The first mitochondrial genome for the butterfly family Riodinidae (*Abisara fylloides*) and its systematic implications

Fang ZHAO[1], Dun-Yuan HUANG[1,2], Xiao-Yan SUN[3], Qing-Hui SHI[1], Jia-Sheng HAO[1,*], Lan-Lan ZHANG[1], Qun YANG[3,*]

1. *College of Life Sciences, Anhui Normal University, Wuhu* 24100, *China*
2. *College of Forestry, Jiangxi Environmental Engineering Vocational College, Ganzhou* 34100, *China*
3. *Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing* 21008, *China*

Abstract: The Riodinidae is one of the lepidopteran butterfly families. This study describes the complete mitochondrial genome of the butterfly species *Abisara fylloides*, the first mitochondrial genome of the Riodinidae family. The results show that the entire mitochondrial genome of *A. fylloides* is 15 301 bp in length, and contains 13 protein-coding genes, 2 ribosomal RNA genes, 22 transfer RNA genes and a 423 bp A+T-rich region. The gene content, orientation and order are identical to the majority of other lepidopteran insects. Phylogenetic reconstruction was conducted using the concatenated 13 protein-coding gene (PCG) sequences of 19 available butterfly species covering all the five butterfly families (Papilionidae, Nymphalidae, Peridae, Lycaenidae and Riodinidae). Both maximum likelihood and Bayesian inference analyses highly supported the monophyly of Lycaenidae+Riodinidae, which was standing as the sister of Nymphalidae. In addition, we propose that the riodinids be categorized into the family Lycaenidae as a subfamilial taxon.

Keywords: *Abisara fylloides*; Mitochondrial genome; Riodinidae; Systematic implication

The typical metazoan mitochondrial genome (mitogenome) contains 37 genes, including 13 protein-coding genes (PCGs), 2 rRNA genes and 22 tRNA genes, and a non-coding area (i.e., the control region or the A+T-rich region) (Wolstenholme, 1992; Boore, 1999). Maternal inheritance, lack of recombination and an accelerated evolutionary rate compared with the nuclear genome have all contributed to the increased use of mitogenomes, which is one of the key methods in fields such as phylogenetics, comparative and evolutionary genomics, molecular evolution and population genetics (Ballard & Whitlock, 2004; Simonsen et al, 2006). At present, mitochondrial genomes have already been determined in a variety of insect groups covering nearly 200 species. However, reported complete mitogenomes are relatively scarce for lepidopterans and especially for butterflies. To our knowledge, as of October 2012 only about 20 butterfly species covering five butterfly families (Table 2) have been reported or deposited into the GenBank, but only one butterfly family, the Riodinidae, still lack corresponding data.

The phylogenetic position and taxonomic ranking of the butterfly family Riodinidae among butterfly lineages are still controversial issues among entomologists. Some scholars suggest that the riodinids are closely related to the lycaenids, considering the similarities in morpholo gical character, behavior, and host plants between the two (slug-like larvae, pupa contigua, ants associated) (Ackery, 1984; Chou, 1998; de Jong et al, 1996; Ehrlich, 1958; Scott, 1985). Moreover, they are usually classified into the Lycaenidae family as a subfamilial taxon. Some consider the riodinids a unique family parallel to the Lycaenidae

Received: 27 November 2012; Accepted: 06 May 2013
Foundation items: This work was supported by the National Natural Science Foundation of China (41172004), Provincial Key Projects of the Natural Science Foundation in the Colleges of Anhui province (KJ2010A142), Opening Funds from the State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences (104143) and the Program for Innovative Research Teams in Anhui Normal University.
* Corresponding author, E-mail: jshaonigpas@sina.com

family (Campbell et al, 2000; de Jong et al, 1996; Kristensen, 1976; Shou et al, 2006), and others propose that the riodinids are more closely related to the nymphalids than to the other butterfly groups, such as the lycaenids (Martin & Pashley, 1992; Robbins, 1987, 1988).

This study sequenced the first complete mitochondrial genome of the *Abisara fylloides*, a representative species of the family Riodinidae, by long PCR and primer walking techniques. Its genetic structure was preliminarily compared with those of other available butterfly species. The maximum likelihood (ML) and Bayesian inference (BI) phylogenetic trees of the *A. fylloides* and other available butterfly representative species were reconstructed based on concatenated DNA sequences of the 13 protein-coding genes (PCGs), and aim to clarify their phylogenetic relationships and further provide new information about the structure, organization and molecular evolution of the lepidopteran mitogenomes.

MATERIALS AND METHODS

Sample collection and DNA extraction

Adult individuals of *A. fylloides* were collected in Jinghong, Yunnan Province, China in August 2006. After collection, sample specimens were preserved in 100% ethanol immediately and stored at −20 °C before DNA extraction (specimen No. ZHF07). Total Genomic DNA was isolated using the proteinase K-SiO_2 method as described by Hao et al (2005).

PCR amplification and sequence determining

The multiple sequence alignments were conducted using the software Clustal X 1.8 based on the mitogenome sequences of *Coreana raphaelis*, *Artogeia melete*, *Troides aeacus* available from GenBank and those of *Argyreus hyperbius*, *Acraea issoria*, *Calinaga davidis*, *Pieris rapae* determined in our laboratory (Thompson et al, 1997). The long PCR primers, which may cover the whole mitogenome, were designed according to the conserved regions by the software Primer premier 5.0 (Singh et al, 1998) (Table 1). Seven short fragment sequences (500-700 bp) of cox1, cox2, cox3, cytb, nad1, rrnL and rrnS were amplified using insect universal primers (Caterino & Sperling, 1999; Simmons & Weller, 2001; Simon et al, 1994). All the primers were synthesized by the Shanghai Sheng gong Biotechnology Co. Ltd.

Table 1 List of PCR primers used in this study

Genes	Forward primers (5′)	Reverse primers (3′)	Annealing temperature (°C)
cox1*	GGTCAACAAATCATAAAGATATTG	TAAACTTCAGGGTGACCAAAAAT	50.0
cox1-cox2	TTATTTGTATGAGCCGTAG	ATAGCAGG AAGATTGTTC	47.5
cox2*	GAGACCATTACTTGCTTTCAGTCACT	CTAATATGGCAGATTATATGTATGG	49.5
cox2-cox3	TTTTATTGCTCTTCCATCT	TTATTCCTCATCGTAATCC	48.5
cox3*	TATTTCAATGATGACGAGAT	CAAATCCAAAATGGTGAGT	49.8
cox3-nad5	TTTATAGCAACAGGATTTC	CATCAACTGGTTTAACTTT	45.5
nad5	AAAACTTCCAGAAAATAATCTC	TTGCTTTATCTACTTTAAGACA	46.5
nad5-cytb	AATTATACCAGCACATAT	TTATCGACTGCAAATC	47.1
cytb*	TATGTACTACCATGAGGACAAATAT	ATTACACCTCCTAATTTATTAGGAAT	47.0
cytb-nad1	TCCTGCTAACCCTTTAGTCA	AGGTAGATTACGGGCTGTT	48.0
nad1*	CGTAAAGTCCTAGGTTATATTCAGATCG	ATCAAAAGGAGCTCGATTAGTTTC	52.0
nad1-rrnL	AGCCCGTAATCTACCTAA	TAAGACGAGAAGACCCTAT	47.0
rrnL*	CGCCTGTTATCAAAAACAT	CCGGTCTGAACTCAGAT	45.5
rrnL-rrnS	AGACTATTGATTATGCTACCT	TAAGAATCTAATGGATTACAA	46.5
rrnS*	CTTCTACTTTGTTACGACTTA T	AATTTTGTGCCAGCAGTTG	50.0
rrnS-nad2	AGAGGGTATCTAATCCGAGTTT	TGGCTGAGAATTAAGCGATA	49.5
nad2-cox1	ATACAGAAGCAGCATTA	AGAAGGAGGAAGAAGTCAA	52.0

*: universal primer.

Seven partial gene sequences were initially sequenced under the following conditions: an initial denaturation at 94 °C for 5 minutes, then denaturation at 94 °C for 1 minute for a total of 35 cycles; annealing at 45−55 °C for 1 minute and extension at 72 °C for 2 minutes plus 30 seconds; final extension at 72 °C for 10 minutes. Long PCRs were performed using TaKaRa LA Taq polymerase with the following cycling parameters: an initial denaturation for 5 minutes at 95 °C; followed by 30 cycles at 95 °C for 55 seconds, 45−55 °C for 2 minutes, 68 °C for 2 min and 30 seconds; and a subsequent final extension step of 68 °C for 10 minutes.

The PCR products were separated by electrophoresis in a 1.2% agarose gel and purified using the DNA gel

extraction kit (TaKaRa). All PCR fragments were sequenced directly after purification with the QIA quick PCR Purification Kit reagents (QIAGEN). Internal primers were applied to complete sequences by primer walking (detailed primer information will be provided upon request). All fragments were sequenced for both strands.

Data analysis

We used DNASIS MAX (Hitachi) for sequence assembly and annotation. Protein-coding genes and rRNA genes were identified by sequence comparison with other available insect mitochondrial sequences. The tRNAs were identified by tRNAscan-SE v.1.21 (Lowe & Eddy, 1997). The putative tRNAs, which were not found by tRNAscan-SE, were identified by a sequence comparison of *A. fylloides* with the other lepidopteran tRNAs. PCGs were aligned with the other available lepidopteran mitogenomes using DAMBE software (Xia & Xie, 2001). The tandem repeats in the A+T-rich region were predicted using the Tandem Repeats Finder online (http://tandem.bu.edu/trf/trf.html) (Benson, 1999). Nucleotide composition was calculated using PAUP 4.0b10 (Swofford, 2002). The mitogenome sequence data have been deposited in GenBank under the accession number HQ259069.

Phylogenetic analysis

Phylogenetic analyses were performed on 19 representative species including *A. fylloides*, covering all the six families of butterflies. The multiple aligning of the concatenated nucleotide sequences of the 13 mitochondrial PCGs of the 19 species (Table 2) was conducted using ClustalX 1.8. The phylogenetic trees were reconstructed with the maximum likelihood (ML) and Bayesian inference (BI) methods, using the moth species Adoxophyes honmai (GenBank accession number of mitogenome: NC014295) as the outgroup. In both phylogenetic analyses, the third codon position of all the sequences was excluded. The ML analyses were conducted in PAUP 4.0b10 by using TBR branch swapping (10 random addition sequences) as a search method. The model GTR+I+Γ was selected as the best fit model using Modeltest 3.06 (Posada & Crandall, 1998) under the AIC scores, and the bootstrap values of the ML tree were evaluated via the bootstrap test with 1 000 iterations. The Bayesian analysis was performed using MrBayes 3.1.2 (Huelsenbeck & Ronquist, 2001) with the partitioned strategy (13 partitions: cox1, cox2, cox3, atp8, atp6, nad1, nad2, nad4, nad4L, nad5, nad6 and cytb), and the best substitution model for each partition 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 with a burn-in of 25% until the average standard deviation of split frequencies to be less than 0.01, which means that convergence was reached.

RESULTS AND DISCUSSION

Genome structure and organization

The complete mitogenome is 15 301 bp in length, encodes 37 genes in all. It contains 13 protein, 22 tRNA, 2 rRNA genes and a non-coding high A+T content region (Figure 1, Table 3). Its structure and organization are identical to those of the majority of other lepidopterans (Bae et al, 2004; Cha et al, 2007; Cameron & Whiting, 2008; Hao et al, 2012; Hong et al, 2008, 2009; Hu et al, 2010; Kim et al, 2009b; Ji et al, 2012; Junqueira et al, 2004; Wang et al, 2011), though a few lepidopterans, such as three *Thitarodes* species, were reported to possess the ancestral gene order trnI-trnQ-trnM instead of the trnM-trnI-trnQ (Cao et al, 2012) .

Eight overlapping sequences totaling 61 bp are located throughout the *A. fylloides* mitogenome, with size ranging from 2 to 35 bp, of which the longest (35 bp) is located between the cox2 and the $tRNA^{Lys}$ genes. In addition, 17 intergenic spacers ranging from 1 to 45 bp in length are found in the mitogenome. Among these spacers, the longest is located between the $tRNA^{Gln}$ and nad2 genes, the other 16 spacers are scattered throughout the whole genome (Figure 1, Table 3). Most of these spacer regions are arranged relatively compactly

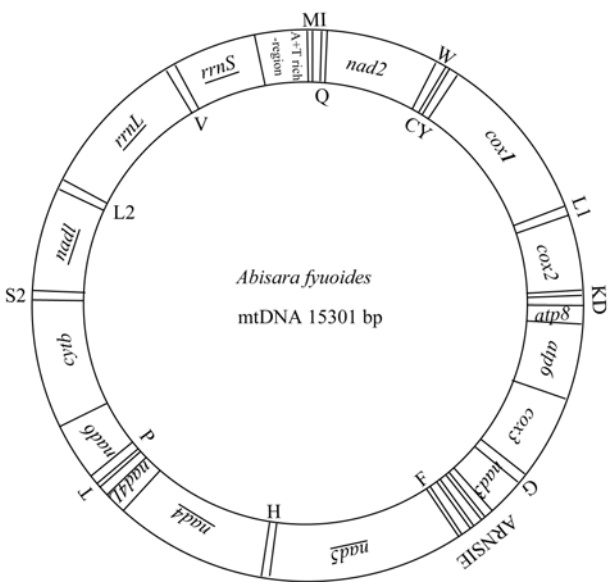

Figure 1 Circular map of the mitochondrial genome of *Abisara fylloides*

Gene names not underlined indicate the direction of transcription from left to right and those underlined indicate right to left. Transfer RNA genes encoded by H and L strands are shown outside and inside the circular gene map, respectively. Transfer RNA genes are indicated by the IUPAC-IUB single letter amino acid codes, while L1, L2, S1, S2 represent tRNA-Leu(UUR), tRNA-Leu(CUN), tRNA-Ser(AGN) and tRNA-Ser(UCN), respectively.

Table 2 The available rhopaloceran mitogenomes and their characteristics

Family	Subfamily	Species	Whole genome			PCG[b]	*rrnL*		*rrnS*		AT-rich region		GenBank accession no.	References
			Size (bp)	(A+T) %	No. codons[a]	(A+T) %	Size (bp)	(A+T) %	Size (bp)	(A+T) %	Size (bp)	(A+T) %		
Papilionidae	Papilioninae	*Teinopalpus aureus*	15 242	79.8	3 720	78.3	1 320	82.4	781	85.6	395	93.1	HM563681	Qin et al, 2012
		Papilio machaon	15 185	80.3	3 707	79.0	1 319	83.5	773	84.2	362	92.5	HM243594	unpublished
	Parnassiinae	*Parnassius bremeri*	15 389	81.3	3 723	80.2	1 344	83.9	773	85.1	504	93.6	FJ871125	Kim et al, 2009b
Hesperiidae	Pyrginae	*Ctenoptilum vasava*	15 468	80.5	3 698	78.9	1 343	84.1	774	86.4	429	88.1	JF713818	Hao et al, 2012
	Hesperiinae	*Parnara guttata*	15 411	80.6	3 718	78.9	1 368	84.5	778	85.1	411	90.8	JX101619	unpublished
	Coeliadinae	*Choaspes benjaminii*	15 272	80.8	3 681	79.0	1 355	85.2	777	86.0	293	92.2	JX101620	unpublished
Pieridae	Pierinae	*Artogeia melete*	15 140	79.8	3 715	78.4	1 319	83.4	777	85.5	351	89.2	NC_010568	Hong et al, 2009
		Pieris rapae	15 157	79.7	3 722	78.3	1 320	84.0	764	85.0	393	91.6	HM156697	Mao et al, 2010
Lycaenidae	Theclinae	*Coreana raphaelis*	15 314	82.7	3 708	81.5	1 330	85.3	777	85.8	375	94.1	DQ102703	Kim et al, 2006
		Protantigius superans	15 248	81.7	3 712	80.3	1 331	85.1	739	85.6	361	93.6	HQ184265	Kim et al, 2011
	Aphnaeinae	*Spindasis takanonis*	15 349	82.4	3 719	81.0	1 333	85.6	777	84.7	371	94.6	HQ184266	Kim et al, 2011
Nymphalidae	Apaturinae	*Apatura ilia*	15 242	80.5	3 711	78.9	1 333	84.0	776	84.9	403	92.5	JF439725	Chen et al, 2012
		Sasakia charonda	15 244	79.9	3 695	78.2	1 323	84.4	775	85.0	380	91.8	NC_014224	unpublished
	Nymphalinae	*Kallima inachus*	15 183	80.3	3 721	79.2	1 335	82.7	774	85.1	376	92.0	JN857943	Qin et al, 2012
	Heliconiinae	*Argyreus hyperbius*	15 156	80.8	3 718	79.5	1 330	84.5	778	85.2	349	95.4	JF439070	Wang et al, 2011
		Acraea issoria	15 245	79.7	3 717	78.0	1 331	83.9	788	83.7	430	96.0	GQ376195	Hu et al, 2010
	Limenitidinae	*Athyma sulpitia*	15 268	81.9	3 729	80.6	1 319	84.7	779	85.7	349	94.6	JQ347260	Tian et al, 2012
	Libytheinae	*Libythea celtis*	15 164	81.2	3 722	80.0	1 335	84.7	775	85.4	328	96.3	HQ378508	unpublished
Riodinidae	Nemeobiinae	*Abisara fylloides*	15 301	81.2	3 730	79.8	1 334	85.4	771	85.6	423	91.0	HQ259069	This study

a: Termination codons were excluded in total codon count. b: Protein-coding genes.

Table 3 Organization of the *Abisara fylloides* mitochondrial genome

Gene	Direction	Position	Size (bp)	Intergenic length*	Start codon	Stop codon
$tRNA^{Met}$	F	1−67	67	0		
$tRNA^{Ile}$	F	68−131	64	−3		
$tRNA^{Gln}$	R	129−197	69	45		
nad2	F	243−1 256	1 014	1	ATT	TAA
$tRNA^{Trp}$	F	1 258−1 324	67	−8		
$tRNA^{Cys}$	R	1 317−1 382	66	3		
$tRNA^{Tyr}$	R	1 386−1 451	66	6		
cox1	F	1 458−2 989	1 532	3	CGA	T
$tRNA^{Leu(UUR)}$	F	2 993−3 059	67	0		
cox2	F	3 060−3 738	679	-3	ATG	T
$tRNA^{Lys}$	F	3 736−3 806	71	0		
$tRNA^{Asp}$	F	3 807−3 871	65	0		
atp8	F	3 872−4 033	162	−7	ATC	TAA
atp6	F	4 027−4 714	688	−2	ATG	TAA
cox3	F	4 713−5 501	789	2	ATG	TAA
$tRNA^{Gly}$	F	5 504−5 569	66	0		
nad3	F	5 570−5 923	354	3	ATT	TAA
$tRNA^{Ala}$	F	5 927−5 997	71	3		
$tRNA^{Arg}$	F	6 001−6 064	64	1		
$tRNA^{Asn}$	F	6 066−6 131	66	16		
$tRNA^{Ser(AGN)}$	F	6 148−6 208	61	1		
$tRNA^{Glu}$	F	6 210−6 275	66	-2		
$tRNA^{Phe}$	R	6 274−6 340	67	0		
nad5	R	6 341−8 108	1 768	0	ATT	T
$tRNA^{His}$	R	8 109−8 179	71	6		
nad4	R	8 186−9 518	1 333	-2	ATG	T
nad4L	R	9 517−9 805	289	2	ATG	TAA
$tRNA^{Thr}$	F	9 808−9 870	63	0		
$tRNA^{Pro}$	R	9 871−9 937	67	5		
nad6	F	9 943−10 467	525	3	ATA	TAA
cytb	F	10 471−11 622	1 152	−2	ATG	TAA
$tRNA^{Ser(UCN)}$	F	11 621−11 685	65	17		
nad1	R	11 703−12 641	939	1	ATG	TAA
$tRNA^{Leu(CUN)}$	R	12 643−12 710	68	0		
rrnL	R	12 711−14 044	1 334	0		
$tRNA^{Val}$	R	14 045−14 107	63	0		
rrnS	R	14 108−14 878	771	0		
A+T-rich region		14 879−15 301	423			

*: In the column intergenic length, the positive number indicates interval base pairs between genes, while the negative number indicates the overlapping base pairs between genes.

compared with other insect mitogenomes (Cameron & Whiting, 2008; Hao et al, 2012; Hong et al, 2009; Hu et al, 2010; Kim et al, 2009b; Ji et al, 2012; Junqueira et al, 2004; Wang et al, 2011).

The *A. fylloides* A+T-rich region is flanked on one side by the rrnS and on the other side by the $tRNA^{Met}$ genes. This region exhibits the highest A+T content (91.0%) (Table 4), and spans 423 bp (Table 3, Table 4). A sequence analysis of the A+T-rich region revealed that it contained some structures typical of other lepidopteran mitogenomes: (1) at 20 bp downstream of the small subunit rRNA gene, there is a structure including a motif 'ATAGA' which is very well conserved in all sequenced lepidopteran insects, and a 18-bp polyT stretch, both of which have been suggested as the origin of minority or light strand replication (O_N) and to play a regulatory role (Kim et al, 2009a; Lutz-Bonengel et al, 2004; Saitou et al, 2005; Yukuhiro et al, 2002). (2) Between the sites 15 151 and 15 168 there is a microsatellite-like $(TA)_9$ element, which is also found in the majority of other lepidopterans; (3) There is another motif "ATTTA" of unknown function located from 15 139 to 15 143 upstream of the $(TA)_9$, which is also typical of the other lepidopterans. In addition, to our great surprise, an unexpected short microsatellite-like repeating region $(TA)_{11}$ downstream of the $(TA)_9$ was detected in the A+T-rich region, and this has not been reported in any other lepidopterans.

Table 4 Nucleotide composition and skewness in different regions of the *Abisara fylloides* mitogenome

Region	Size (bp)	Nucleotide composition (%)					AT-skew	GC-skew
		T	C	A	G	A+T		
Whole genome	15 301	41.7	11.3	39.5	7.5	81.2	−0.027	−0.202
Major-strand PCGs	6 927	44.5	12.1	34.1	9.3	78.6	−0.132	−0.131
Minor-strand PCGs	4 329	47.6	6.3	34.1	12.0	81.7	−0.165	0.311
Whole PCGs	11 224	45.8	9.9	34.0	10.3	79.8	−0.148	0.020
Whole tRNA	1 461	40.0	7.7	41.5	10.7	81.5	0.018	0.163
rrnL	1 334	38.4	4.8	47.0	9.8	85.4	0.101	0.342
rrnS	771	39.9	4.9	45.7	9.5	85.6	0.068	0.319
A+T-rich region	423	48.9	4.5	42.1	4.5	91.0	−0.075	0.000

Base composition bias

The base composition of the *A. fylloides* mitogenome shows an A and T bias (Table 4). The whole A+T content of the mitogenome is up to 81.2%, ranging from that of *Argyreus hyperbius* (80.81%) to that of *Coreana raphaelis* (82.66%). Like most other metazoan mitogenomes, the A+T content of the A+T-rich region, which is located between the rrnS and $tRNA^{Met}$ genes, is the highest (91.0%) in all known butterfly species except for *P. rapae* (Pieridae) to date. The base contents of A, T, C, G are 39.5%, 41.7%, 11.3%, 7.5%, respectively, indicating a relatively higher A+T content (81.2%). These phenomena commonly exist in the protein-coding genes, which have a relatively lower A+T content (79.8%), and the tRNA and rRNA genes in insects.

Protein-coding genes

The sequences of the 13 *Abisara fylloides* PCGs are 11 224 bp in length, including 3 730 codons (excluding termination codons). Twelve of the 13 PCGs use standard ATN as their start codon except for the cox1 gene, and eight of these 12 PCGs begin with ATA or ATG (Methionine) and the other four begin with ATT or ATC (Isoleucine) (Table 3). The start codons for cox1 gene of lepidopteran insects have usually been a controversial issue. In general, there is no typical start codon, especially for the cox1 gene, which usually uses CAG (R) as start codon in insects, including the lepidopterans. However, some scholars reported unusual start codons, such as the trinucleotide TTG (Bae et al, 2004; Hong et al, 2008), ACG (Lutz-Bonengel et al, 2004), GCG (Nardi et al, 2003), the tetranucleotide ATAA, ATCA and ATTA (Clary & Wolstenholme, 1983; de Bruijn,1983; Kim et al, 2006), and the hexanucleotides TATTAG (Flook et al, 1995), TTTTAG (Yukuhiro et al, 2002), TATCTA (Coates et al, 2005), ATTTAA (Beard et al, 1993; Mitchell et al, 1993) for cox1 in some other insect species. However, the CGA is present as a conserved region for all lepidopteran insects reported, and thus we tend to consider that CGA is the cox1 start codon for *A. fylloides* as Kim et al (2009b) suggested. As for stop codons, 9 of the 13 PCGs use standard TAA except for the cox1, cox2, nad4 and nad5 genes, all of which terminate at a single thymine (Table 3), and this case was also found in all other lepidopterans reported to date. For more details on this phenomenon, please refer to the discussions of Kim et al. (2010).

The PCG amino acid sequence variation analysis showed that there were 3 755 homologous sites, of which 1 964 are conserved, 1 791 are variable, and 1 128 are parsimony informative. Among the twenty amino acids in the 13 PCGs, six (Leu, Met, Ile, Phe, Asn, and Tyr) were used more frequently than the others, and their usage frequencies were higher than the average, whereas the other 14 amino acids less used (Table 5).

rRNA and tRNA genes

The large subunit rRNA and small subunit rRNA genes of the *A. fylloides* are 1 334 and 771 bp in length, respectively. As in other lepidopterans, these two genes are located between $tRNA^{Leu(UUR)}$ and $tRNA^{Val}$, and between $tRNA^{Val}$ and A+T-rich region respectively (Cameron & Whiting, 2008; Hao et al, 2012; Hu et al, 2010; Kim et al, 2009a; Salvato et al, 2008; Wang et al, 2011; Yang et al, 2009).

The *A. fylloides* mitogenome harbors 22 tRNA genes, which are scattered throughout the whole genome and ranged in length from 61 to 71 bp. Except for the $tRNA^{Ser(AGN)}$, which lacks the DHU loop, all tRNAs are shown to be folded into the cloverleaf secondary structures, within which all amino acid acceptor stems have 7 base pairs, and all anticodon stems have 5 base pairs. This was also found in all the other lepidopterans determined to date (Cameron & Whiting, 2008; Hao et al, 2012; Hu et al, 2010; Kim et al, 2009a).

A total of 27 pairs of base mismatches were detected in all the predicted tRNA secondary structures, among which 17 are GU, 6 are UU, 2 are AA, and 2 are AC. The AA mismatches occur at the anticodon stem of

Table 5 Codon usage of the protein-coding genes of the *Abisara fylloides* mitogenomes

Codon	N (RSCU)	Codon	N (RSCU)	Codon	N (RSCU)	Codon	N (RSCU)
UUU (F)	356 (1.86)	UCU (S)	108 (2.59)	UAU (Y)	175 (1.83)	UGU (C)	28 (1.81)
UUC (F)	26 (0.14)	UCC (S)	16 (0.38)	UAC (Y)	16 (0.17)	UGC (C)	3 (0.19)
UUA (L)	488 (5.30)	UCA (S)	94 (2.25)	UAA (*)	0 (0)	UGA (W)	84 (1.87)
UUG (L)	12 (0.13)	UCG (S)	2 (0.05)	UAG (*)	0 (0)	UGG (W)	6 (0.13)
CUU (L)	34 (0.37)	CCU (P)	70 (2.33)	CAU (H)	62 (1.85)	CGU (R)	14 (1.04)
CUC (L)	3 (0.03)	CCC (P)	9 (0.3)	CAC (H)	5 (0.15)	CGC (R)	0 (0)
CUA (L)	14 (0.15)	CCA (P)	38 (1.27)	CAA (Q)	57 (1.93)	CGA (R)	37 (2.74)
CUG (L)	1 (0.01)	CCG (P)	3 (0.1)	CAG (Q)	2 (0.07)	CGG (R)	3 (0.22)
AUU (I)	424 (1.88)	ACU (T)	74 (2.1)	AAU (N)	251 (1.81)	AGU (S)	25 (0.6)
AUC (I)	27 (0.12)	ACC (T)	7 (0.2)	AAC (N)	26 (0.19)	AGC (S)	0 (0)
AUA (M)	283 (1.81)	ACA (T)	59 (1.67)	AAA (K)	97 (1.92)	AGA (S)	89 (2.13)
AUG (M)	29 (0.19)	ACG (T)	1 (0.03)	AAG (K)	4 (0.08)	AGG (S)	0 (0)
GUU (V)	74 (2.39)	GCU (A)	73 (2.52)	GAU (D)	55 (1.83)	GGU (G)	56 (1.17)
GUC (V)	4 (0.13)	GCC (A)	8 (0.28)	GAC (D)	5 (0.17)	GGC (G)	1 (0.02)
GUA (V)	45 (1.45)	GCA (A)	33 (1.14)	GAA (E)	65 (1.78)	GGA (G)	100 (2.08)
GUG (V)	1 (0.03)	GCG (A)	2 (0.07)	GAG (E)	8 (0.22)	GGG (G)	35 (0.73)

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

Stop codons were excluded in total codon counts.

$tRNA^{Lys}$ and the TψC loop of $tRNA^{Trp}$; the AC mismatches occur at the amino acid acceptor stem of $tRNA^{Tyr}$ and the anti codon stem of $tRNA^{Leu(UUR)}$. Among the 6 UU mismatches, two occur at the amino acid acceptor stems of $tRNA^{Leu(UUR)}$ and $tRNA^{Ala}$, two at the anti codon stems of $tRNA^{Leu(UUR)}$ and $tRNA^{Gln}$, and another two at the anti codon stem of $tRNA^{Ser(UCN)}$, respectively (Figure 2).

Phylogenetic analysis

At present, for the phylogenetic positions of the riodinids within papilionid butterflies, most morphological studies place them as most closely related to the lycaenids and identify the nymphalids as the closest relatives to this riodinid+lycaenid clade (de Jong et al, 1996; Ehrlich & Ehrlich, 1967; Kristensen, 1976; Scott & Wright, 1990). These relationships have been inferred using a variety of phylogenetic methods and are supported by a number of adult, larval and pupal synapomorphies. Additionally, molecular (DNA sequence of the mitochondrial NADH1 gene) or molecular plus morphological evidence also result in a monophyletic interpretation of the Riodinidae+ Lycaenidae, and their sister relationship to the Nymphalidae (Wahlberg et al, 2005; Weller et al, 1996). However, based on a cladistic analysis of four foreleg characters with nine character states, Robbins (1988) suggested that the Riodinidae are more closely related to the Nymphalidae than to the Lycaenidae, and this result is supported by the nuclear 28S rRNA gene sequence data (Martin & Pashley, 1992).

There are two opinions regarding the taxonomic rank of riodinids. First, some previous studies suggest that the riodinids should be categorized into the family lycaenids as a subfamilial taxon in light of their morphological characters (Chou, 1998; de Jong et al, 1996; Ehrlich, 1958; Kristensen, 1976; Scott & Wright, 1990), and this opinion is supported by the molecular studies of Zou et al (2009) and Hao et al (2007). Other studies postulate that the riodinids should be classified as a separate family parallel to Lycaenidae (Harvey, 1987; Martin & Pashley, 1992; Robbins, 1988; Weller et al, 1996), and this view is supported by molecular phylogenetic studies based on data from the *wingless* gene by Campbell et al (2000) and the combined analysis of sequences of the nuclear Ef-1a, wingless, and mitochondrial COI genes by Wahlberg et al (2005).

The ML and Bayesian trees of this study (Figure 3) showed that all the butterfly taxa in this study did not form a monophyletic unit, and a similar case was reported by Hao et al (2012). Nonetheless, both the ML and BI trees indicated that all the butterfly species were grouped into five distinct lineages: 1) the Papilionidae, including papilionids and parnassids; 2) Hesperiidae; 3) Pieridae; 4) Nymphalidae; 5) Lycaenidae+Riodinidae. The monophyly of Lycanidae + Riodinidae was strongly supported with a 100% bootstrap value in ML, and with 1.00 posterior probability value in BI. Thus, considering the results of Heikkilä et al (2012), which indicate the monophylies of lycaenids and riodinids, it is reasonable to propose that the two groups may be sisters, though the taxa sampling of riodinids in this analysis is extremely limited. Additionally, based their congruent genetic divergences compared with those between other butterfly subfamilies, the riodinids should be categorized into the Lycaenidae family as a subfamilial taxon.

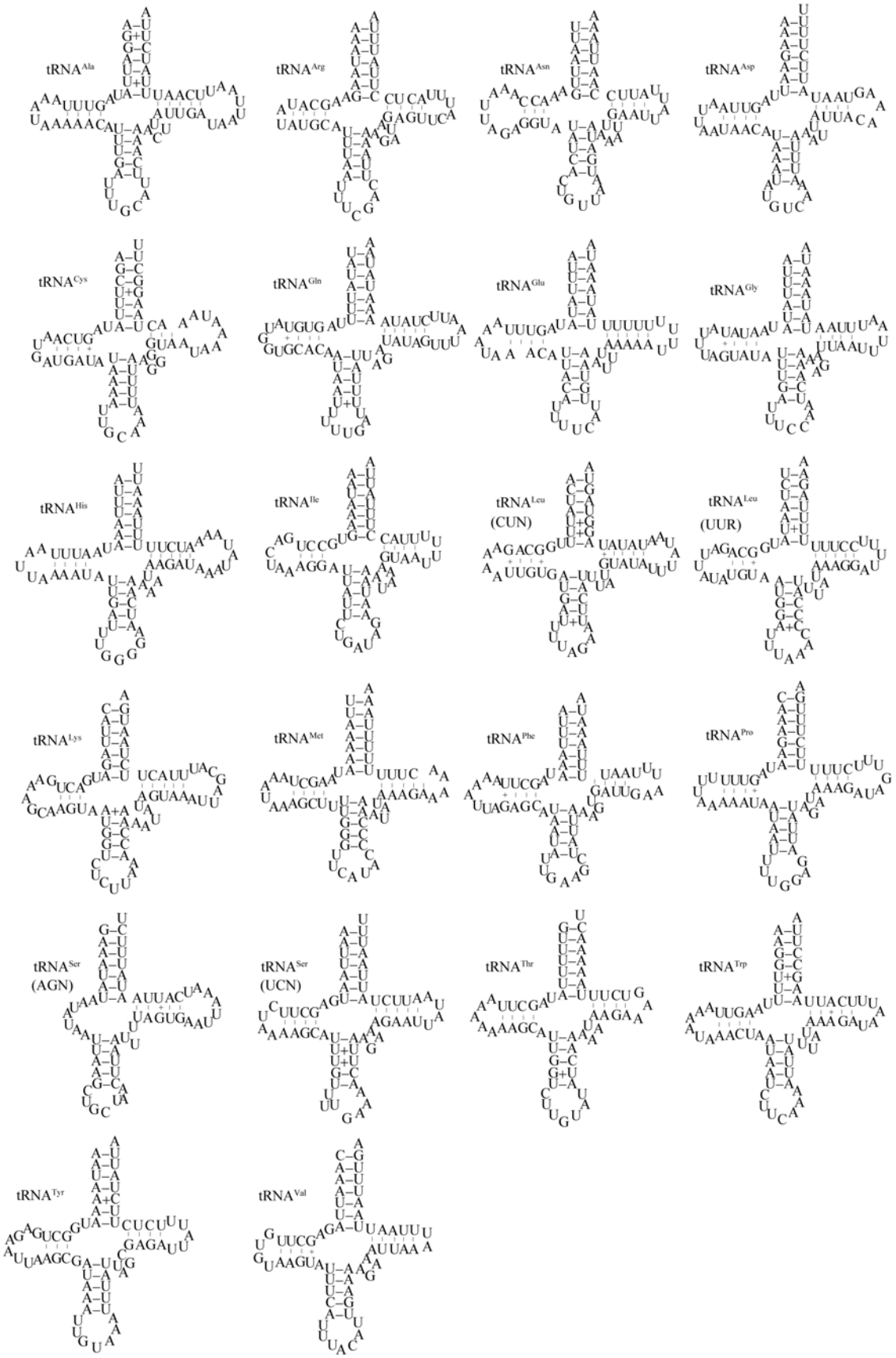

Figure 2 Predicated clover-leaf secondary structures for the mitochondrial tRNA genes of *Abisara fylloides*

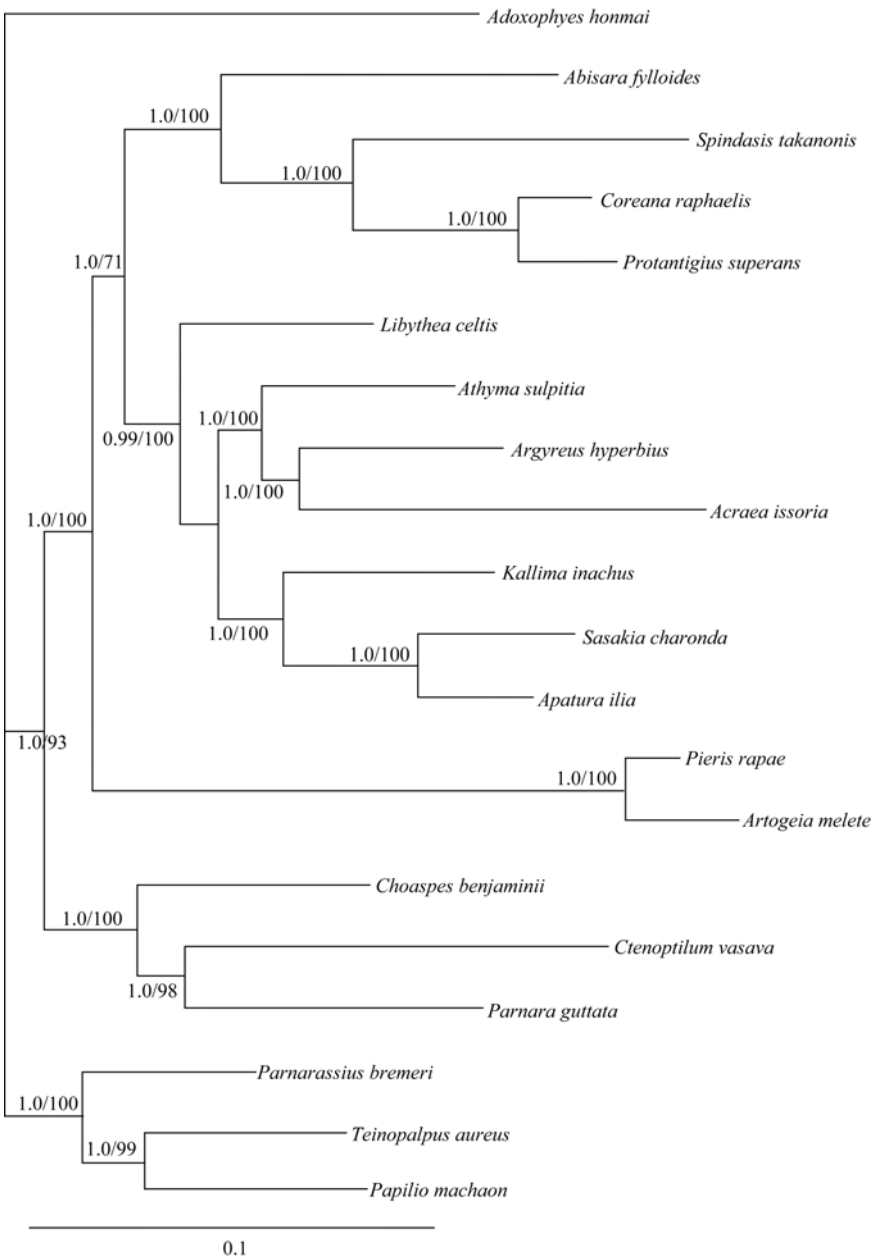

Figure 3 The Bayesian inference (BI) and maximum likelihood (ML) phylogenetic trees of main butterfly lineages based on 13 protein-coding gene sequences (Numbers on each node correspond to the posterior probability values of the BI analysis and the ML bootstrap percentage values for 1 000 replicates of ML analysis)

References

Ackery PR. 1984. Systematic and faunistic studies on butterflies. *In*: Vane-Wright RI, Ackery PR. The Biology of Butterflies. London: Academic Press.

Bae JS, Kim I, Sohn HD, Jin BR. 2004. The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization, and phylogenetic analysis with other insects. *Molecular Phylogenetics and Evolution,* **32**(3): 978-985.

Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. *Molecular Ecology,* **13**(4): 729-744.

Beard CD, Hamm SM, Collins FH. 1993. The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. *Insect Molecular Biology,* **2**(2): 103-124.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research,* **27**(2): 573-580.

Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Research,* **27**(8): 1767-1780.

Cao YQ, Ma C, Chen JY, Yang DR. 2012. The complete mitochondrial genomes of two ghost moths, *Thitarodes renzhiensis* and *Thitarodes*

yunnanensis: the ancestral gene arrangement in Lepidoptera. *BMC Genomics*, **13**(1): 276.

Cameron SL, Whiting MF. 2008. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene*, **408**(1-2): 112-123.

Campbell DL, Brower AVZ, Pierce NE. 2000. Molecular evolution of the wingless gene and its implications for the phylogenetic placement of the butterfly family Riodinidae (Lepidoptera: Papilionoidea). *Molecular Biology and Evolution,* **17**(5): 684-696.

Caterino MS, Sperling FAH. 1999. *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. *Molecular Phylogenetics and Evolution,* **11**(1): 122-137.

Cha SY, Yoon HJ, Lee EM, Yoon MH, Hwang JS, Jin BR, Han YS, Kim I. 2007. The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae). *Gene,* **392**(1-2): 206-220.

Chen M, Tian LL, Shi QH, Cao TW, Hao JS. 2012. Complete mitogenome of the Lesser Purple Emperor *Apatura ilia* (Lepidoptera: Nymphalidae: Apaturinae) and comparison with other nymphalid butterflies. *Zoological Research,* **33**(2): 191-201.

Chou I. 1998. Classification and Identification of Chinese Butterflies[M]. Zhengzhou: Henan Scientific and Technological Publishing House. (in Chinese)

Coates BS, Sumerford DV, Hellmich RL, Lewis LC. 2005. Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis*. *International Journal of Biological Sciences,* **1**: 13-18.

Clary DO, Wolstenholme DR. 1983. Genes for cytochrome *c* oxidase subunit I, URF2, and three tRNAs in Drosophila mitochondrial DNA. *Nucleic Acids Research,* **11**(19): 6859-6872.

De Bruijn MHL. 1983. *Drosophila melanogaster* mitochondrial DNA, a novel organization and genetic code. *Nature*, **304**(5932): 234-241.

De Jong R, Vane-Wright RI, Ackery PR. 1996. The higher classification of butterflies (Lepidoptera): problems and prospects. *Insect Systematics & Evolution*, **27**(1): 65-103.

Ehrlich PR. 1958. The Comparative Morphology, Phylogeny and Higher Classification of the Butterflies (Lepidoptera: Papilionoidea). *University of Kansas Science Bulletin*, **39**(8**)**: 305-370.

Ehrlich PR, Ehrlich AH. 1967. The phenetic relationships of the butterflies. I. Adult taxonomy and the nonspecificity hypothesis. *Systematic Biology,* **16**(4): 301-317.

Flook PK, Rowell CH, Gellissen G. 1995. The sequence, organization, and evolution of the *Locusta migratoria* mitochondrial genome. *Journal of Molecular Evolution,* **41**(6): 928-941.

Hao JS, Li CX, Sun XY, Yang Q. 2005. Phylogeny and divergence time estimation of cheilostome bryozoans based on mitochondrial 16S rRNA sequences. *Chinese Science Bulletin,* **50**(12): 1205-1211.

Hao JS, Su CY, Zhu GP, Chen N, Wu DX, Pan HC, Zhang XP. 2007. Mitochondrial 16S rDNA molecular morphology of the main butterflies' lineages and its phylogenetic significance. *Journal of Genetics and Molecular Biology,* **18**(2): 109-121.

Hao JS, Sun QQ, Zhao HB, Sun XY, Gai YH, Yang Q. 2012. The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperiidae: Pyrginae) and its phylogenetic implication. *Comparative and Functional Genomics,* **2012**: 328049, doi: 10.1155/2012/328049.

Harvey DJ. 1987. The Higher Classification of the Riodinidae. Ph. D dissertation, University of Texas, Austin.

Heikkilä M, Kaila L, Mutanen M, Peña C, Wahlberg N. 2012. Cretaceous origin and repeated tertiary diversification of the redefined butterflies. *Proceedings of the Royal Society B: Biological Sciences,* **279**(1731): 1093-1099.

Hong GY, Jiang ST, Yu M, Yang Y, Li F, Xue FS, Wei ZJ. 2009. The complete nucleotide sequence of the mitochondrial genome of the cabbage butterfly, *Artogeia melete* (Lepidoptera: Pieridae). *Acta Biochimica et Biophysica Sinnica,* **41**(6): 446-455.

Hong MY, Lee EM, Jo YH, Park HC, Kim SR, Hwang JS, Jin BR, Kang PD, Kim KG, Han YS, 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.

Hu J, Zhang DX, Hao JS, Huang DY, Cameron S, Zhu CD. 2010. The complete mitochondrial genome of the yellow coaster, *Acraea issoria* (Lepidoptera: Nymphalidae: Heliconiinae: Acraeini): sequence, gene organization and a unique tRNA translocation event. *Molecular Biology Reports,* **37**(7): 3431-3438.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics,* **17**(8): 754-755.

Ji LW, Hao JS, Wang Y, Huang DY, Zhao JL, Zhu CD. 2012. The complete mitochondrial genome of the dragon swallowtail, *Sericinus montela* Gray (Lepidoptera: Papilionidae) and its phylogenetic implication. *Acta Entomologica Sinica,* **55**(1): 91-100.

Junqueira ACM, Lessinger AC, Torres TT, da Silva FR, Vettore AL, Arruda P, Espin AMLA. 2004. The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae). *Gene*, **339**: 7-15.

Kim I, Lee EM, Seol KY, Yun EY, Lee YB, Hwang JS, Jin BR. 2006. The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). *Insect Molecular Biology,* **15**(2): 217-225.

Kim SR, Kim MI, Hong MY, Kim KY, Kang PD, Hwang JS, Han YS, Jin BR, Kim I. 2009a. The complete mitogenome sequence of the Japanese oak silkmoth, *Antheraea yamamai* (Lepidoptera: Saturniidae). *Molecular Biology Reports,* **36**(7): 1871-1880.

Kim MI, Baek JY, Kim MJ, Jeong HC, Kim KG, Bae CH, Han YS, Jin BR, Kim I. 2009b. Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. *Molecules and Cells,* **28**(4): 347-363.

Kim MJ, Wan XL, Kim KJ, Hwang JS, Kim I. 2010. Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). *African Journal of Biotechnology,* **9**(5): 735-754.

Kim MJ, Kang RA, Jeong HC, Kim KG, 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). *Molecular Phylogenetics and Evolution,* **61**(2): 436-445.

Kristensen NP. 1976. Remarks on the family-level phylogeny of butterflies (Insecta, Lepidoptera, Rhopalocera). *Journal of Zoological Systematics and Evolutionary Research,* **14**(1): 25-33.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research,* **25**(5): 955-964.

Lutz-Bonengel S, Sänger T, Pollak S, Szibor R. 2004. Different methods to determine length heteroplasmy within the mitochondrial control region. *International Journal of Legal Medicine,* **118**(5): 274-281.

Martin JA, Pashley DP. 1992. Molecular systematic analysis of butterfly family and some subfamily relationships (Lepidoptera: Papilionoidea). *Annals of the Entomological Society of America,* **85**(2): 127-139.

Mitchell SE, Cockburn AF, Seawright JA. 1993. The mitochondrial genome of *Anopheles quadrimaculatus* species A: complete nucleotide sequence and gene organization. *Genome,* **36**(6): 1058-1073.

Mao ZH, Hao JS, Zhu GP, Hu J, Si MM, Zhu CD. 2010. Sequencing and analysis of the complete mitochondrial genome of *Pieris rapae* Linnaeus (Lepidoptera: Pieridae). *Acta Entomologica Sinica,* **53**(11): 1295-1304. (in Chinese)

Nardi F, Carapelli A, Dallai R, Frati F. 2003. The mitochondrial genome of the olive fly *Bactrocera oleae*; two haplotypes from distant geographical locations. *Insect Molecular Biology,* **12**(6): 605-611.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics,* **14**(9): 817-818.

Qin F, Jiang GF, Zhou SY. 2012a. Complete mitochondrial genome of the *Teinopalpus aureus guangxiensis* (Lepidoptera: Papilionidae) and related phylogenetic analyses. *Mitochondrial DNA,* **23**(2): 123-125.

Qin XM, Guan QX, Zeng DL, Qin F, Li HM. 2012b. Complete mitochondrial genome of *Kallima inachus* (Lepidoptera: Nymphalidae: Nymphalinae): Comparison of *K. inachus* and *Argynnis hyperbius*. *Mitochondrial DNA,* **23**(4): 318-320.

Robbins RK. 1987. Logic and phylogeny: a critique of Scott's phylogenies to the butterflies and Macrolepidoptera. *Journal of the Lepidopterists' Society,* **41**: 214-216.

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

Saitou S, Tamura K, Aotsuka T. 2005. Replication origin of mitochondrial DNA in insects. *Genetics,* **171**(4): 1695-1705.

Salvato P, Simonato M, Battisti A, Negrisolo E. 2008. The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). *BMC Genomics*, **9**: 331.

Scott JA. 1985. The phylogeny of butterflies (Papilionoidea and Hesperioidea). *Journal of Research on the Lepidoptera,* **23**(4): 241-281.

Scott JA, Wright DM. 1990. Butterfly phylogeny and fossils // Kudrna O. Butterflies of Europe. Vol 2. Wiesbaden: Aula-Verlag.

Shou JX, Chou I, Li YF. 2006. Systematic Butterfly Names of the World. Xi'an: Shaanxi Science and Technology Press. (in Chinese)

Simmons RB, Weller SJ. 2001. Utility and evolution of cytochrome *b* in insects. *Molecular Phylogenetics and Evolution,* **20**(2): 196-210.

Simon C, Frati F, Bekenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America,* **87**(6): 651-701.

Simonsen TJ, Wahlberg N, Brower AVZ, de Jong R. 2006. Morphology, molecules and fritillaries: approaching a stable phylogeny for Argynnini (Lepidoptera: Nymphalidae). *Insect Systematics and Evolution,* **37**(4): 405-418.

Singh VK, Mangalam AK, Dwivedi S, Naik S. 1998. Primer premier: program for design of degenerate primers from a protein sequence. *Biotechniques*, **24**(2): 318-319.

Swofford DL. 2002. PAUP* 4.0b10: Phylogenetic Analysis Using Parsimony (* and Other Methods), Beta Version. Sunderland: Sinauer Associates.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DJ. 1997. The CLUSTAL.X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research,* **25**(24): 4876-4882.

Tian LL, Sun XY, Chen M, Gai YH, Hao JS, Yang Q. 2012. Complete mitochondrial genome of the Five-dot Sergeant *Parathyma sulpitia* (Nymphalidae: Limenitidinae) and its phylogenetic implications. *Zoological Research,* **33**(2): 133-143.

Wahlberg N, Braby MF, Brower AVZ, de Jong R, Lee MM, Nylin S, Pierce N, Sperling FA, Vila R, Warren AD, Zakharov E. 2005. Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. *Proceedings of the Royal Society B: Biological Sciences,* **272**(1572): 1577-1586.

Wang XC, Sun XY, Sun QQ, Zhang DX, Hu J, Yang Q, Hao JS. 2011. Complete mitochondrial genome of the laced fritillary *Argyreus hyperbius* (Lepidoptera: Nymphalidae). *Zoological Research,* **32**(5): 465-475.

Weller SJ, Pashley DP, Martin JA, Constable JL. 1996. Reassessment of butterfly family relationships using independent genes and morphology. *Annals of the Entomological Society of America,* **89**(2): 184-192.

Wolstenholme DR. 1992. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology,* **141**: 173-216.

Xia X, Xie Z. 2001. DAMBE: Software package for data analysis in molecular biology and evolution. *Journal of Heredity,* **92**(4): 371-373.

Yang L, Wei ZJ, Hong GY, Jiang ST, Wen LP. 2009. The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae). *Molecular Biology Reports,* **36**(6): 1441-1449.

Yukuhiro K, Sezutsu H, Itoh M, Shimizu K, Banno Y. 2002. Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkmoth, *Bombyx mandarina*, and its close relative, the domesticated silkmoth, *Bombyx mori*. *Molecular Biology and Evolution,* **19**(8): 1385-1389.

Zou FZ, Hao JS, Huang DY, Zhang DX, Zhu GP, Zhu CD. 2009. Molecular phylogeny of 12 families of the Chinese butterflies based on mitochondrial ND1 and 16S rRNA gene sequences (Lepidoptera: Ditrysia: Rhopalocera). *Acta Entomologica Sinica,* **52**(2): 191-201. (in Chinese).