrnS*) were detected in *L. curia*. The lrRNA and srRNA genes of the *L. curia* mitogenome were 1334 and 785 bp in length, respectively. They were located between $tRNA^{Leu}$ (CUN) and $tRNA^{Val}$ and between $tRNA^{Val}$ and the A + T-rich region, respectively (Fig. 1). The length of the lrRNA gene was determined to be 1334 bp, which was very well within the size range of 470 bp in *Bemisia tabaci* (Thao et al., 2004) to 1426 bp in *Hyphantria cunea* (Liao et al., 2010), as observed in the corresponding gene of other insects sequenced previously. Similarly, the length of the srRNA gene was determined to be 785 bp, which again was well within the size range of 434 bp in *Ostrinia nubilalis* (Clary and Wolstenholme, 1985) to 827 bp in *Locusta migratoria*, as observed in the corresponding genes of other completely sequenced insects (Flook et al., 1995).

A + T-rich region

The A + T-rich region of *L. curia* was located between the srRNA and $Trna^{Met}$ genes. The 469 bp long A + T-rich region exhibited the highest A + T contents (89.8%) than any other

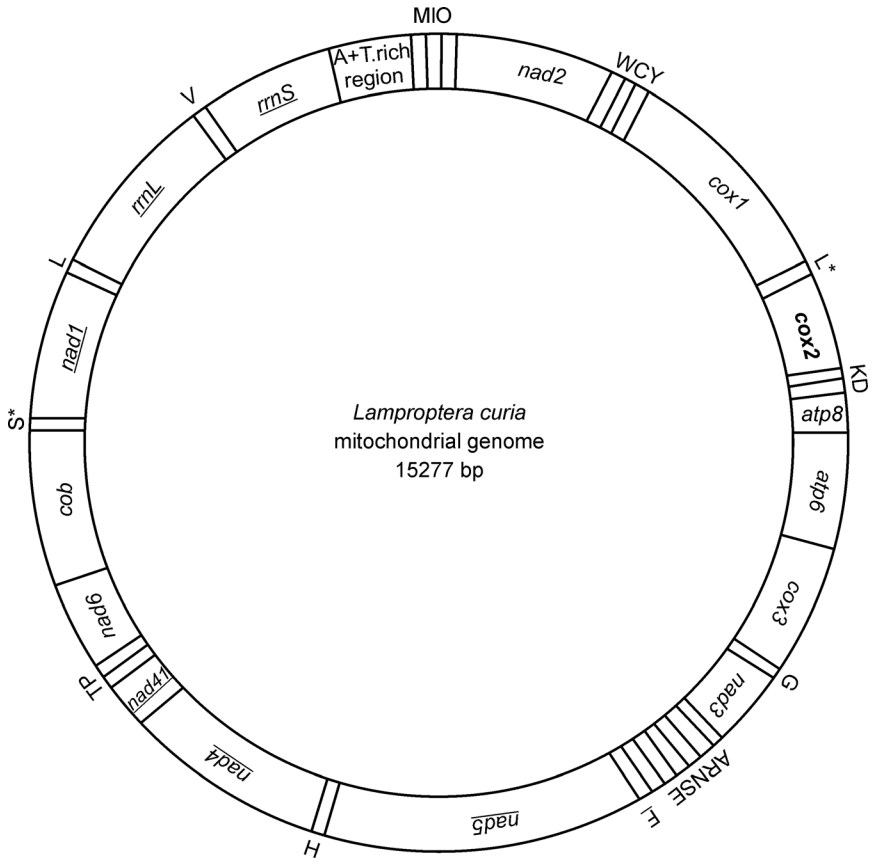


Figure 1 Circular map of the mitochondrial genome of *Lamproptera curia*. *cox1 – 3*: cytochrome oxidase subunit 1 – 3 genes; *atp6*, *atp8*: ATP synthase subunits 6 and 8 genes; *cob*: cytochrome oxidase b gene; *nad1 – 6* and *nad4L*:NADH dehydrogenase subunits 1 – 6 and 4L. tRNA genes are denoted 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 and with underlines of counter clockwise.

regions of *L. curia* mitogenome. A conserved sequence was found in the 5′-end of A + T-rich region, which consisted of 2 repeats of a unit (5′-ATAGATTTTTTTTTTTTTTT-3′). Additionally, other short microsatellite-like repeat regions were also observed throughout the A + T-rich region, without noticeable macro-repeats.

Phylogenetic analysis

The newly obtained mitogenomes of *L. curia* was used for phylogenetic analyses, together with the mitogenomes of 45 other lepidopteran, representing seven lepidopteran superfamilies (Papilionoidea, Hesperioidea, Bombycoidea, Geometroidea, Noctuoidea, Pyraloidea and Tortricoidea). The phylogenetic analyses were carried out using Bayesian Inference (BI) and maximum likelihood (ML) algorithms for the concatenated nucleotide and amino acid sequences of 13 protein-coding genes. Four phylogenetic trees were constructed but they all produced one very similar topology structure, involving seven superfamilies clustered into three clades (Figs. 3–6). The first clade included the families of Papilionoidea and Hesperioidea. The second one included the superfamilies of Bombycoidea, Geometroidea, Noctuoidea and Pyraloidea. In particular, Bombycoidea and Geometroidea in the second clade, was first clustered into a small branch of (Bombycoidea + Geometroidea). This branch of (Bombycoidea + Geometroidea) was in turn made up for a larger clade with the Noctuoidea. Finally the sub-branch was clustered with Pyraloidea. The Tortricoidea species constituted the third clade.

Papilionoidea: Kristensen (1976) suggested a close relationship between the Nymphalidae and Lycaenidae groups with the Pieridae based on morphological characteristics alone. The proposal was further supported by subsequent studies using a combination of morphological and molecular characteristics (Wahlberg et al., 2005; Kim et al., 2011).

In the present study, when the Hesperiidae was excluded, the molecular phylogenetic relationship among the true butterfly families (Nymphalidae, Lycaenidae, Pieridae and Papilionidae) was demonstrated as (((Lycaenidae + Pieridae) + Nymphalidae) + Papilionidae), but with the Lycaenidae being identified as a sister group of Pieridae with low nodal support at 43% when the ML method was used based on the amino acid sequence data (Fig. 4). However,the molecular phylogenetic relationship among true butterfly families was

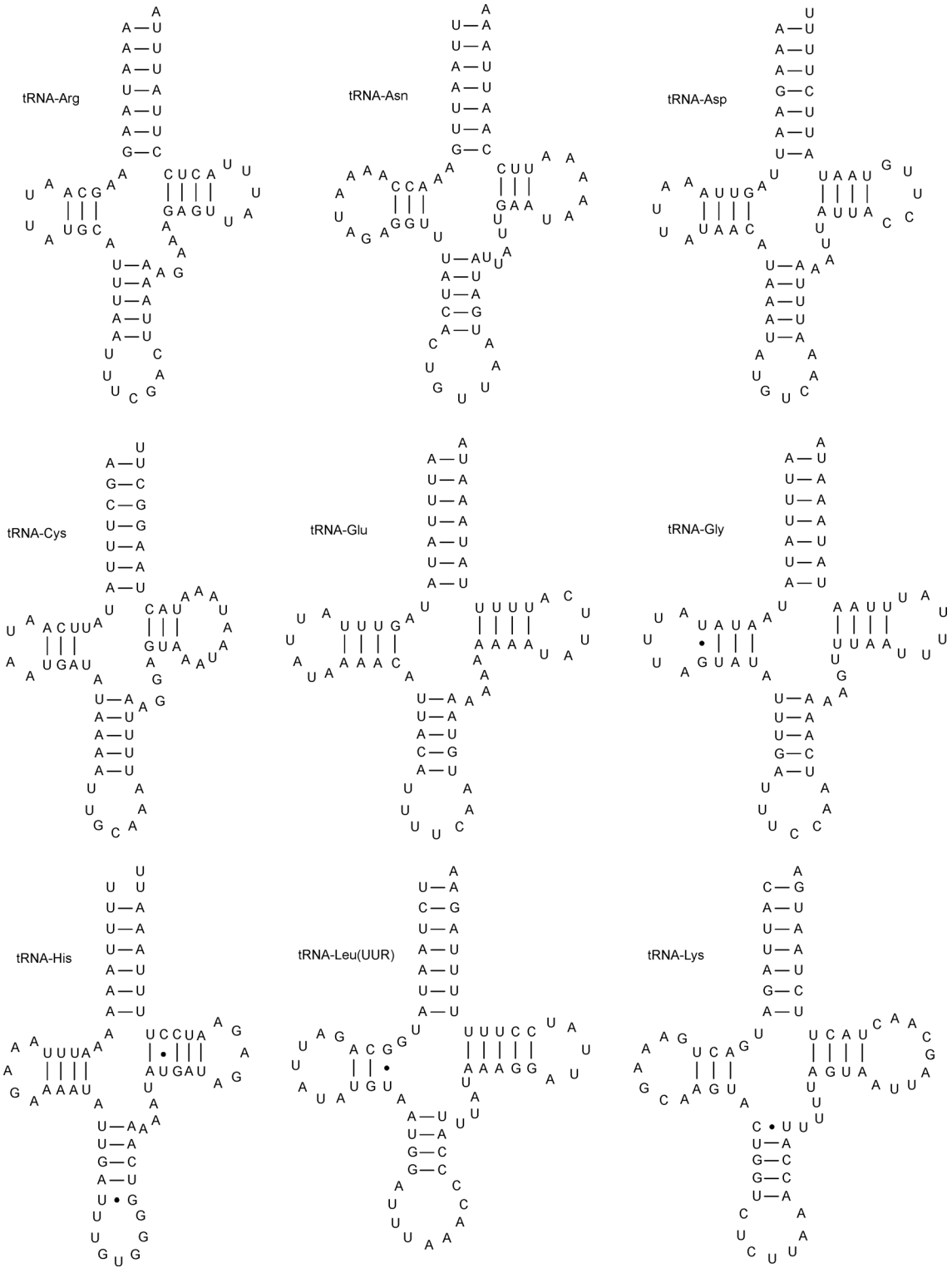
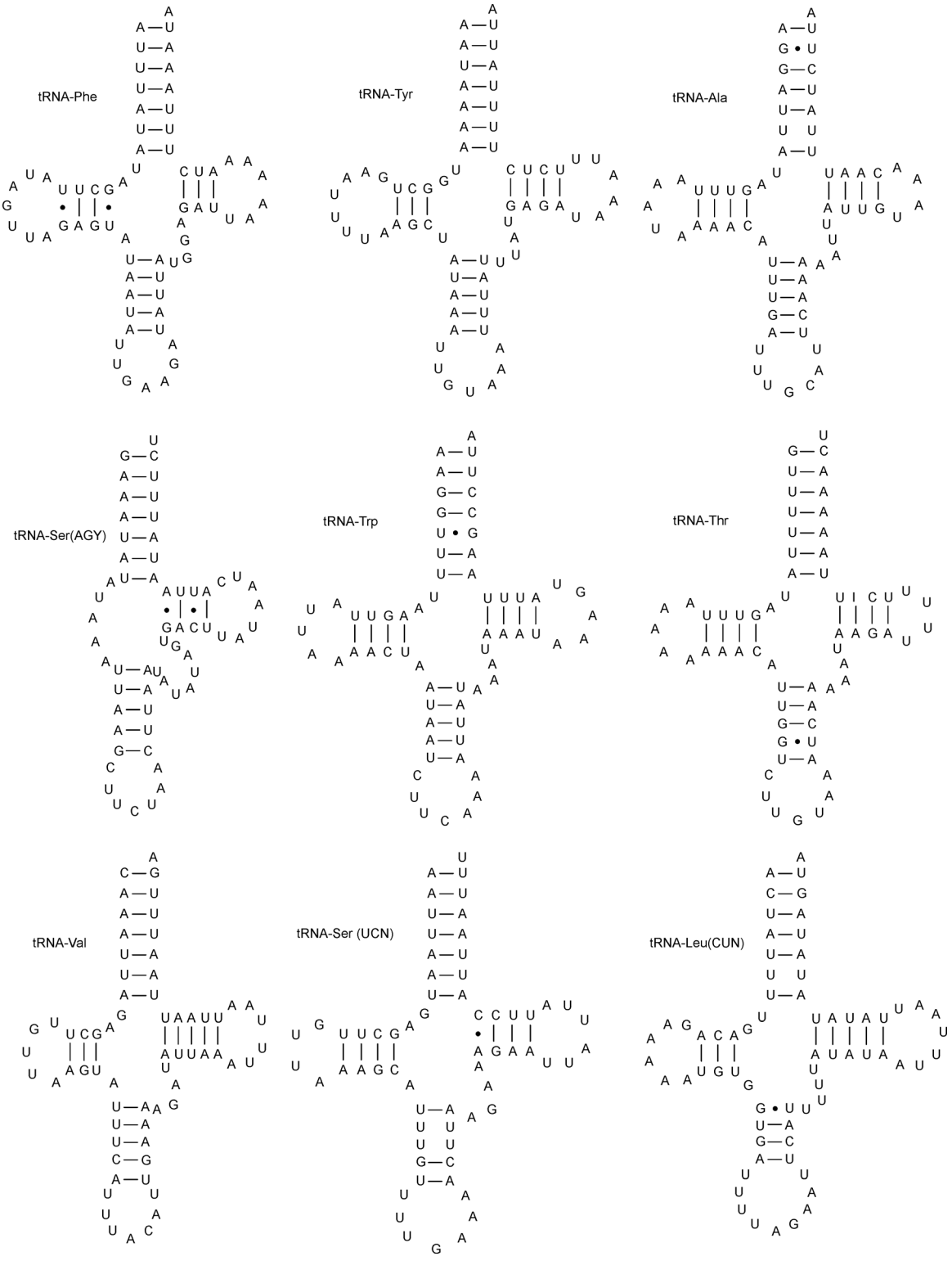
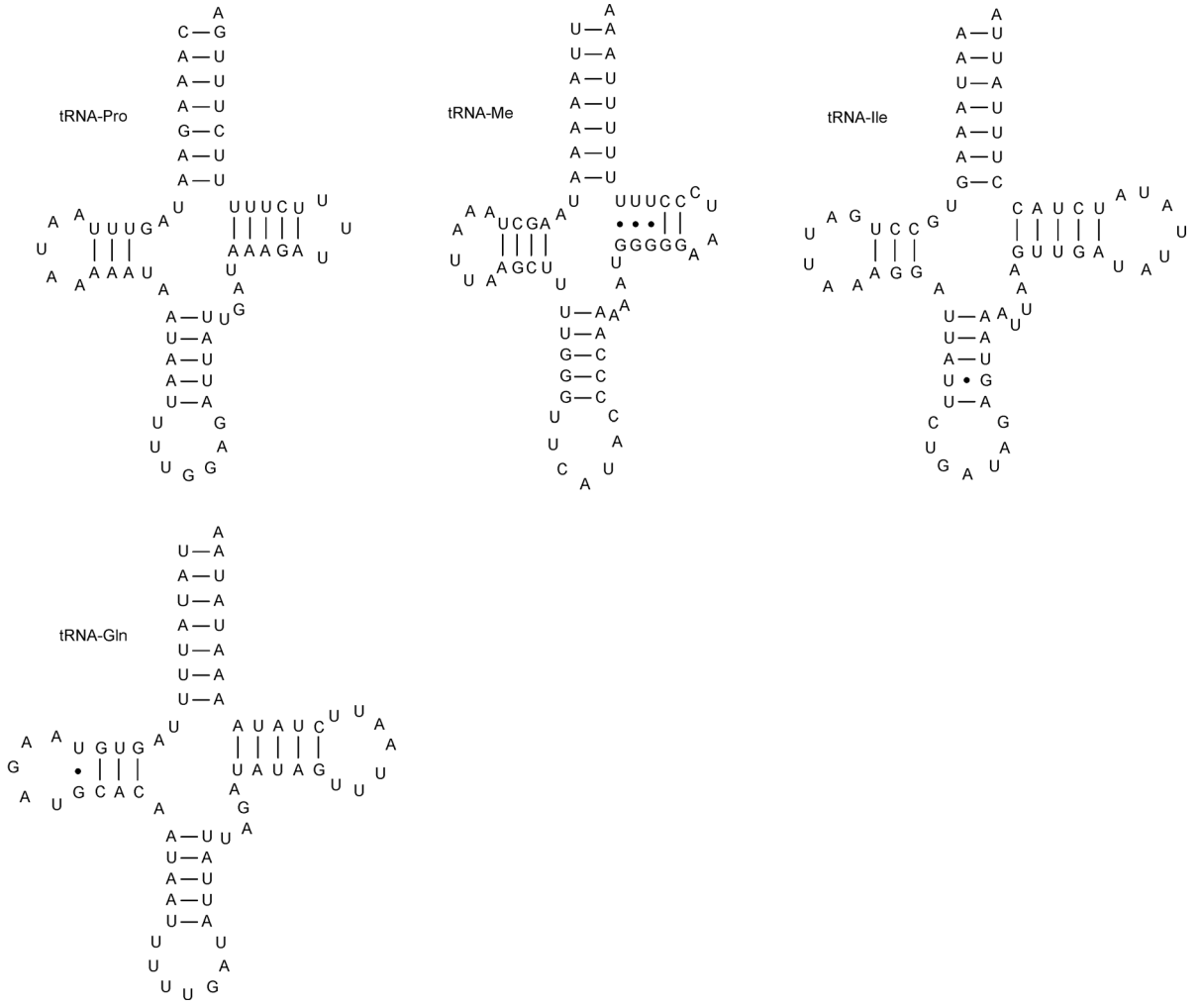


Figure 2 Predicted secondary structures for the 22 tRNA genes of *L. curius*.



(Fig. 2 continued)



(Fig. 2 continued)

showed in a same structure in the three remaining phylogenetic trees as (((Nymphalidae + Lycaenidae) + Pieridae) + Papilionidae) (Figs. 3, 5 and 6). Nevertheless, this relationship was well in agreement with similar findings of previous studies (Kristensen, 1976; Wahlberg et al., 2005; Kim et al., 2010). Additionally,the sister group relationship existed between Nymphalidae and Lycaenidae was found to have high nodal supports (ML: 82%; BI: 0.98 and 1.00).

Hesperioidea: Morphological classification currently divides the true butterfly into 2 superfamilies: Papilionoidea and Hesperioidea. Within the Hesperioidea, there is only one family i.e. "Hesperiidae" (Harvey, 1991; Ackery et al., 1999). However, the above mentioned morphological classification is yet to be supported by molecular phylogenetic studies. In the study of Wahlberg et al.(2005) Hesperioidea was found to be within Papilionoidea using Bayesian Inference (BI), having a sister relationship with((Pieridae + Nymphalidae + (Riodinidae + Lycaenidae)). Similar conclusions were also given by Regier et al.(2008) and Mutanen et al.(2010) in their respective studies:Hesperioidea as a sister branch of ((Pieridae + Lycaenidae) + Nymphalidae))existed within Papilionoidea. Likewise, both phylogenetic trees constructed using either ML and BI methods in the present study showed a finding differing to the results of morphological studies. Again, Hesperioidea was found to be within the superfamily of Papilionoidea. All these suggest to the authors that further studies in molecular phylogenetics are warranted in order to determine the evolution position of Hesperioidea.

Bombycoidea: All four phylogenetic trees generated in the present study support a relationship of ((Saturniidae + Sphingidae) + Bombycidae). This relationship was in total agreement with previous findings, based on either morphological studies (Minet, 1991) or molecular phyolgenetic analyses (Kawahara et al., 2009).

Noctuoide: The findings of the present study reinforced the notion that among the Noctuoid, Lymantriidae met first with Arctiidae, subsequently with Noctuidae and lastly with Notodontidae. This convergent structure(((Lymantriidae +

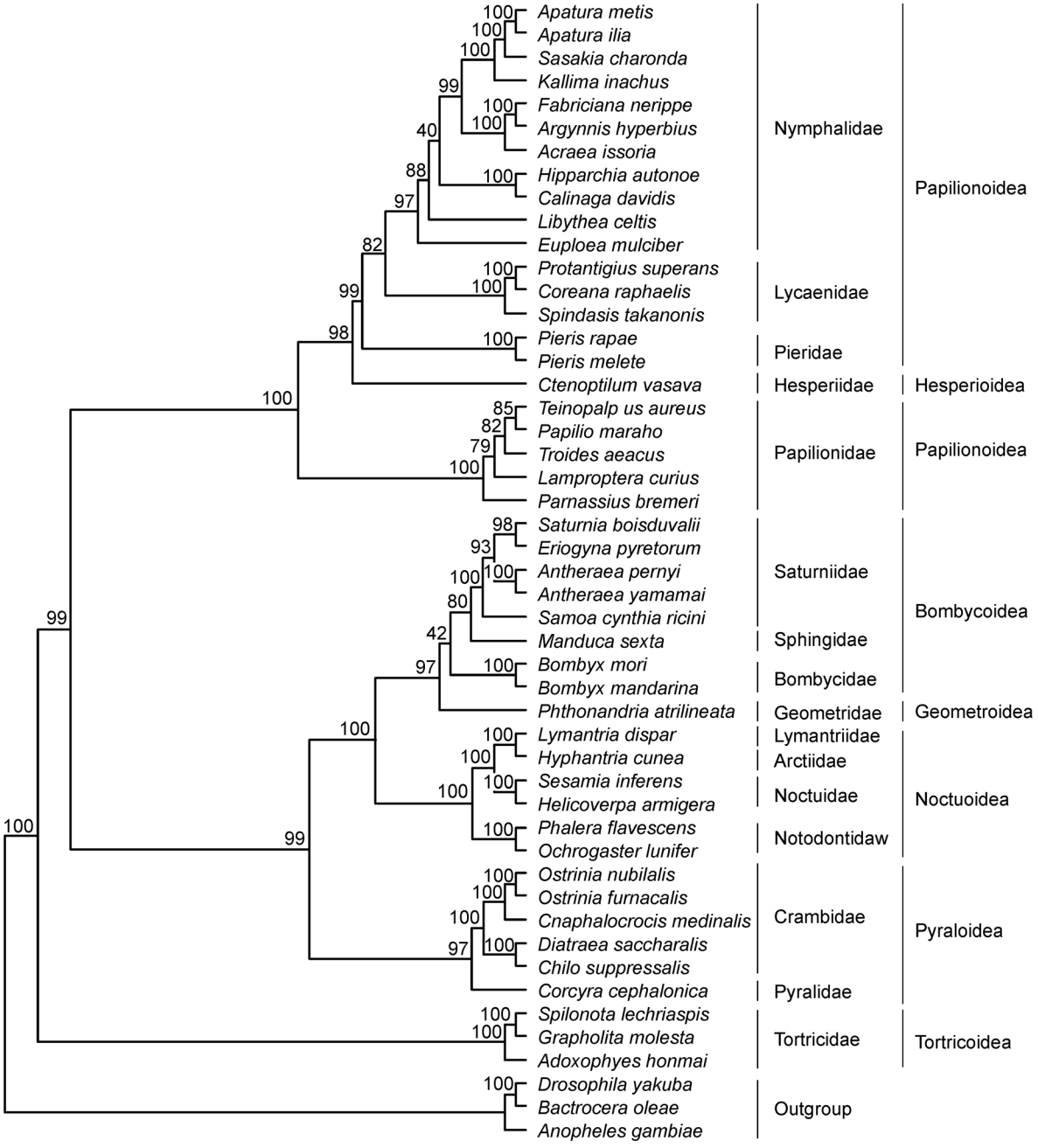


Figure 3 ML trees based on the nucleotide sequences of the concatenated 13 protein coding genes.

Arctiidae) + Noctuidae) + Notodontidae)was highly evident by both phylogenetic trees obtained using ML method and BI approach with 100% nodal support. Furthermore, our results were also in agreement with those of morphological studies (Yin et al., 2008).

Pyraloidea: Currently, the proposal of Munroe and Solis (1999) on morphological classification of Pyraloidea, i.e. Pyraloidea included Crambidae and Pyralidae; are widely accepted in the field. Based on the phyogenetic trees of ML and BI generated by the present study, Crambidae and Pyralidae were sister branches of each other, providing yet again more evidence bolstering the above mentioned suggestion.

Tortricoidea: Morphological classification placed Tortricoidea superfamily as one family(Tortricidae) consisting of 3 sub-families (Razowski, 1976; Horak, 1999). Based on our analyses, 3 species under the super-family of Tortricoidea converged to a branch as showed in both ML tree and BI tree. Among, two species from the sub-family of Olethreutinae: *Grapholita molesta* and *Spilonota lechriaspis* joined one another first. This branch was then joined by *Adoxophyes honmai* from the sub-family of Tortricinae. The nodal support

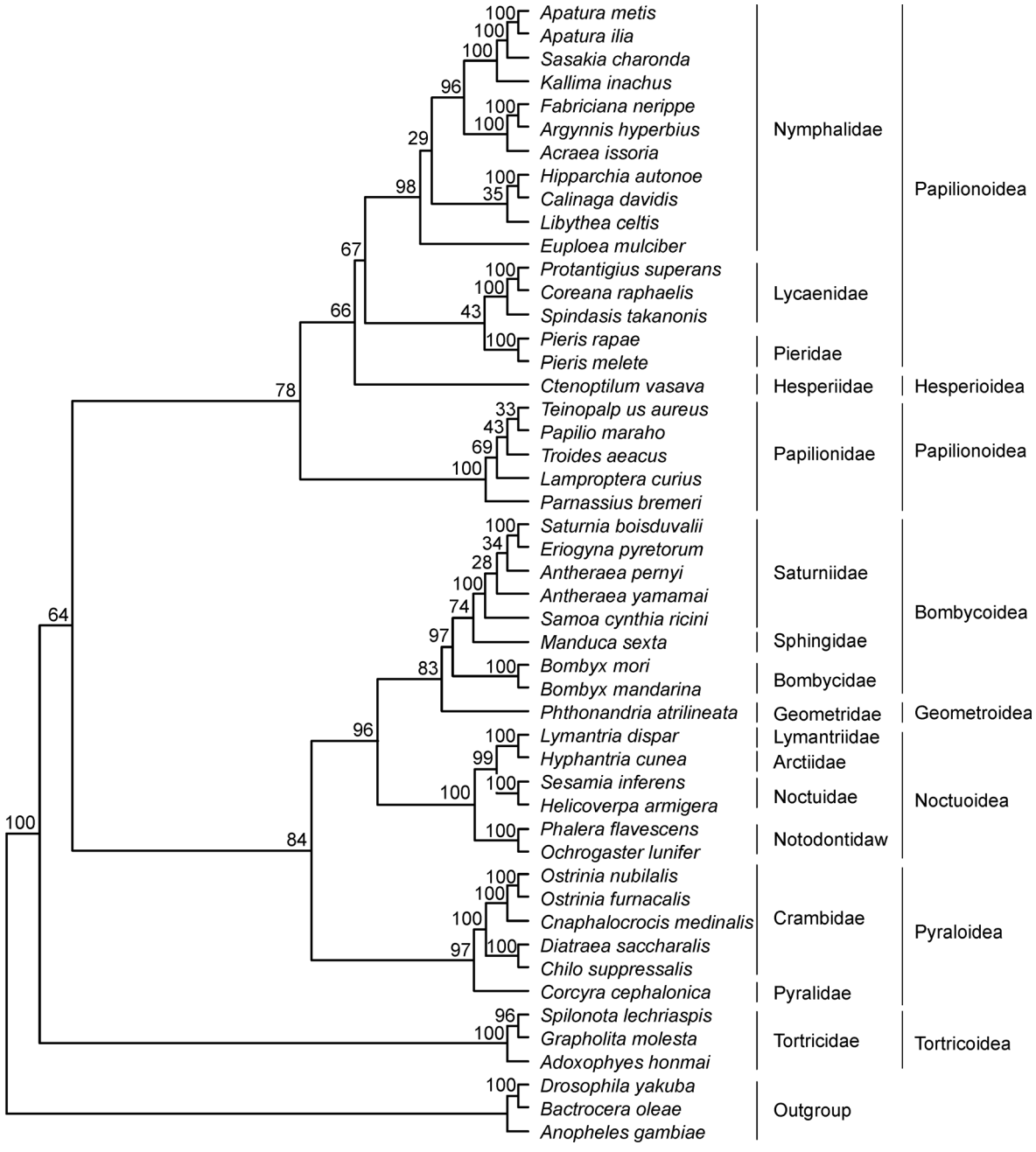


Figure 4 ML trees based on the amino acid sequences of the concatenated 13 protein coding genes.

obtained for the above structures was significantly high with ML ranged from 96 to 100% (Figs. 3 and 4) and BI at 100% (Figs. 5 and 6) respectively. The results of our research supported the findings of previous morphological studies.

Geometroidea: Phylogenetic trees showed that Geometroidea has the closest relationship with Bombycoidea, as compared to other super family.

Molecular phylogenetic analyses based on mitogenome in recent years have become a hot topic of researches on molecular classification of lepidopteran insect, with good progresses made. Lee et al. (2006) used 7 protein-coding genes of the lepidoptera mtogenome for their phylogenetic analyses, with all findings in support of single lineage (monophyletic) of Lepidoptera. The internal relationship among the classification groups under Lepidoptera(Aoditrysia(Obtectomera(Macrolepidoptera))), was found by the same study to be in consistent with that defined by the traditional classification. In 2010, Feng et al. (2010) carried out a similar study on the phylogenetic relationship among the insects of Lepidoptera based on 12 PCGs. Their study supported the current theory of morphological studies, i.e., all of the families: Bombycoidea, Pyraloidea and Papilionoidea

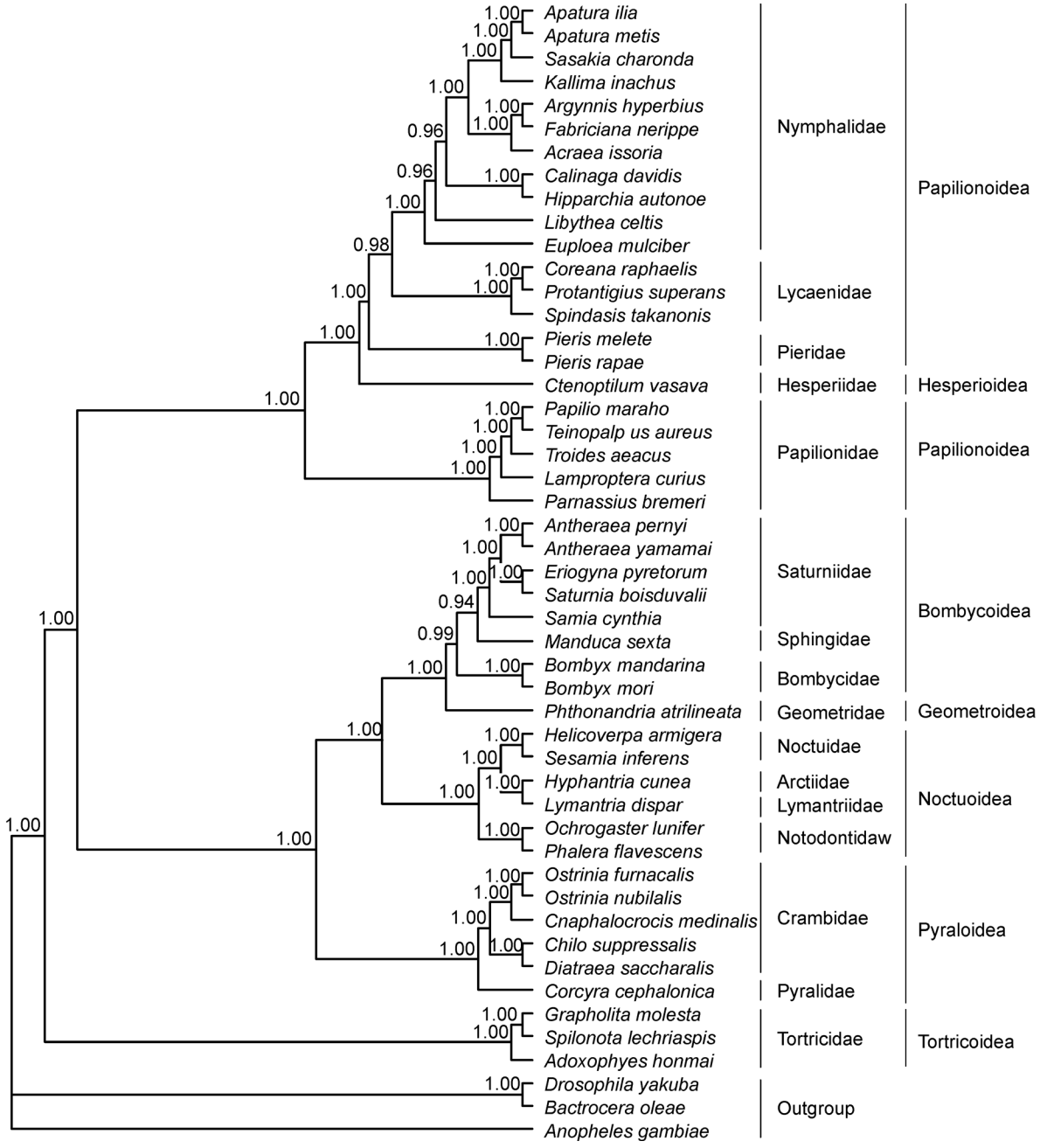


Figure 5 BI trees based on the nucleotide sequences of the concatenated 13 protein coding genes.

is monophyletic.

The findings of the present study is in support of a monophyletic origin for Bombycoidea, Noctuoide, Geometroidea, Pyraloidea and Tortricoidea respectively. Six superfamilies converged to one signle branch of Obtectomera,with the exception of Tortricoidea. This finding is in agreement with traditional classification.

In conclusion, we have sequenced the complete mitogenomes of *L. curia.* Phylogenetic analyses principally yielded the relationships (((((Bombycoidea + Geometroidea) + Noctuoide) + Pyraloidea) + (Papilionoidea + Hesperioidea)) + Tortricoidea). Within the true butterfly families, the relationships (Nymphalidae + Lycaenidae + Pieridae + Hesperiidae + Papilionidae)were supported by the majority of data sets. Hesperioidea was found to be within the superfamily of Papilionoidea. To further evaluate the Hesperioidea phylogenetic relationships among the true butterflies, a larger number of complete mitogenome sequences that encompass more of the Hesperioidea mitogenome will be required.

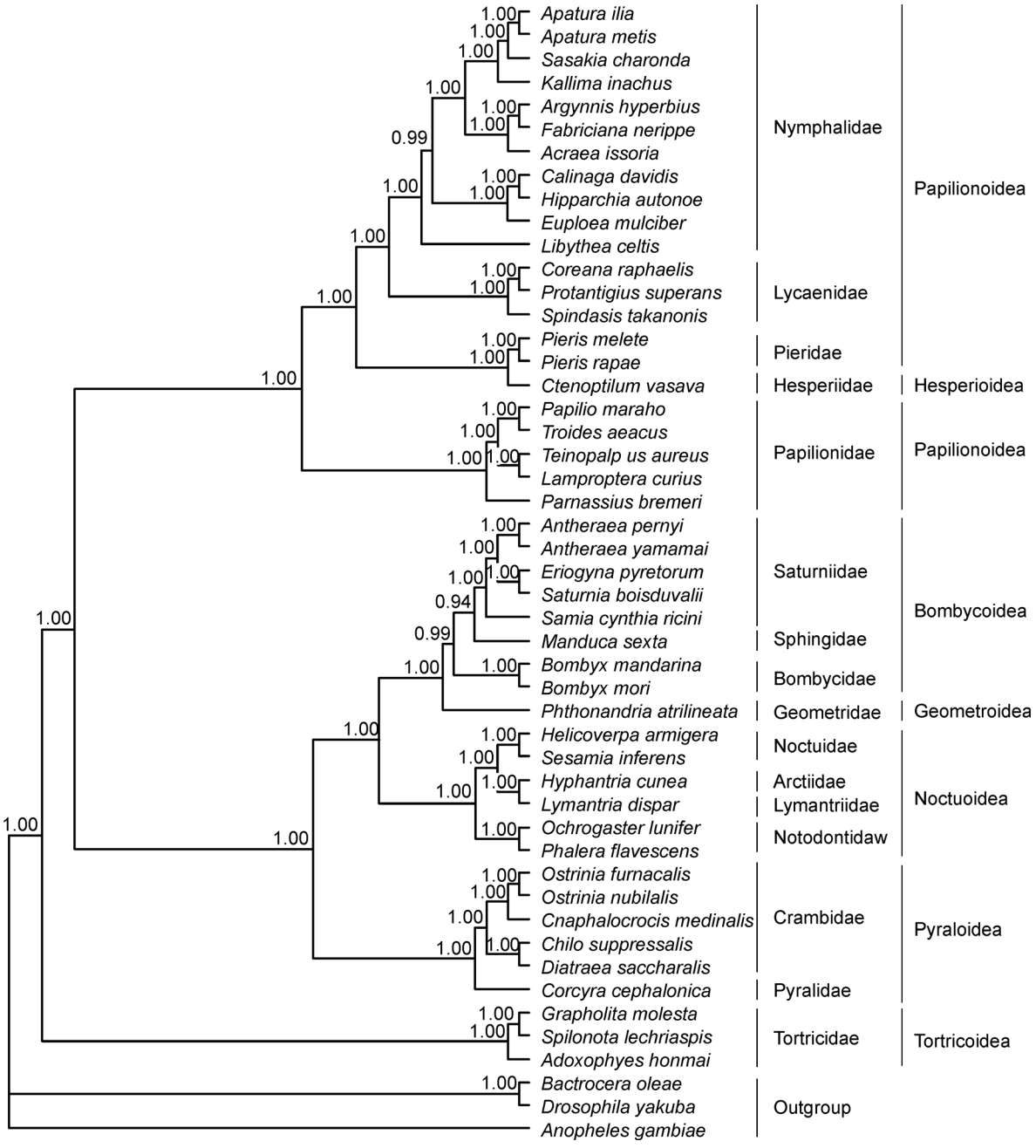


Figure 6 BI trees based on the amino acid sequences of the concatenated 13 protein coding genes.

Author contributions

Qin Xin-Min conceived the project. Guan Qing-Xin performed mitochondrial genome sequencing. Li Hui-Min performed genomic DNA samples. Zhang Yu, Liu Yu-Ji and Guo Dan-Ni analyzed the data.

Acknowledgements

The project was supported by Guangxi Key Laboratory of Rare and Endangered Animal Ecology Research Foundation (1401z006), Guangxi Normal University, China.

Compliance with ethics guidelines

Xin-Min Qin, Qin-Xin Guan, Hui-Min Li, Yu Zhang, Yu-Ji Liu and Dan-Ni Guo declare that no conflict of interest.

All institutional and national guidelines for the care and use of laboratory animals were followed.

References

Abascal F, Posada D, Zardoya R (2007). MtArt: a new model of amino

acid replacement for arthropoda. Mol Biol Evol, 1: 1–5

Ackery P R, De Jong R, Vane-Wright R I (1999). The butterflies: Hedyloidea, Hesperoidea, and Papilionoidea. In: Kristensen NP (Ed.), The butterflies: Hedyloidea, Hesperoidea, and Pilionoidea. De Gruyter, Berlin, 263–300

Altschul S F, Gish W, Miller W, Myers E W, Lipman D J (1990). Basic local alignment search tool. J Mol Biol, 3(3): 403–410

Anderson S, Bankier A T, Barrell B G, de Bruijn M H, Coulson A R, Drouin J, Eperon I C, Nierlich D P, Roe B A, Sanger F, Schreier P H, Smith A J, Staden R, Young I G (1981). Sequence and organization of the human mitochondrial genome. Nature, 290(5806): 457–465

Bae J S, Kim I, Sohn H D, Jin B R (2004). The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization, and phylogenetic analysis with other insects. Mol Phylogenet Evol, 3(3): 978–985

Boore J L (1999). Animal mitochondrial genomes. Nucleic Acids Res, 8 (8): 1767–1780

Clary D O, Wolstenholme D R (1985). The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization and genetic code. J Mol Evol, 22(3): 252–271

Clayton D A (1992). Transcription and replication of animal mitochondrial DNA. Int Rev Cytol, 141: 217–232

Feng X, Liu D F, Wang N X, Zhu C D, Jiang G F (2010). The mitochondrial genome of the butterfly *Papilio xuthus* (Lepidoptera: Papilionidae) and related phylogenetic analyses. Mol Biol Rep, 8(8): 3877–3888

Flook P K, Rowell C H F, Grellissen G (1995). The sequence organisation, and evolution of the *Louocsta migratoria* mitochondrial genome. J Mol Evol, 6: 928–941

Guindon S, Lethiec F, Duroux P, Gascuel O (2005). PHYML Online-a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res, 33(Web Server): W557–W559

Harvey D J (1991). Higher classification of the Nymphalidae [A], Appendix B. In: Nijhout HF. Higher classification of the Nymphalidae. Washington DC: Smithsonian Institution Press, pp 200–273

Hong M Y, Lee E M, Jo Y H, Park H C, Kim S R, Hwang J S, Jin B R, Kang P D, Kim K G, Han Y S, Kim I (2008). Complete nucleotide sequence and organization of the mitogenome of the silk moth *Caligula boisduvalii* (Lepidoptera: Saturniidae) and comparison with other lepidopteran insects. Gene, 413(1–2): 49–57

Horak M (1999). The Tortricoidea. In: Kristensen N P. Lepidoptera: Motha and Butterflies. 1. Evolution,Systematics, and Biogegraphy. Handbook of Zoology Vo1.IV, Part 35: 199–215

Hu J, Zhang D X, Hao J S, Huang D Y, Cameron S, Zhu C D (2010). The complete mitochondrial genome of the yellow coaster, *Acraea issoria* (Lepidoptera: Nymphalidae: Heliconiinae: Acraeini):sequence, gene organization and a unique tRNA translocation event. Mol Biol Rep, 7 (7): 3431–3438

Jiang S T, Hong G Y, Yu M, Li N, Yang Y, Liu Y Q, Wei Z J (2009). Characterization of the complete mitochondrial genome of the giant silkworm moth, *Eriogyna pyretorum* (Lepidoptera:Saturniidae). Int J Biol Sci, 4: 351–365

Kawahara A Y, Mignault A A, Regier J C, Kitching I J, Mitter C (2009). Phylogeny and Biogeography of *Hawk moths* (Lepidoptera: Sphingidae): Evidence from Five Nuclear Genes. PLoS ONE, 5: 1–11

Kim M I, Beak J Y, Kim M J (2009). Complete nucleotide sequence and organization of the mitogenome of the red-spotted Apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. Mol Cell, 4(4): 347–363

Kim M J, Kang A R, Jeong H C, Kim K G, Kim I (2011). Reconstructing intraordinal relationships in Lepidoptera using mitochondrial genome data with the description of two newly sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera: Lycaenidae). Mol Phy Evol, 2(2): 436–445

Kim M J, Wan X L, Kim K G, Hwang J S, Kim I (2010). Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). Afr J Biotechnol, 5: 735–754

Kristensen N P (1976). Remarks on the family-level phylogeny of butterflies (Insecta Lepidoptera, Rhopalocera). Zeit Zool Syst Evol Forsch, 14(1): 25–33

Kristensen N P (1999). 11 Evolution, Systematics, and Biogeography, Vol IV. In: Kristensen NP. Lepidoptera: Moths and Butterflies. Berlin and New York: DeGruyter, pp: 491.

Kristensen N P, Skalski A W (1999). Phylogeny and palaeontology. In: Kristensen N P (Ed.), Handbuch der Zoologie, vol. IV, Arthropoda: Insecta, Part 35, Lepidoptera, Moths and Butterflies, vol. 1: Evolution, Systematics, and Biogeography. Walter de Gruyter, Berlin & New York, pp. 7–25

Lavrov D V, Brown W M, Boore J L (2000). A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. Proc Natl Acad Sci USA, 25(25): 13738–13742

Lee E S, Shin K S, Kim M S, Park H, Cho S, Kim C B (2006). The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidse). Gene, 373: 52–57

Liao F, Wang L, Wu S, Li Y P, Zhao L, Huang G M, Niu C J, Liu Y Q, Li M G (2010). The mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). Int J Biol Sci, 2: 172–186

Liu Y, Li Y, Pan M, Dai F, Zhu X, Lu C, Xiang Z (2008). The complete mitochondrial genome of the Chinese oak silkmoth, *Antheraea pernyi* (Lepidoptera: Saturniidae). Acta Biochim Biophys Sin (Shanghai), 8(8): 693–703

Lowe T M, Eddy S R (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res, 5(5): 955–964

Minet J (1991). Tentative Reconstruction of the ditrysian phylogeny (Lepidiptera, Gloassata). Entomol Scand, 1(1): 69–95

Munroe E, Solis M A (1999). The Pyraloidea. In: Kristensen NP. Lepidoptera: Moths and Butterflies. I. Evolution, Systematics, and Biogeography. Handbook of Zoology, 35: 233–256(Vol IV. In: Kristensen NP. Lepidoptera: Moths and Butterflies. Berlin and New York: DeGruyter, pp: 233–256

Mutanen M, Wahlerg N, Kaila L (2010). Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. Proc Biol Sci, 277(1695): 2839–2848

Nielsen E S (1989) Phylogeny of major lepidopteran groups. In: (eds) The Hierarchy of Life. Amsterdam, Elsevier. pp. 281–294

Ojala D, Montoya J, Attardi G (1981). tRNA punctuation model of RNA proeessing in Human mitochondria. Nature, 290(5806): 470–474

Razowski J (1976). Phylogeny and system of Tortricidae (Lepidoptera). Acta zool Cracov, 5: 73–120

Regier J C, Cook C, Mitter C, Hussey A (2008). A phylogenetic study of the 'bombycoid complex' (Lepidoptera) using five protein-coding nuclear genes, with comments on the problem of macrolepidopteran phylogeny. Syst Entomol, 33(1): 175–189

Regier J C, Zwick A, Cummings M P, Kawahara A, Cho S, Weller S, Roe A, Baixeras J, Brown J W, Parr C, Davis D R, Epstein M, Hallwachs W, Hausmann A, Janzen D H, Kitching I J, Solis M A, Yen S H, Bazinet A L, Mitter C (2009). Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): an initial molecular study. BMC Evol Biol, 9(1): 280

Ronquist F, Huelsenbeck J P (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 12 (12): 1572–1574

Simon C, Frati F, Bekenbach A (1994). Evolution, weighting, and phylogenetic utility of mitochondrialgene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am, 6(6): 651–701

Singh V K, Mangalam A K, Dwivedi S, Naik S (1998). Primer premier: Program for design of degenerate primers from a protein sequence. Biotechniques, 2: 318–319

Smart P (1989). The illustrated encyclopedia of the butterfly world (New York, USA: Chartwell Books

Thao M L, Baumann L, Baumann P(2004). Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). BMC Evol Biol, 4: 25–37

Wahlberg N, Braby M F, Brower A V Z, de Jong R, Lee M M, Nylin S, Pierce N E, Sperling F A H, Vila R, Warren A D, Zakharov E (2005). Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. Proc Biol Sci, 272(1572): 1577–1586

Weller S J, Pashely D P (1995). In search of butterfly origins. Mol Phylogenet Evol, 3(3): 235–246

Wolstenholme D R (1992). Animal mitochondrial DNA: structure and evolution. Int Rev Cytol, 141: 173–216

Xia X, Xie Z (2001). DAMBE: Data analysis in molecular biology and evolution. J Hered, 4(4): 371–373

Yang Z, Rannala B (1997). Bayesian phylogenetic inference using DNA sequences: A Markov chain Monte Carlo method. Mol Biol Evol, 7 (7): 717–724

Yin W Y, Song D X, Yang X K (2008). Study on phylogeny of the Hexapoda (Insect). Beijing, Science Press, pp:193–194

Zhou Z J, Huang Y, Shi F M (2007). The mitochondrial genome of the *Ruspolia dubia* (Orthoptera: Conocephalidae): a short A + T-rich region with 70 bp in length Genome. Genome, 50(9): 855–866