The Complete Mitochondrial Genome of *Leucoptera malifoliella* Costa (Lepidoptera: Lyonetiidae)

Yu-Peng Wu,^{1,2,3} Jin-Liang Zhao,^{2,*} Tian-Juan Su,^{2,4,*} Jie Li,⁵ Fang Yu,² Douglas Chesters,²
Ren-Jun Fan,⁶ Ming-Chang Chen,⁷ Chun-Sheng Wu,² and Chao-Dong Zhu²

The mitochondrial genome (mitogenome) of *Leucoptera malifoliella* (= *L. scitella*) (Lepidoptera: Lyonetiidae) was sequenced. The size was 15,646 bp with gene content and order the same as those of other lepidopterans. The nucleotide composition of *L. malifoliella* mitogenome is highly A + T biased (82.57%), ranked just below *Coreana raphaelis* (82.66%) (Lepidoptera: Lycaenidae). All protein-coding genes (PCGs) start with the typical ATN codon except for the *cox1* gene, which uses CGA as the initiation codon. Nine PCGs have the common stop codon TAA, four PCGs have the common stop codon T as incomplete stop codons, and *nad4l* and *nad6* have TAG as the stop codon. Cloverleaf secondary structures were inferred for 22 tRNA genes, but *trnS1*(AGN) was found to lack the DHU stem. The secondary structure of *rrnL* and *rrnS* is generally similar to other lepidopterans but with some minor differences. The A + T-rich region includes the motif ATAGA, but the poly (T) stretch is replaced by a stem-loop structure, which may have a similar function to the poly (T) stretch. Finally, there are three long repeat (154 bp) sequences followed by one short repeat (56 bp) with four (TA)_n intervals, and a 10-bp poly-A is present upstream of *trnM*. Phylogenetic analysis shows that the position of Yponomeutoidea, as represented by *L. malifoliella*, is the same as traditional classifications. Yponomeutoidea is the sister to the other lepidopteran superfamilies covered in the present study.

Introduction

MITOGENOMES ARE MATERNALLY inherited, with a covalently closed double-stranded DNA structure, a relatively stable arrangement of genes, and are broadly applied in phylogenetic reconstruction, phylogeography, and molecular evolution (Zhang *et al.*, 1995; Nardi *et al.*, 2003; Arunkumar *et al.*, 2006). Animal mitogenomes are generally 14–20 kb in length, including 37 genes—13 being protein-coding genes (PCGs), 22 transfer RNAs (tRNA), and 2 genes for ribosomal RNA (rRNA). They also contain an A + T-rich noncoding area that regulates transcription and replication of the mitogenome (Boore, 1999; Taanman, 1999) and influences the length of mitogenome (Wolstenholme, 1992). As sequencing technology develops, there is a rapid growth of mitogenome data in GenBank. To date, there have been more than 30 Lepidoptera species sequenced in their entirety or to near completion (Coates *et al.*, 2005; Kim *et al.*, 2006; Lee *et al.*, 2006; Cha *et al.*, 2007; Cameron and Whiting, 2008;

Hong *et al.*, 2008; Liu *et al.*, 2008; Pan *et al.*, 2008; Salvato *et al.*, 2008; Hong *et al.*, 2009; Jiang *et al.*, 2009; Kim MI *et al.*, 2009; Hu *et al.*, 2010; Li *et al.*, 2010; Liao *et al.*, 2010; Zhao *et al.*, 2010; Kim MJ *et al.*, 2010; Margam *et al.*, 2011).

The Lepidoptera includes moths and butterflies, with more than 160,000 described species distributed worldwide in 124 families (Kristensen *et al.*, 2007). *Leucoptera malifoliella* Costa belongs to Lyonetiidae (Lepidoptera). Its junior synonym is *Leucoptera scitella* Zeller (Mey, 1994). Lyonetiidae is a microlepidopteran family with more than 500 described species. These moths are small and slender, with a very narrow forewing and a wingspan, which rarely exceeds 1 cm. Their larvae are generally leaf miners. *L. malifoliella* damage apples, pears, and other plants, and withering fruit tree leaves. Recent studies focused on the sex pheromone for prevention and control (Francke *et al.*, 1987; Koutinkova *et al.*, 1999), but few studies have been carried out on their mitogenome, and there is a lack of mitogenome data.

¹Institute of Loess Plateau, Shanxi University, Taiyuan, China.

²Key Laboratory of Zoological Systematics and Evolution (CAS), Institute of Zoology, Chinese Academy of Sciences, Beijing, China.

³Plant Protection and Quarantine Station of Shanxi Province, Taiyuan, China.

⁴College of Life Sciences, Capital Normal University, Beijing, China.

⁵Pomology Institute, Shanxi Academy of Agricultural Sciences, Taigu, China.

⁶Institute of Plant Protection, Shanxi Academy of Agricultural Sciences, Taiyuan, China.

⁷Shanxi Academy of Agricultural Sciences, Taiyuan, China.

*These two authors contributed equally to this work.

In this study, we described the complete mitogenome of *L. malifoliella*, the first sequenced specie of Lyonetiidae, and compared its features with other available lepidopteran mitogenomes. Finally, the phylogenetic relationships among the lepidopteran superfamilies were reconstructed using complete mitochondrial genomes.

Materials and Methods

DNA sample extraction

Larvae were collected from an orchard in Beijing, China, *L. malifoliella* were identified according to Mey (1994), and raised in laboratory. The hatched moths were collected, preserved in 100% ethanol, and stored at -20°C . Total DNA was extracted and isolated from single specimens using the DNeasy Tissue kit (QIAGEN) according to the manufacturer's instructions.

Primer design, polymerase chain reaction, and sequencing

Short fragment amplifications were performed using the universal polymerase chain reaction (PCR) primers from Simon *et al.* (1994). The degenerate and specific primer pairs were designed based on the known mitochondrial sequences in Lepidoptera, or designed by Primer5.0 software on the fragments that we previously sequenced (Table 1). All the primers were synthesized by Shanghai Sangon Biotechnology Co., Ltd. (Beijing, China). For fragments of length less than 2 kb, PCR conditions were as follows: 95°C for 5 min; 34 cycles of 94°C for 30 s; 50°C – 55°C (depending on primer combinations), 1–3 min (depending on putative length of the fragments) at 68°C ; and a final extension step of 72°C for 10 min. For fragments of length more than 2 kb, PCR conditions were as follows: 92°C for 2 min; 40 cycles of 92°C for 30 s, 50°C – 55°C for 30 s (depending on primer combinations), 60°C for 12 min; and a final extension step of 60°C for 20 min.

The entire mitogenome of *L. malifoliella* was amplified in 14 fragments. For most fragments, we used $2\times$ Taq PCR MasterMix (Tiangen Biotech Co., Ltd., Beijing, China) in the amplification; for fragments longer than 2 kb (*cox2-nad5* and *nad5-cob*) and with higher AT contents (A+T-rich region), amplification used Takara LA Taq (Takara Co., Dalian, China). All amplifications were performed on an Eppendorf Mastercycler and Mastercycler gradient in 50- μL reaction volumes. The reaction volume of $2\times$ Taq PCR MasterMix contains 22 μL sterilized distilled water, 25 μL $2\times$ Master Mix, 1 μL of each primer (10 μM), and 1 μL of DNA template; the one of Takara LA Taq consists of 26.5 μL of sterilized distilled water, 5 μL of $10\times$ LA PCR Buffer II (Takara), 5 μL of 25 mM MgCl_2 , 8 μL of dNTPs Mixture, 2 μL of each primer (10 μM), 1 μL of DNA template, and 0.5 μL (1.25 U) of TaKaRa LA Taq polymerase (Takara).

The PCR products were detected via electrophoresis in 1% agarose gel, purified using the 3S Spin PCR Product Purification Kit, and sequenced directly with ABI-377 automatic DNA sequencer. All fragments were sequenced from both strands. Short amplified products were sequenced directly by internal primers, long amplified products were sequenced completely by primer walking, but the *rrnS-nad2* regions were sequenced after cloning. The purified PCR products were ligated to the *pEASY-T3* Cloning Vector (Beijing

TABLE 1. REGION, PRIMERS, AND SEQUENCES FOR POLYMERASE CHAIN REACTIONS IN THIS STUDY

Region	Upstream primer	Upstream primer sequence (5'-3')	Downstream primer	Downstream primer sequence (5'-3')	Size (bp)
<i>trnQ-cox1</i>	Gln-J-518 ^a	ATTAAGGCCAATTATGGGAGCT	Cl-N-1391 ^a	AATCTGAGTATCGACGTGGTA	2000
<i>trnQ-nad2</i>	Gln0486 ^b	TAAACTATATCTAATAATATCAAAAATTATTGTG	ND2-N-784 ^a	TTTAATCCTCCGATAGCTCCAAT	600
<i>nad2</i>	Gln-J-171 ^a	ATTTACTTAGATTATCCCCC	C1-N-1031 ^a	GAAGGGGGAGAGTCAGAAAT	1200
<i>cox1</i>	LCO1490 ^a	GGTCAACAATCATATAAGATATTGG	HCO2198 ^a	TAACTTCAGGGTGACCAAAAATCA	650
<i>cox1-cox2</i>	C1-J-2183 ^a	CAACATTATTGATTTTGG	C2-N-3494 ^a	GGTAAAACACTACGATTATCAAC	1300
<i>trnL2-trnK</i>	Leu-J-3029 ^c	CTAATATGGCAGACTATATGTAATGGA	Lys14111R ^e	GACCAATTACTTCTTCAGTCATCTAATG	750
<i>cox2-nad5</i>	C2-J-2085 ^a	TAGATAATCGCATTTTACCCCT	N5-N-2826 ^a	TGCTTTCAGGGGATATTTTGTTA	5000
<i>nad5-cob</i>	N5-J-532 ^a	ATGAGCAACTGAAGAATAAGC	Cytb-N-989 ^a	ATAATAGGGATGGAAAGGAAT	2500
<i>Cob</i>	CB-J-10933 ^c	TATGTTTCTCTGAGGACAAATATC	CB-N-11367 ^c	TAACTCCTCTAAATTTATGGGA	460
<i>cob-nad1</i>	Cytb-J-625 ^a	TCGACCTGTAGAAGAACCCCTA	N1-N-57 ^a	CAAAATTATGCTTTATTAGGAGGT	730
<i>cob-rrnL</i>	CB-5971 ^b	CAACACAGGATCTAATAACCTTTAGG	LR-N-12866 ^d	ACATGATCTGAGTTCAACCCGG	1800
<i>rrnL-rrnS</i>	Lr-J-2275 ^a	AACCTCTATAGGCTCTCTCGT	Sr-N-3095 ^a	GAGATACTGGAAAGTGTTC	1100
<i>rrnS</i>	SR-J-14233 ^c	GAAAGCAGCGGCAATATG	SR-N-14588 ^d	AACTAGGATTAGTACCCCTATTAT	350
<i>rrnS-nad2</i>	Sr-J-6512 ^a	TCTCTACTTTGTACGACTTA	Gln-N-193 ^a	AAGGGGATAAATCTAAGTAA	1100

^aPrimers newly designed for this genome.

^bPrimers from Lee *et al.* (2006).

^cPrimers from Zhao *et al.* (2010).

^dPrimers from Simon *et al.* (1994).

TABLE 2. CHARACTERISTICS OF THE LEPIDOPTERAN MITOGENOMES

Superfamily	Species	Size (bp)	PCG A+T content (%)		rmL A+T content (%) 16S		rmS A+T content (%) 12S		A+T-rich region		GenBank accession no.
			A+T content (%)	A+T (%)	Size (bp)	A+T (%)	Size (bp)	A+T (%)			
Yponomeutoidea	<i>Corcyra cephalonica</i>	15273	82.57	80.24	1351	85.49	87.14	770	733	95.36	JN790955
	<i>Chilo suppressalis</i>	15395	80.43	78.96	1355	82.95	85.86	778	351	96.58	HQ897685
Pyraloidea			80.67	78.9	1383	84.24	86.17	788	348	95.4	JF339041
	<i>Diatraea saccharalis</i>	15490	80.02	77.9	1412	84.77	85.53	781	335	94.43	FJ240227
	<i>Ostrinia nubilalis</i>	14535	80.17	79.16	1339	84.91	82.03	434			AF442957
	<i>Ostrinia furnacalis</i>	14536	80.37	79.42	1339	84.99	82.76	435			AF467260
	<i>Antheraea pernyi</i>	15566	80.16	78.53	1369	83.86	84.13	775	552	90.4	AY242996
	<i>Antheraea yamanai</i>	15338	80.29	78.94	1380	83.99	84.41	776	334	89.52	EU726630
	<i>Bombyx mandarina</i>	15928	81.68	79.64	1377	84.75	85.95	783	747	95.91	AB070263
Bombycoidea	<i>Bombyx mori</i>	15643	81.32	79.57	1375	84.36	85.57	783	499	95.39	AF149768
	<i>Eriogyna pyretorum</i>	15327	80.82	79.41	1338	84.6	84.45	778	358	92.18	FJ685653
	<i>Manduca sexta</i>	15516	81.79	80.3	1391	85.26	85.71	777	324	95.37	EU286785
	<i>Saturnia boisduvalii</i>	15360	80.63	79.15	1391	84.76	84.11	774	330	91.52	EF622227
	<i>Phthonandria atrilineata</i>	15499	81.02	79.1	1400	85.71	86.03	803	457	98.25	EU569764
Tortricoidea	<i>A doxophyes honmai</i>	15680	80.39	78.48	1387	83.56	85.37	779	490	94.29	DQ073916
	<i>Grapholita molesta</i>	15717	80.87	78.89	1377	84.75	85.36	772	771	95.85	HQ392511
	<i>Spilonota lechriaspis</i>	15368	81.19	79.72	1382	85.17	86.25	778	441	92.74	HM204705
	<i>Helicoverpa armigera</i>	15347	80.97	79.43	1395	84.73	85.89	794	328	95.12	GU188273
	<i>Hyphantria cunea</i>	15481	80.39	78.95	1426	84.99	84.53	808	357	94.96	GU592049
	<i>Lymantria dispar</i>	15569	79.88	77.84	1351	84.23	85.23	799	435	96.09	FJ617240
	<i>Ochrogaster lunifer</i>	15593	77.84	75.73	1351	81.5	83.25	806	319	93.42	AM946601
Noctuoidea	<i>Sesamia inferens</i>	15413	80.24	78.62	1385	83.39	85.33	784	311	95.82	JN039362
	<i>Acraea issoria</i>	15245	79.76	78.11	1331	83.85	83.76	788	430	96.05	GQ376195
	<i>Apatura metis</i>	15236	80.44	78.96	1333	84.47	84.85	779	394	92.89	JF801742
	<i>Artogeia melete</i>	15140	79.78	78.52	1319	83.47	85.46	777	351	89.17	EU597124
	<i>Pieris rapae</i>	15157	79.74	78.28	1320	84.01	84.95	764	393	91.61	GQ398376
	<i>Calinaga dauidis</i>	15267	80.45	78.94	1337	83.84	85.9	773	389	92.03	HQ658143
	<i>Coreana raphaelis</i>	15314	82.66	81.51	1330	85.26	85.84	777	375	94.13	DQ 102703
Papilionoidea	<i>Hipparchia autonoe</i>	15489	79.09	76.91	1335	83.67	85.29	775	678	94.54	GQ868707
	<i>Papilio maraho</i>	16094	80.5	78.17	1333	83.72	85.49	779	1270	94.62	FJ810212
	<i>Parnassius breneri</i>	15389	81.27	80.18	1344	83.93	85.12	773	504	93.65	FJ871125
	<i>Sasakia charonda</i>	15244	79.87	78.22	1323	84.35	85.03	775	380	91.84	AP011824
	<i>Sasakia charonda kuriyamaensis</i>	15222	79.89	78.3	1311	84.21	85.03	775	380	91.84	AP011825
	<i>Teinopalpus aureus</i>	15242	79.81	78.31	1320	82.42	85.66	781	395	93.16	HM563681

TABLE 3. SUMMARY OF MITOGENOME OF *LEUCOPTERA MALIFOLIELLA*

Gene	Direction	Location	Size (bp)	Anticodon	Start codon	Stop codon
<i>trnM</i>	F	1–66	66	CAT		
<i>trnI</i>	F	68–139	72	GAT		
<i>trnQ</i>	R	137–205	69	TTG		
Spacer 1	N/A	206–248	43			
<i>nad2</i>	F	249–1257	1009		ATT	T
<i>trnW</i>	F	1258–1325	68	TCA		
<i>trnC</i>	R	1318–1392	75	GCA		
<i>trnY</i>	R	1395–1459	65	GTA		
<i>cox1</i>	F	1464–2994	1531		CGA	T
<i>trnL2(UUR)</i>	F	2995–3062	68	TAA		
<i>cox2</i>	F	3063–3744	682		ATA	T
<i>trnK</i>	F	3745–3815	71	TTT		
<i>trnD</i>	F	3818–3883	66	GTC		
<i>atp8</i>	F	3884–4045	162		ATT	TAA
<i>atp6</i>	F	4039–4716	678		ATG	TAA
<i>cox3</i>	F	4720–5508	789		ATG	TAA
<i>trnG</i>	F	5515–5581	67	TCC		
<i>nad3</i>	F	5582–5935	354		ATT	TAA
Spacer2		5936–5954	19			
<i>trnA</i>	F	5955–6017	63	TGC		
<i>trnR</i>	F	6017–6083	67	TCG		
<i>trnN</i>	F	6083–6151	69	GTT		
<i>trnS1(AGN)</i>	F	6150–6211	68	GCT		
<i>trnE</i>	F	6214–6280	67	TTC		
<i>trnF</i>	R	6279–6343	65	GAA		
<i>nad5</i>	R	6344–8069	1726		ATT	T
Spacer3	N/A	8070–8090	21			
<i>trnH</i>	R	8091–8157	67	GTG		
<i>nad4</i>	R	8157–9497	1341		ATG	TAA
<i>nad4L</i>	R	9503–9781	279		ATG	TAG
<i>trnT</i>	F	9793–9859	67	TGT		
<i>trnP</i>	R	9860–9924	65	TGG		
<i>nad6</i>	F	9927–10448	522		ATA	TAG
<i>cob</i>	F	10460–11614	1155		ATG	TAA
Spacer4	N/A	11615–11628	14			
<i>trnS2(UCN)</i>	F	11629–11695	67	TGA		
Spacer5	N/A	11696–11714	19			
<i>nad1</i>	R	11715–12650	936		ATA	TAA
<i>trnL1(CUN)</i>	R	12651–12720	70	TAG		
<i>rrnL</i>	R	12721–14071	1351			
<i>trnV</i>	R	14072–14143	72	TAC		
<i>rrnS</i>	R	14144–14913	770			
A + T-rich region		14914–15646	733			

TransGen Biotech Co., Ltd., Beijing, China), and then sequenced by M13-F and M13-R primers and walking. Sequencing was performed using ABI BigDye ver 3.1 dye terminator sequencing technology and run on ABI PRISM 3730×1 capillary sequencers.

Analysis and annotation

Sequence annotation was performed using the DNASTar package (DNASTar Inc. Madison, WI). The tRNA genes were identified using the tRNAscan-SE v.1.21 software (Lowe and Eddy, 1997). The putative tRNAs were then confirmed by sequence alignment with other insects of lepidoptera using the Bioedit (Hall, 1999). Secondary structure was inferred using DNA-SIS 2.5 (Hitachi Engineering, Tokyo, Japan). The *trnS1(AGN)* secondary structure was developed as proposed by Steinberg and Cedergren (1994). *rrnL* and *rrnS*

secondary structures were drawn by XRNA (developed by B. Weiser and available at <http://rna.ucsc.edu/rnacenter/xrna/xrna.html>). Helix numbering follows the convention established at the CRW site (Cannone *et al.*, 2002) and *Grapholita molesta* rRNA secondary structure (Gong *et al.*, 2011), with minor modification. The stem-loop structure of A + T-rich region was determined by the Mfold Web Server (Zuker, 2003; <http://mfold.rna.albany.edu/?q=mfold>). PCGs and rRNAs were identified by similarity to other lepidopterans. The nucleotide sequences of PCGs were translated based on the invertebrate mtDNA genetic code. Nucleotide composition and codon usage were calculated using MEGA5.0 (Tamura *et al.*, 2011).

Phylogenetic analysis

To infer the phylogenetic relationships of lepidopterans, other available complete mitogenomes in Lepidoptera were

obtained from GenBank. *Bactrocera oleae* (NC_005333) (Nardi *et al.*, 2003) and *Anopheles gambiae* (NC_002084) (Beard *et al.*, 1993) were used as outgroups. The alignment of the amino acid sequences and nucleotide sequences of each of the 13 mitochondrial PCGs was performed with MUSCLE (Edgar, 2004) using default settings, and concatenated into an amino acid (3,872 sites in length) and nucleotide (11,616 sites in length) matrix. The concatenated set of amino acid sequences and nucleotide sequences were used in phylogenetic analyses, using Bayesian Inference (BI) and Maximum Likelihood (ML) methods. Substitution model selection was conducted via a comparison of Akaike Information Criterion scores (Akaike, 1974), calculated using the programs ProTest ver. 1.4 (Abascal *et al.*, 2005) for amino acid sequence alignment and Modeltest ver. 3.7 (Posada and Crandall, 1998) for nucleotide sequence alignment. The MtRev (Adachi and Hasegama, 1996)+I+G model and GTR (Lanave *et al.*, 1984)+I+G model were chosen as the best-fitting model for amino acid sequences and nucleotide sequences, respectively, for BI analyses and ML analyses. The BI analysis was conducted using MrBayes 3.1 (Huelsenbeck and Ronquist,

2001) with four independent Markov chains run for 1,000,000 metropolis-coupled MCMC generations, with tree sampling every 100 generations and a burn-in of 2000 trees. The ML analysis was performed using RAXML (Stamatakis, 2006) with 1000 bootstrap replicates.

Results and Discussion

Genome structure and organization

The *L. malifoliella* mitogenome is a circular molecule of 15,646 bp in length, deposited in GenBank under accession number JN790955. The *L. malifoliella* mitogenome showed the typical metazoan gene content, containing 13 PCGs, 2 rRNAs, 22 tRNAs, and noncoding regions. The gene order in *L. malifoliella* is A+T-rich region-*trnM-trnI-trnQ*, whereas the ancestral gene order for the Lepidoptera is A+T-rich region-*trnI-trnQ-trnM* (Junqueira *et al.*, 2004). This placement of *trnM* may be a molecular feature exclusive to lepidopteran mitogenomes (Cameron and Whiting, 2008).

The *L. malifoliella* mitogenome is biased toward A+T (82.57%) with the value falling into lepidopteran range of

	<i>trnY</i> ←	<i>cox1</i> →
<i>Leucoptera malifoliella</i>	TTTATCGCTTAACTCAGCCATTTTATTT	TTG CGA AAATGATTT
<i>Saturnia boisduvalii</i>	TTTATCGCTTTTACTCAGCCATTTTATTAATTATTTAAT	TTG CGA AAATGACTT
<i>Phthonandria atrilineata</i>	TTTATCGCTTAAATCTCAGCCATTTTATT	TTG CGA AAATGACTA
<i>Manduca sexta</i>	TTTATCGCTTAAACTCAGCCATTTTATTT	TTG CGA AAATGACTT
<i>Chilo suppressalis</i>	TTTATCGCTTAAACCTCAGCCATTTTATTTTA	TTG CGA AAATGACTT
<i>Corcyra cephalonica</i>	TTTATCGCTTATACCTCAGCCATTTTATTTT	TTG CGA AAATGACTA
<i>Calinaga davidis</i>	TTTATCGCTTATACCTCAGCCATTTTAT	TTG CGA AAATGACTT
<i>Grapholita molesta</i>	TTTATCGCCTAAACTTCAGCCATTTTATTCTAT	TAG CGA AAATGACTT
<i>Pieris melete</i>	TTTATCGCTTATAACTCAGCCATTTTAT	TAG CGA AAATGACTT
<i>Pieris rapae</i>	TTTATCGCTTATACCTCAGCCATTTTATTTA	AAG CGA AAATGACTT
<i>Spilonota lechriaspis</i>	TTTATCGCTTATTCTCAGCCATTTTATT	TAG CGA AAATGAATC
<i>Helicoverpa armigera</i>	TTTATCGCTTAAACTCAGCCATTTTAT	TAG CGA AAATGACTT
<i>Hipparchia autonoe</i>	TTTATCGCTTATCCCTCAGCCATTTTACTATTTT	AAG CGA AAATGACTT
<i>Acraea issoria</i>	TTTATCGCTTATTAACTCAGCCATTTTAT	TTG CGA AAATGACTT
<i>Teinopalpus aureus</i>	TTTATCGCTTAAATTTTCAGCCATTTTAT	TAG CGA AAATGATTT
<i>Hyphantria cunea</i>	TTTATCGCTTATAACTCAGCCATTTTATTTTTTTTTT	TAG CGA AAATGACTT
<i>Papilio maraho</i>	TTTATCGCTTATATTTTCAGCCATTTTAT	TAG CGA AAATGACTT
<i>Parnassius bremeri</i>	TTTATCGCTTATACTCAGCCATTTTAT	TAG CGA AAATGACTT
<i>Diatraea saccharalis</i>	TTTATCGCTTATAACTCAGCCATTTTATT	TAG CGA AAATGACTA
<i>Antheraea yamamai</i>	TTTATCGCTTATTAACCTCAGCCATTTTATTAAT	TAG CGA AAATGACTT
<i>Lymantria dispar</i>	TTTACCGCTTAAACTCAGCCATTTTATTAATAAAT	AAG CGA AAATGATTA
<i>Eriogyna pyretorum</i>	TTTATCGCTTATTAACTCAGCCATTTTATTAATAAATTATA	TAG CGA AAATGACTT
<i>Antheraea pernyi</i>	TTTATCGCTTTAAACCTCAGCCATTTTATTAAT	TAG CGA AAATGACTT
<i>Ostrinia furnacalis</i>	TTTATCGCTTAAATCTCAGCCATTTTATTTTAATT	TAG CGA AAATGACTA
<i>Ostrinia nubilalis</i>	TTTATCGCTTAAATCTCAGCCATTTTATTTTAATT	TAG CGA AAATGACTA
<i>Ochrogaster lunifer</i>	TTTATCGCCTACCCTCAGCCATTTTATTTATTTT	TAG CGA AAATGACTA
<i>Coreana raphaelis</i>	TTTATCGCTTATATCTCAGCCATTTTATT	TAG CGA AAATGACTT
<i>Adoxophyes honmai</i>	TTTATCGCTTAAACCTCAGCCATTTTATTCTAT	TAG CGA AAATGACTT
<i>Sesamia inferens</i>	TTTATCGCTTATTAACTCAGCCATTTTATTTT	TAG CGA AAATGACTT
<i>Apatura metis</i>	TTTATCGCTTATAACTCAGCCATTTTATTAT	AAG CGA AAATGATTA
<i>Bombyx mandarina</i>	TTTATCGCTTATTCTTCAGCCATTTTATTACTTACTTTATT	TAG CGA AAATGAATT
<i>Bombyx mori</i>	TTTATCGCTTATTATTCAGCCATTTTATTTTTTTCTTTT	TAG CGA AAATGAATT
<i>Sasakia charonda</i>	TTTATCGCTTATAACTCAGCCATTTTATTAT	AAG CGA AAATGATTA
<i>Sasakia charonda kuryamaensis</i>	TTTATCGCTTATAACTCAGCCATTTTATTAT	AAG CGA AAATGATTA

FIG. 1. Alignment result of *trnY* and *cox1* in 34 Lepidopterans. The dotted line and underline indicate the locations of *cox1* and *trnY*, respectively. The overlapping base between *trnY* and *cox1* is marked gray.

FIG. 2. Putative secondary structures for the tRNA genes of *Leucoptera malifoliella* mitochondrial genome.

77.84% (*Ochrogaster lunifer*, Salvato *et al.*, 2008) to 82.66% (*Coreana raphaelis*, Kim *et al.*, 2006). The A+T content was 80.24% in PCGs, 85.49% in *rrnL* genes, 87.14% in *rrns* genes, and 95.36% in the A+T-rich region. These values were also high in other lepidopterans reported (Table 2).

Protein-coding genes

The initial and termination codons of 13 PCGs are shown in Table 3. Twelve PCGs start with a typical ATN codon (ATT for *nad2*, *nad3*, *nad5*, *atp8*; ATA for *cox2*, *nad6*, *nad1*; ATG for *atp6*, *cox3*, *nad4*, *nad4l*, *cob*). The exception is the *cox1* gene, which uses CGA as the start codon. Seven PCGs have the common stop codon TAA, *nad4l* and *nad6* have the common stop codon TAG, and four PCGs have the codon T as incomplete stop codons, which was also found in other animal mitochondrial genes (Clary and Wolstenholme, 1985).

The putative start codons of PCGs in the *L. malifoliella* mitogenome are ATN, except for the CGA start codon of the *cox1* gene. The start codon of the *cox1* gene is controversial in

many studies. The putative codon CGA is common across insects (*Anabrus simplex*, Fenn *et al.*, 2007; *Adoxophyes hnmiai*, Lee *et al.*, 2006; *Manduca sexta*, Cameron and Whiting, 2008; *O. lunifer*, Salvato *et al.*, 2008; *Eriogyna pyretorum*, Jiang *et al.*, 2009; *Phthonandria atrilineata*, Yang *et al.*, 2009; *Hyphantria cunea*, Liao *et al.*, 2010; *Artogeia melete*, Hong *et al.*, 2009; *Antheraea yamamai*, Kim SR *et al.*, 2009; *Eumenis autonoe*, Kim *et al.*, 2010). The tetranucleotides TTAG and hexanucleotide TATTAG have also been proposed as start codons for the *cox1* gene (*Parnassius bremeri*, Kim *et al.*, 2009; *C. raphaelis*, Kim *et al.*, 2006; *Antheraea pernyi*, Liu *et al.*, 2008; *Ostrinia nubilalis*, *Ostrinia furnacalis*, Coates *et al.*, 2005; *Bombyx mandarina*, Yukuhiro *et al.*, 2002; *Papilio xuthus*, Feng *et al.*, 2010). However, TTAG lacks absolute conservation and may serve alternative functions, not always as an initiation codon. Alignment of the mitogenome sequence from all Lepidopterans had shown that an arginine (CGR) functions as the start codon for the *cox1* gene (Margam *et al.*, 2011). In our study the start codon of the *cox1* gene is CGA according to the alignment of the lepidoterans (Fig. 1).

Transfer and ribosomal RNA genes

The 22 tRNA genes ranged from 63 to 75 nucleotides. Fourteen tRNAs are coded on the J-strand and 8 on the N-strand, as with other Lepidoptera. The *trnK* anticodon is TTT, which is unusual in this insect order. Complete cloverleaf secondary structures could be inferred for 21 of the 22 tRNAs. The secondary structure of *trnS1*(AGN) was incomplete, lacking the DHU arm (Fig. 2). A total of 43 unmatched base pairs were scattered throughout the 21 tRNA genes, including 15 pairs in the DHU stems, 11 pairs in the amino acid acceptor stems, 9 pairs in the TΨC stems, and 8 pairs in the anticodon stems. Twenty-one of these are G-U pairs, which form a stable hydrogen-bonded pair. The remaining were C-A, C-U, G-G, G-A, and U-U mismatches.

As in the other insect mitogenome sequences, two rRNA genes were present in *L. malifoliella*. The *rrnL* gene (1351 bp) was found between *trnL*(CUN) and *trnV*, and the *rrnS* (770 bp) between *trnV* and the A+T-rich region. Both the secondary structure of *rrnL* and *rrnS* conform to the models proposed for other insects (Cameron and Whiting, 2008; Wei *et al.*, 2009; Wei *et al.*, 2010). Forty-nine helices are present in *rrnL* of *L. malifoliella*, as in *G. molesta* (Gong *et al.*, 2011), *M. sexta* (Cameron and Whiting, 2008), *Drosophila melanogaster* (Schnare *et al.*, 1996), and *Apis mellifera* (Gillespie *et al.*, 2006). There is a large internal loop among H991, H1057, and H1087, which is similar to *G. molesta*, and differs from *M. sexta*. The microsatellite sequence of (TA)_n inserted in the loop region of H2347 in *Adoxophyes honmai* (Lee *et al.*, 2006), *Spilonota lechriaspis* (Zhao *et al.*, 2010), *G. molesta*, which belong to Tortricidae, is not present in *L. malifoliella* (Fig. 3).

FIG. 4. Predicted *rrnS* secondary structure in *Leucoptera malifoliella* mitogenome. Tertiary interactions and base triples are shown connected by continuous lines. Base pairing is indicated as follows: Watson-Crick pairs by lines, wobble GU pairs by plus, and other noncanonical pairs by circles.

TABLE 4. THE CODON NUMBER AND RELATIVE SYNONYMOUS CODON USAGE IN *LEUCOPTERA MALIFOLIELLA* MITOCHONDRIAL PROTEIN CODING GENES

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	384	1.9	UCU(S)	118	2.97	UAU(Y)	184	1.8	UGU(C)	27	1.74
UUC(F)	21	0.1	UCC(S)	7	0.18	UAC(Y)	20	0.2	UGC(C)	4	0.26
UUA(L)	472	5.29	UCA(S)	85	2.14	UAA(*)	7	1.56	UGA(W)	85	1.79
UUG(L)	16	0.18	UCG(S)	1	0.03	UAG(*)	2	0.44	UGG(W)	10	0.21
CUU(L)	26	0.29	CCU(P)	53	1.83	CAU(H)	63	1.88	CGU(R)	16	1.31
CUC(L)	3	0.03	CCC(P)	13	0.45	CAC(H)	4	0.12	CGC(R)	4	0.33
CUA(L)	18	0.2	CCA(P)	49	1.69	CAA(Q)	55	1.86	CGA(R)	28	2.29
CUG(L)	0	0	CCG(P)	1	0.03	CAG(Q)	4	0.14	CGG(R)	1	0.08
AUU(I)	457	1.92	ACU(T)	76	2.16	AAU(N)	237	1.84	AGU(S)	19	0.48
AUC(I)	19	0.08	ACC(T)	4	0.11	AAC(N)	20	0.16	AGC(S)	1	0.03
AUA(M)	282	1.87	ACA(T)	61	1.73	AAA(K)	116	1.92	AGA(S)	76	1.91
AUG(M)	19	0.13	ACG(T)	0	0	AAG(K)	5	0.08	AGG(S)	11	0.28
GUU(V)	57	2.11	GCU(A)	72	2.62	GAU(D)	51	1.82	GGU(G)	59	1.25
GUC(V)	2	0.07	GCC(A)	0	0	GAC(D)	5	0.18	GGC(G)	3	0.06
GUA(V)	46	1.7	GCA(A)	35	1.27	GAA(E)	69	1.86	GGA(G)	110	2.33
GUG(V)	3	0.11	GCG(A)	3	0.11	GAG(E)	5	0.14	GGG(G)	17	0.36

A total of 3721 codons were analyzed, excluding the initiation and termination codons.

The amino acids encoded by codons are labeled according to the IUPAC-IUB single-letter amino acid codes.

RSCU, relative synonymous codon usage.

Twenty-nine helices present in *rrnS* of *L. malifoliella* belong to three domains, as in *G. molesta*, *M. sexta*, and *A. mellifera*. The structures of Helix H47, H673, H1303, H1047, H1068, H1074, and H1113 are different from *M. sexta*, but similar to *G. molesta*, with the exception of H47, which has a shorter loop length in *L. malifoliella* compared to *G. molesta* (Fig. 4).

Codon usage

Relative synonymous codon usage values of the *L. malifoliella* mitogenome are summarized in Table 4. The codons CUG, ACG, and GCC were not represented in the coding sequences. The most frequent amino acids in *L. malifoliella* mitochondrial proteins are leucine (14.4%), isoleucine (12.8%), phenylalanine (10.9%), and serine (8.6%).

Noncoding and overlapping region

The *L. malifoliella* mitogenome harbors 16 noncoding regions, ranging from 1 to 43 bp. Intergenic spacer sequences have 5 regions with a length of more than 14 bp. The remaining intergenic spacers were less than 11 bp.

Spacer 1 (43 bp) is located between the *trnQ* and *nad2* genes. This spacer can be taken as lepidopteran feature, not found in other insects. Kim *et al.* (2009) detected high sequence identity between the intergenic spacer sequence and the neighboring *nad2* from several lepidopteran insects; this indicated that the spacer sequence may have originated from a partial duplication of the *nad2* gene.

Spacer 2 (19 bp) is found between *nad3* and *trnA* gene; the spacer is longest in lepidopterans sequenced, and in others it is only 1 to 2 bp. Additionally, this region is generally overlapped in other lepidopterans, such as *B. mandarina*, *B. mori*, *M. sexta*, *O. furnacalis*, *O. nubilalis*.

Spacer 3 (21 bp) is found between the *nad5* and *trnH* genes; the spacer is also found in *A. honmai* (23 bp), *E. pyretorum* (18 bp), *B. mandarina* (18 bp), *B. mori* (21 bp), *A. melete* (18 bp), and *C. raphaelis* (16 bp). Spacer 4 (14 bp) is found between the *cob* and *trnS2(UCN)* genes. This spacer is also present in *A. pernyi* (15 bp), *A. yamamai* (24 bp), *S. boisduvalii* (41 bp), and *M. sexta* (21 bp), and it is shorter in other lepidopterans.

Spacer 5 (19 bp) is between the *trnS2(UCN)* and *nad1* genes, commonly detectable in lepidopterans with size

FIG. 5. The structure of the A + T-rich region of *Leucoptera malifoliella* mitogenome.

FIG. 6. Alignment of motif and Poly(T) in the A+T-rich region of 34 lepidopterans. The Poly(T) stretch is marked gray. Marked box is motif ATAGA. Underline is the stem-loop structure of *Leucoptera malifoliella*.

16–38 bp. This intergenic spacer is conserved for all insects. Most lepidopterans harbor the motif (ATACTAA), except for ATACTAT in *Corcyra cephalonica* (unpublished, HQ897685) and ATCATAT in *Sesamia inferens* (unpublished NC_015835). Similarly, in Hymenoptera there is a 6-bp conserved motif (THACWW) (Wei *et al.*, 2010). In Coleoptera, there is a 5-bp

conserved motif (TACTA) (Sheffield *et al.*, 2008). The motif has been suggested to be a possible mitochondrial transcription termination peptide-binding site (Taanman, 1999).

Overlapping sequences had a total length of 25 bp from 1 to 8 bp, spread over 14 regions. The longest overlapping sequence AAGCCTTA (8 bp) is located between the *trnW* gene and the

FIG. 7. Phylogeny of lepidopteran insects. (A) Phylogenetic trees inferred from amino acid sequences and nucleotide sequences of 13 protein-coding genes (PCGs) of the mitogenome using ML analysis; the numbers above branches are bootstrap percentages; the first and second values are from amino acid sequences and nucleotide sequences, respectively. *Bactrocera oleae* and *Anopheles gambiae* were used as outgroups. (B) Phylogenetic trees inferred from amino acid sequences and nucleotide sequences of 13 PCGs of the mitogenome using Bayesian Inference (BI) analysis; the numbers above branches give posterior probabilities. In the BI tree inferred from amino acid sequences *C. cephalonica* is sister to the clade (Pyraloidea + (Noctuoidea + (Geometroidea + Bombycoidea))); the posterior probabilities are 100% and 90%, and the posterior probabilities of *C. cephalonica* in the BI tree from nucleotide sequences is 100% (not labeled). The other information on the tree is the same as in (A).

A

FIG. 7.

B

FIG. 7. (Continued)

trnC gene. The seven-nucleotide overlap (ATGATAA) is located between *atp8* and *atp6*, which is common in other insects. The remaining overlapping sequences are less than 3 bp.

A+T-rich region

The A+T-rich region of *L. malifoliella* mitogenome is located between *rrnS* and *trnM*, with 95.36% AT nucleotides and a length of 733 bp. There is a motif ATAGA downstream of *rrnS*, but not followed by the typical poly (T) stretch, but replaced by a stem-loop structure (Fig. 5). There are three long repeat (154 bp) sequences followed by one short repeat (56 bp), each preceded by (TA)_n microsatellite regions (Fig. 5). Finally, a 10-bp poly-A is present upstream of *trnM*, a feature common across lepidopterans.

The stem-loop structure in the A+T-rich region was also observed in other insect orders, including Orthoptera, Dip-tera, Plecoptera, Hymenoptera, and Phthiraptera (Brehm *et al.*, 2001; Schultheis *et al.*, 2002; Cameron *et al.*, 2007; Cha *et al.*, 2007; Ye *et al.*, 2008). The stem-loop structure in the A+T-rich region of *Drosophila* was suggested as the site of the initiation of light strand synthesis (Clary and Wolstenholme, 1987), but the position of the stem-loop structure in *L. malifoliella* is found to be same as the poly (T) stretch of other lepidopterans (Fig. 6), a feature only found in Leucoptera. Two species (*A. yamamai* and *S. boisduvalii*) in Lepidoptera have a stem-loop structure, but also possess a poly (T) stretch, and the flanking sequence of the stem-loop structure are conserved, with consensus TATA sequences at the 5' and G(A)_nT at the 3'. The feather is also among other insects (Zhang *et al.*, 1995; Schultheis *et al.*, 2002). In contrast to these insects, there are no conserved sequences flanking both sides of the *L. malifoliella* stem-loop structure, and the location of stem-loop structure is closer to *rrnS*. Ye *et al.* (2008) suggested that the stem-loop structure might have the same function as the poly (T) stretch, if the latter feature is absent. Therefore, the stem-loop structure in *L. malifoliella* may play an important role in recognition of the light strand replication origin, but determining the function needs additional research.

Phylogenetic Relationships

To place the *L. malifoliella* mitogenome relative to other lepidopterans mitogenomes and investigate the phylogenetic relationships among the superfamilies in Lepidoptera, two data sets containing the concatenated amino acid sequences and nucleotide sequences of 13 PCGs were generated. These 34 sequences represent seven superfamilies: Bombycoidea, Geometroidea, Noctuoidea, Papilionoidea, Pyraloidea, Tortricidae, and Yponomeutoidea. According to the most recent consensus view of lepidopteran relationships in Kristensen and Skalski (1999), Papilionoidea, Bombycoidea, Noctuoidea, and Geometroidea are designated as the Macrolepidoptera; Pyraloidea together with Macrolepidoptera are designated as Obectomera; Tortricidae together with Obectomera are designated as Apoditrysia; Yponomeutoidea is the sister to the remaining lepidopteran superfamilies covered in the present study. The BI and ML analyses generate similar topologies, and most major groups were consistently monophyletic apart from Pyraloidea. Three trees all support that *C. cephalonica* is grouped with Pyraloidea; this is same as traditional classifications (Solis, 1997). However, in the BI tree inferred from

amino acid sequences, *C. cephalonica* is sister to the clade (Pyraloidea + (Noctuoidea + (Geometroidea + Bombycoidea))).

In our phylogenetic results, the placement of Yponomeutoidea (as represented by *L. malifoliella*) is the same as the traditional classification, basal to all Lepidoptera, with full nodal support in BI (100%/100%) and ML analyses (100%/100%). Bombycoidea and Geometroidea are sister groups with high nodal support on BI (100%/100%) and ML analyses (90%/92%) (Fig. 7A, B), which is consistent with Yang *et al.* (2009), but differs to the typical morphological results, which give a sister group relationship between the Papilionoidea and Geometroidea. Papilionoidea is the sister of the remaining macrolepidopteran families, in accordance with other studies (Jiang *et al.*, 2009; Yang *et al.*, 2009; Liao *et al.*, 2010). Pyraloidea has a closer relationship to most Macrolepidoptera than Papilionoidea (butterflies), a result confirmed by a recent study (Regier *et al.*, 2009), but different from the traditional classification.

Acknowledgments

Prof. Qi-lian Qin and his lab members (Institute of Zoology, Chinese Academy of Sciences) kindly provided advice and facilities in sequence cloning. We also thank Shu-jun Wei (Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences) and Xiao-he Wang (Institute of Zoology, Chinese Academy of Sciences) for their kind help in data analysis.

This work was supported mainly by grants from the Knowledge Innovation Program of Chinese Academy of Sciences (Grant No. KSXC2-EW-B-02), Public Welfare Project from the Ministry of Agriculture, China (Grant No. 201103024), the National Science Foundation, China (NSFC Grant No. 30870268, 31172048, J0930004) to Chao-dong Zhu and NSFC Grant (No. 31172129) to Chun-Sheng Wu.

Disclosure Statement

No competing financial interests exist.

References

- Abascal, F., Zardoya, R., and Posada, D. (2005). ProTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105.
- Adachi, J., and Hasegawa, M. (1996). Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J Mol Evol* **42**, 459–468.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans Autom Contr* **19**, 716–723.
- Arun Kumar, K.P., Metta, M., and Nagaraju, J. (2006). Molecular phylogeny of silkmoths reveals the origin of domesticated silkmoth, *Bombyx mori* from Chinese *Bombyx mandarina* and paternal inheritance of *Antheraea proylei* mitochondrial DNA. *Mol Phylogenet Evol* **40**, 419–427.
- Beard, C.B., Mills, D., and Collins, F.H. (1993). The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization and comparisons with mitochondrial sequences of other insects. *Insect Mol Biol* **2**, 103–124.
- Boore, J.L. (1999). Animal mitochondrial genomes. *Nucleic Acids Res* **27**, 1767–1780.
- Brehm, A., Harris, D.J., Hernández, M., Cabrera, V.M., Larruga, J.M., Pinto, F.M., and González, A.M. (2001). Structure and evolution of the mitochondrial DNA complete control region

- in the *Drosophila subobscura* subgroup. *Insect Mol Biol* **10**, 573–578.
- Cameron, S.L., Johnson, K.P., and Whiting, M.F. (2007). The mitochondrial genome of the screamer Louse *Bothriometopus* (Phthiraptera: Ischnocera): effects of extensive gene rearrangements on the evolution of the genome. *J Mol Evol* **65**, 589–604.
- Cameron, S.L., and Whiting, M.F. (2008). The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene* **408**, 112–123.
- Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., and Muller, K.M. (2002). The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**, 1471–2105.
- Cha, S.Y., Yoon, H.J., Lee, E.M., Yoon, M.H., Hwang, J.S., Jin, B.R., Han, Y.S., and Kim, I. (2007). The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera: Apidae). *Gene* **392**, 206–220.
- Clary, D.O., and Wolstenholme, D.R. (1985). The mitochondrial DNA molecular of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. *J Mol Evol* **22**, 252–271.
- Clary, D.O., and Wolstenholme, D.R. (1987). *Drosophila* mitochondrial DNA: conserved sequences in the A+T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. *J Mol Evol* **25**, 116–125.
- Coates, B.S., Sumerford, D.V., Hellmich, R.L., and Lewis, L.C. (2005). Partial mitochondrial genome sequence of *Ostrinia nubilalis* and *Ostrinia furnicalis*. *Int J Biol Sci* **1**, 13–18.
- Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.
- Feng, X., Liu, D.F., Wang, N.X., Zhu, C.D., and Jiang, G.F. (2010). The mitochondrial genome of the butterfly *Papilio xuthus* (Lepidoptera: Papilionidae) and related phylogenetic analyses. *Mol Biol Rep* **37**, 3877–3888.
- Fenn, J.D., Cameron, S.L., and Whiting, M.F. (2007). The complete mitochondrial genome of the Mormon cricket (*Anabrus simplex*: Tettigoniidae: Orthoptera) and an analysis of control region variability. *Insect Mol Biol* **16**, 239–252.
- Francke, W., Franke, S., Toth, M., Szöcs, G., Guerin, P., and Arn, H. (1987). Identification of 5,9-dimethylheptadecane as a sex pheromone of the moth *Leucoptera scitella*. *Naturwissenschaften* **74**, 143–144.
- Gillespie, J.J., Johnston, J.S., Cannone, J.J., and Gutell, R.R. (2006). Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Mol Biol* **15**, 657–686.
- Gong, Y.J., Shi, B.C., Kang, Z.J., Zhang, F., and Wei, S.J. (2011). The complete mitochondrial genome of the oriental fruit moth *Grapholita molesta* (Busck) (Lepidoptera: Tortricidae). *Mol Biol Rep* **39**, 2893–2900.
- Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95–98.
- Hong, G.Y., Jiang, S.T., Yu, M., Yang, Y., Li, F., Xue, F.S., and Wei, Z.J. (2009). The complete nucleotide sequence of the mitochondrial genome of the cabbage butterfly, *Artogeia melete* (Lepidoptera: Pieridae). *Acta Biochim Biophys Sin* **41**, 446–455.
- 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., and 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* **30**, 413–49–57.
- Hu, J., Zhang, D.X., Hao, J.S., Huang, D.Y., Cameron, S., and 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* **37**, 431–438.
- Huelsenbeck, J.P., and Ronquist, F. (2001). MrBayes: bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755.
- Jiang, S., Hong, G., Yu, M., Li, N., Yang, Y., Liu, Y., and Wei, Z. (2009). Characterization of the complete mitochondrial genome of the giant silkworm moth, *Eriogyna pyretorum* (Lepidoptera: Saturniidae). *Int J Biol Sci* **5**, 351–365.
- Junqueira, A.C.M., Lessinger, A.C., Torres, T.T., da Silva, F.R., Vettore, A.L., Arruda, P., and Espin, A.M.L.A. (2004). The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae). *Gene* **339**, 7–15.
- Kim, I., Lee, E.M., Seol, K.Y., Yun, E.Y., Lee, Y.B., Hwang, J.S., and Jin, B.R. (2006). The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). *Insect Mol Biol* **15**, 217–225.
- Kim, M.I., Baek, J.Y., Kim, M.J., Jeong, H.C., Kim, K.G., Bae, C.H., Han, Y.S., Jin, B.R., and Kim, I. (2009). Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other Lepidopteran insects. *Mol Cells* **28**, 347–363.
- Kim, M.J., Wan, X.L., Kim, K.G., Hwang, J.S., and Kim, I. (2010). Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). *African J Biotech* **9**, 735–754.
- Kim, S.R., Kim, M.I., Hong, M.Y., Kim, K.Y., Kang, P.D., Hwang, J.S., Han, Y.S., Jin, B.R., and Kim, I. (2009). The complete mitogenome sequence of the Japanese oak silkworm, *Antheraea yamamai* (Lepidoptera: Saturniidae). *Mol Biol Rep* **36**, 1871–1880.
- Koutinkova, H., Andreev, R., Subchev, M., Tóth, M., and Szöcs, G. (1999). Monitoring of the leafminer *Leucoptera scitella* Zell (Lepidoptera: Lyonetidae) by pheromone traps in Bulgaria. *Acta Phytopathologica et Entomologica Hungarica* **34**, 327–331.
- Kristensen, N.P., Scoble, M.J., and Karsholt, O. (2007). Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. *Zootaxa* **1668**, 699–747.
- Kristensen, N.P., and Skalski, A.W. (1999). Phylogeny and paleontology. In: *Lepidoptera: Moths and Butterflies, Evolution, Systematics, and Biogeography, Handbook of Zoology Vol IV*, Part 35. N.P. Kristensen, ed. (De Gruyter, Berlin and New York), pp. 7–25.
- Lanave, C., Preparata, G., Saccone, C., and Serio, G. (1984). A new method for calculating evolutionary substitution rates. *J Mol Evol* **20**, 86–93.
- Lee, E.S., Shin, K.S., Kim, M.S., Park, H., Cho, S., and Kim, C.B. (2006). The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae). *Gene* **373**, 52–57.
- Li, W., Zhang, X., Fan, Z., Yue, B., Huang, F., King, E., and Ran, J. (2010). Structural characteristics and phylogenetic analysis of the mitochondrial genome of the sugarcane borer, *Diatraea saccharalis* (Lepidoptera: Crambidae). *DNA Cell Biol* **00**, 1–6.
- Liao, F., Wang, L., Wu, S., Li, Y.P., Zhao, L., Huang, G.M., Niu, C.J., Liu, Y.Q., and Li, M.G. (2010). The complete mitochon-

- drial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *Int J Biol Sci* **6**, 172–186.
- Liu, Y.Q., Li, Y.P., Pan, M.H., Dai, F.Y., Zhu, X.W., Lu, C., and Xiang, Z.H. (2008). The complete mitochondrial genome of the Chinese oak silkworm, *Antheraea pernyi* (Lepidoptera: Saturniidae). *Acta Biochim Biophys Sin* **40**, 693–703.
- Lowe, T.M., and Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**, 955–964.
- Margam, V.M., Coates, B.S., Hellmich, R.L., Agunbiade, T., Seufferheld, M.J., Sun, W., Ba, M.N., Sanon, A., Binso-Dabire, C.L., Baoua, I., Ishiyaku, M.F., Covas, F.G., Srinivasan, R., Armstrong, J., Murdock, L.L., and Pittendrigh, B.R. (2011). Mitochondrial genome sequence and expression profiling for the legume pod borer *Maruca vitrata* (Lepidoptera: Crambidae). *PLoS One* **6**, e16444.
- Mey, W. (1994). Taxonomische Bearbeitung der westpalaearktischen Arten der Gattung Leucoptera Hübner, [1825], s.l. (Lep.: Lyonetiidae). *Dtsch Entomologische Z N F* **41**, 173–234.
- Nardi, F., Caeapelli, A., Dallai, R., and Frati, F. (2003). The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographic locations. *Insect Mol Biol* **12**, 605–611.
- Pan, M.H., Yu, Q.Y., Xia, Y.L., Dai, F.Y., Liu, Y.Q., Lu, C., Zhang, Z., and Xiang, Z.H. (2008). Characterization of mitochondrial genome of Chinese wild mulberry silkworm, *Bombyx mandarina* (Lepidoptera: Bombycidae). *Sci China Ser C-Life Sci* **51**, 693–701.
- Posada, D., Crandall, K.A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.
- Regier, J.C., Zwick, A., Cummings, M.P., Kawahara, A.Y., 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., and Mitter, C. (2009). Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): an initial molecular study. *BMC Evol Biol* **9**, 280.
- Salvato, P., Simonato, M., Battisti, A., and Negrisol, E. (2008). The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera: Notodontidae). *BMC Genomics* **9**, 331–345.
- Schnare, M.N., Damberger, S.H., Gray, M.W., and Gutell, R.R. (1996). Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23S-like) ribosomal RNA. *J Mol Biol* **256**, 701–719.
- Schultheis, A.S., Weigt, L.A., and Hendricks, A.C. (2002). Arrangement and structural conservation of the mitochondrial control region of two species of Plecoptera: utility of tandem repeat-containing regions in studies of population genetics and evolutionary history. *Insect Mol Biol* **11**, 605–610.
- Sheffield, N.C., Song, H., Cameron, L., and Whiting, M.F. (2008). A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Mol Biol Evol* **25**, 2499–2509.
- Simon, C., Frati, F., Bekenbach, A., Crespi, B., Liu, H., and Flook P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain-reaction primers. *Ann Entomol Soc Am* **87**, 651–701.
- Solis, M.A. (1997). Michaelshaffera gen. n. - a pyraloid taxon lacking an abdominal tympanal organ (Lepidoptera: Pyralidae). *Entomol Scand* **28**, 391–402.
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690.
- Steinberg, S., and Cedergren, R. (1994). Structural compensation in a typical mitochondrial tRNAs. *Nat Struct Biol* **1**, 507–510.
- Taanman, J.W. (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochim Biophys Acta* **1410**, 103–123.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.
- Wei, S.J., Shi, M., He, J.H., Sharkey, M.J., and Chen, X.X. (2009). The complete mitochondrial genome of *Diadegma semiclausum* (Hymenoptera: Ichneumonidae) indicates extensive independent evolutionary events. *Genome* **52**, 308–319.
- Wei, S.J., Tang, P., Zheng, L.H., Shi, M., and Chen, X.X. (2010). The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Evaniidae) has low A + T content and a long intergenic spacer between *atp8* and *atp6*. *Mol Biol Rep* **37**, 1931–1942.
- Wolstenholme, D.R. (1992). Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol* **141**, 173–216.
- Yang, L., Wei, Z.J., Hong, G.Y., Jiang, S.T., and Wen, L.P. (2009). The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* (Lepidoptera: Geometridae). *Mol Biol Rep* **36**, 1441–1449.
- Ye, W., Dang, J.P., Xie, L.D., and Huang, Y. (2008). Complete mitochondrial genome of *Teleogryllus emma* (Orthoptera: Gryllidae) with a new gene order in Orthoptera. *Zoolog Res* **29**, 236–244.
- Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K., and Banno, Y. (2002). Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkworm, *Bombyx mandarina*, and its close relative, the domesticated silkworm, *Bombyx mori*. *Mol Biol Evol* **19**, 1385–1389.
- Zhang, D.X., Szymura, J.M., and Hewitt, G.M. (1995). Evolution and structural conservation of the control region of insect mitochondrial DNA. *J Mol Evol* **40**, 382–391.
- Zhao, J.L., Zhang, Y.Y., Luo, A.R., Jiang, G.F., Cameron, S.L., and Zhu, C.D. (2010). The complete mitochondrial genome of *Spilonota lechriaspis* Meyrick (Lepidoptera: Tortricidae). *Mol Biol Rep* **38**, 3757–3764.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406–3415.

Address correspondence to:
Chao-Dong Zhu, Ph.D.
Key Laboratory of Zoological Systematics and Evolution (CAS)
Institute of Zoology
Chinese Academy of Sciences
Beijing 100101
China

E-mail: zhucd@ioz.ac.cn

Received for publication January 31, 2012; received in revised form July 3, 2012; accepted July 3, 2012.