The complete mitochondrial genome of *Papilio glaucus* and its phylogenetic implications

Jinhui Shen, Qian Cong, Nick V. Grishin

Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390-9050, USA

Department of Biophysics and Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390-8816, USA

**Abstract**

Due to the intriguing morphology, lifecycle, and diversity of butterflies and moths, Lepidoptera are emerging as model organisms for the study of genetics, evolution and speciation. The progress of these studies relies on decoding Lepidoptera genomes, both nuclear and mitochondrial. Here we describe a protocol to obtain mitogenomes from Next Generation Sequencing reads performed for whole-genome sequencing and report the complete mitogenome of *Papilio (Pterourus) glaucus*. The circular mitogenome is 15,306 bp in length and rich in A and T. It contains 13 protein-coding genes (PCGs), 22 transfer-RNA-coding genes (tRNA), and 2 ribosomal-RNA-coding genes (rRNA), with a gene order typical for mitogenomes of Lepidoptera. We performed phylogenetic analyses based on PCG and RNA-coding genes or protein sequences using Bayesian Inference and Maximum Likelihood methods. The phylogenetic trees consistently show that among species with available mitogenomes *Papilio glaucus* is the closest to *Papilio (Agehana) maraho* from Asia.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

**Keywords:** *Papilio glaucus*, Mitochondrial genome, Illumina sequencing, Phylogeny

**Introduction**

The order Lepidoptera contains approximately 160,000 described and half a million estimated species (Kristensen et al., 2007), and it represents one of the most diverse and fascinating group of insects. Recent studies have revealed their potential as model organisms to study the genetics of interesting phenotypic traits in butterflies, such as the Batesian mimicry in swallowtails (Clarke and Sheppard, 1972; Nishikawa et al.,...
List of taxa analyzed in present paper.

| Family       | Species          | Length (bp) | Accession number | References          |
|--------------|------------------|-------------|------------------|---------------------|
| Papilionidae | Papilio glaucus  | 15,306      | KR822739         | This study          |
|              | Papilio bianor   | 15,357      | NC_018040.1      | Unpublished         |
|              | Papilio dardanus | 15,337      | JX313686.2       | Unpublished         |
|              | Papilio maackii  | 15,357      | KC433408.1       | Dong et al. (2013)  |
|              | Papilio maraho   | 15,094      | FJ102121.2       | Unpublished         |
|              | Papilio polytes  | 15,256      | KM014701.1       | Wang et al. (2014a) |
|              | Papilio syfanius | 15,359      | KJ396621.1       | Dong et al. (2014)  |
|              | Parnassius apollo | 15,404      | KF746065.1       | Chen et al. (2014a) |
|              | Parnassius bremeri| 15,389      | NC_014053.1      | Kim et al. (2009)   |
|              | Parnassius imperator | 15,424   | KM507326.1       | Wang et al. (2014b) |
|              | Sericinus montela | 15,242      | HQ259122.1       | Ji et al. (2012)    |
|              | Luehdorfia taibai | 15,553      | KC952673.1       | Lian-Xi et al. (2014)|
|              | Teinopalpus aureus | 15,242    | HM563681.1       | Qin et al. (2012a)  |
| Lycaenidae   | Coreana raphaelis| 15,314      | DQ102703.1       | Kim et al. (2006)   |
|              | Cupido argiades  | 15,330      | KC310728.1       | Zhang et al. (2013b)|
|              | Curetis bulis    | 15,162      | JX262888.1       | Zhang et al. (2013c)|
|              | Lycana hylophila| 15,280      | JX262887.1       | Zhang et al. (2013c)|
|              | Protantigius superans | 15,248  | HQ184265.1       | Kim et al. (2011a)  |
| Riodinidae   | Abisara fylloides | 15,301      | HQ259069.1       | Unpublished         |
|              | Anthocharis bambusarum | 15,180    | HQ259069.1       | Unpublished         |
| Pieridae     | Anthocharis bambusarum | 15,180    | KM102732.1       | Unpublished         |
|              | Aporia crataegi  | 15,140      | JN796473.1       | Park et al. (2012)  |
|              | Apodemia mormo   | 15,262      | JF437925.1       | Chen et al. (2012)  |
|              | Apatura ilia     | 15,134      | JF437925.1       | Chen et al. (2012)  |
|              | Arge palpicornis | 15,131      | KF590547.1       | Wu et al. (2014)    |
|              | Athyma asura     | 15,181      | KF590542.1       | Wu et al. (2014)    |
|              | Athyma cama      | 15,269      | KF590526.1       | Wu et al. (2014)    |
|              | Athyma kasa      | 15,230      | KF590524.1       | Wu et al. (2014)    |
|              | Athyma opalina   | 15,240      | KF590511.1       | Wu et al. (2014)    |
|              | Athyma perius    | 15,277      | KF590528.1       | Wu et al. (2014)    |
|              | Athyma selenophora| 15,208      | KF590529.1       | Wu et al. (2014)    |
|              | Athyma sulphita  | 15,268      | KJ347260.1       | Tian et al. (2012)  |
|              | Bhatadusta austenia | 15,615     | KF590545.1       | Wu et al. (2014)    |
|              | Calinaga davia    | 15,267      | HQ68143.1        | Xia et al. (2011)   |

(continued on next page)
2013), migration in the monarch (Zhan et al., 2011, 2014) and wing pattern development in longwings (Hines et al., 2012; Surridge et al., 2011). The profound diversity and the recent evolutionary radiation of Lepidoptera provide rich materials to study evolution, speciation and adaptation (Engsontia et al., 2014; Zhang et al., 2013a). These studies benefit significantly from decoding the genomes of various Lepidoptera species.

Recently, we published the draft genome of Eastern Tiger swallowtail Papilio (Pterourus) glaucus using next generation sequencing techniques (Cong et al., 2015). This nuclear whole genome was the first reported from the Papilionidae family. Traditional genome assemblers failed to automatically assemble the mitogenome, probably due to the difficulty in distinguishing NGS reads of the mitogenome from those of

### Table 1 (continued)

| Family       | Species                        | Length (bp) | Accession number | References                      |
|--------------|--------------------------------|-------------|------------------|---------------------------------|
| Danaus       | chrysippus                     | 15,236      | KF990637.1       | Gan et al. (2014a)              |
| Danaus       | plexippus                      | 15,314      | KC836923.1       | Servin-Garciduenas and Martinez-Romero (2014) |
| Dicborragia  | nesmachus                      | 15,355      | KF990541.1       | Wu et al. (2014)                |
| Dophila      | evelina                        | 15,320      | KF990532.1       | Wu et al. (2014)                |
| Euploea      | core                           | 15,192      | KF990546.1       | Wu et al. (2014)                |
| Euploea      | midamus                        | 15,187      | KJ865207.1       | Unpublished                     |
| Euploea      | angularis                      | 15,166      | HQ378507.1       | Hao et al. (2013b)              |
| Eusthila     | irrorviscens                   | 15,365      | KF990527.1       | Wu et al. (2014)                |
| Fabriciana   | nerippe                        | 15,200      | JF504707.1       | Kim et al. (2011b)              |
| Hamadryas    | epinome                        | 15,207      | KM378244.1       | Cally et al. (2014)             |
| Heliconius   | cydno                          | 15,367      | KM0068091.1      | Shen and Wang (2014)            |
| Heliconius   | hecale                         | 15,328      | KP100653.1       | Heliconius Genome, 2012; Meng et al., 2014 |
| Heliconius   | melpomone                      | 15,359      | KF990538.1       | Wu et al. (2014)                |
| Hipparchia   | autonoe                        | 15,489      | GQ686507.1       | Kim et al. (2010)               |
| Idenopsis    | similis                        | 15,200      | KJ476729.1       | Gan et al. (2014b)              |
| Issoria      | lathonia                       | 15,172      | HM243590.1       | Unpublished                     |
| Junonia      | almana                         | 15,256      | KF990539.1       | Wu et al. (2014)                |
| Junonia      | orthyra                        | 15,214      | KF998691.2       | Shi et al. (2013a)              |
| Kallima      | inachus                        | 15,183      | JN857943.1       | Qin et al. (2012b)              |
| Lexias       | diera                          | 15,258      | KF990531.1       | Wu et al. (2014)                |
| Libythea     | celtis                         | 15,164      | HQ378508.1       | Unpublished                     |
| Melanargia   | asiatica                       | 15,142      | KF900486.1       | Huang et al. (2014b)            |
| Melanitis    | leda                           | 15,122      | JF90546.1        | Shi et al. (2013b)              |
| Melitaea     | cinxia                         | 15,162      | HM243592.1       | Unpublished                     |
| Neope        | pulaha                         | 15,209      | KF990543.1       | Wu et al. (2014)                |
| Neptis       | philyra                        | 15,164      | KF990552.1       | Wu et al. (2014)                |
| Neptis       | soma                           | 15,130      | KF990533.1       | Wu et al. (2014)                |
| Pandirus     | diptera                        | 15,257      | KF990530.1       | Wu et al. (2014)                |
| Pantoporia   | hordonia                       | 15,603      | KF990534.1       | Wu et al. (2014)                |
| Parantica    | sita                           | 15,211      | KF990544.1       | Wu et al. (2014)                |
| Pararge      | aegeria                        | 15,240      | KJ547676.1       | Teixeira da Costa (2014)        |
| Parasarpia  | dudu                           | 15,236      | KF990537.1       | Wu et al. (2014)                |
| Parthenos    | sylvia                         | 15,249      | KF990550.1       | Wu et al. (2014)                |
| Polyura      | arja                           | 15,363      | KF990540.1       | Wu et al. (2014)                |
| Sasaokia     | charonda                       | 15,233      | JX119051.1       | Wang et al. (2012)              |
| Sasaokia     | funebris                       | 15,233      | JX131328.1       | Wang et al. (2013b)             |
| Tanacra      | julii                          | 15,316      | KF990548.1       | Wu et al. (2014)                |
| Timelaeae    | maculata                       | 15,178      | KC572151.1       | Cao et al. (2013)               |
| Tirumala     | limniace                       | 15,285      | KJ784473.1       | Gan et al. (2014c)              |
| Triphysa     | phryne                         | 15,143      | KF906487.1       | Zhang et al. (2014a)            |
| Yoma         | sabina                         | 15,330      | KF990535.1       | Wu et al. (2014)                |
| Yptihma      | akragas                        | 15,227      | KF990533.1       | Wu et al. (2014)                |
| Cossidae     | echistis                       | 15,431      | KY318843.1       | Gong et al. (2014)              |
| Bombycidae   | bombyx mori                    | 15,643      | KM279431.1       | Zhang et al. (2014b)            |
| Hespiadae    | thitarodes renziensis          | 16,173      | NC_018094.1      | Cao et al. (2012)               |
nuclear genome, the presence of nuclear copies of mitochondrial (NUMT) DNA (which can lead to conflicts in assembly), and the poor signal-to-noise ratio caused by the high coverage of mitochondrial DNA (Hahn et al., 2013). However, a dedicated effort should be able to assemble the mitogenome from whole-genome sequencing reads. The mitogenome sequences are expected to be useful for phylogenetic studies, and they have been obtained for many species. The Papilionidae family has over 570 species worldwide (C.A. Bridges, 1988), while only 17 species have complete mitogenomes currently available in GenBank (included in Table 1, accessed on: 11/28/2014).

The insect mitogenome is a circular DNA of 14–19 kilobases (kb), containing 13 protein-coding genes (PCGs), 2 ribosomal-RNA-coding genes (rRNAs), 22 transfer-RNA-coding genes (tRNAs), and an A + T rich displacement loop (D-loop) control region (Cameron, 2014). Because of their maternal inheritance, compact structure, lack of genetic recombination, and relatively fast evolutionary rate, mitogenomes have been used widely in molecular phylogenetics and evolution studies (Cameron, 2014; Moritz et al., 1987). Here, we reconstruct and annotate the complete mitogenome of Papilio glaucus from next generation sequencing reads, and perform phylogenetic analyses of P. glaucus mitogenome with available complete mitogenomes of butterflies.

Materials and methods

Library preparation and Illumina sequencing

A male P. glaucus was caught and freshly frozen from Lake Ray Roberts State Park, Greenbelt Corridor along the Elm Fork of the Trinity River, 33.2536, − 97.0434, Denton County, Texas, USA (date 4-VIII-2013). The specimen will be deposited in the National Museum of Natural History, Smithsonian Institution, Washington, DC, USA (USNM). Detailed procedures and protocols for library preparation were described in Cong et al. (2015).

Briefly, we extracted genomic DNA from a piece of muscle dissected from the butterfly thorax using the ChargeSwitch gDNA mini tissue kit (Life Technologies, Grand Island, NY, USA). 250 bp and 500 bp paired-end libraries were made following the Illumina TruSeq DNA sample preparation guide using enzymes from NEBNext Modules (New England Biolabs, Ipswich, MA, USA). These libraries were sequenced at the genomics core facility in UT Southwestern Medical Center for 150 bp from both ends with a rapid run on Illumina HiSeq1500.

Sequence assembly

Sequencing reads were processed sequentially by MIRABAIT (Chevreux et al., 1999) to remove contamination from sequence adapters, by Fastq_quality_trimmer (http://hannonlab.cshl.edu/fastx_toolkit/) to trim low-quality regions at both ends and by QUAKE to correct errors (Cong et al., 2015; Kelley et al., 2010). From either the 250 bp or 500 bp library, we used mitochondrial baiting and iterative mapping (MITObim) v1.6 (Hahn et al., 2013) to assemble the mitogenome using two approaches: (1) using mitogenomes of Papilio maackii (KC433408.1), Papilio polytes (KM014701.1) and Papilio maraho (FJ810212.1) as references to guide the assembly; (2) using a short COI barcode sequence (a segment of about 600 bp from the mitochondrial gene cytochrome oxidase I) of P. glaucus (GU090087.1) as the starting seed. We used the default parameters for MITObim except for setting the – kbait to be 35 instead of 31.

The genome assemblies directly produced by the reference-guided mode in MITObim did not directly consider that the mitogenome should be a circular DNA and that the reads mapped to the N-terminus of the reference sequence could overlap with reads mapped to the C-terminus of the reference mitogenome. Therefore,

| Fragment location on genome | Primer | Primer sequence (5′ to 3′) |
|-----------------------------|--------|---------------------------|
| 14,387–15,306, 1–104        | D-loopF | GCAACTGCTGGCACAAAAT       |
|                             | D-loopR | CAACTTCAACATCCCCAACATCA   |
|                             | ND4F    | CTATCTCTACCCAAGATCCACC    |
|                             | ND4R    | TAGCTGTTCCTCTTTTATG       |
| 7428–8136                   |         |                           |
Fig. 1. Coverage of *Papilio glaucus* mitogenome by sequencing reads (250 bp and 500 bp) mapped to them by Bowtie2. (A). Coverage at each base position. (B). Histogram of coverage distribution. (C). Negative correlation between the coverage and A + T contents of 50 bp windows in the genome.
it produced sequences whose C-terminal segment (usually about several hundred base pairs) is a duplication of the N-terminal segment. We detected such duplicated regions by aligning the N-terminal half and C-terminal half with BLASTN and manually removed the redundant segments. We also adjusted the linear representation of the circular DNA by circular permutation so that the sequence starts with the coding gene for ND2, which is the convention for most Lepidoptera sequences deposited in the database.

We aligned the mitogenome sequences produced by different methods with MAFFT (Katoh and Standley, 2013) (Supplementary data). These assemblies mostly agree with each other, with most discrepancies located in the D-loop region. We derived our final mitogenome sequence from the alignment of these different assemblies by taking the dominant nucleotide or gap at each position.

**Annotation and analysis of the mitochondrial genome**

The mitogenome sequence was annotated using the MITOS web server (Bernt et al., 2013). We translated the sequences of PCGs to protein sequences using the genetic code for invertebrate mitogenomes. Secondary structures of tRNA genes were predicted using the same server.

**Assembly quality assessment**

We first checked if the assembly was well-supported by the sequencing reads by mapping the reads to the mitogenome using bowtie2 v2.2.3 (Langmead and Salzberg, 2012). The alignments were combined into one single SAM-format file, processed with SAMtools (Li et al., 2009) and visualized in Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Number of mapped reads (coverage by reads) at each position was calculated using bedtools v2.20.1 (Quinlan and Hall, 2010) and the histogram of the coverage was prepared in IBM SPSS Statistics v21 and Microsoft Excel 2010.

Second, we assessed the quality of our assembly by its consistency with other published *Papilio* sequences in the protein-, rRNA- and tRNA-coding regions. We aligned the rRNA- and tRNA-coding sequences directly and aligned translated sequences for PCGs to the corresponding proteins of other published *Papilio* mitogenomes. Alignments confirmed that our sequences are consistent with the majority of these published mitogenomes, and gaps are only in regions that are poorly conserved among other *Papilio* species.

Finally, we confirmed the assembly by comparing with Sanger sequencing results. We compared the mitogenome sequence with a reported 2291 bp partial mitogenome sequence of *P. glaucus* (accession: AF044013) (Caterino and Sperling, 1999) using BLAST. In addition, we sequenced the D-loop region and one arbitrarily selected coding region that partly covers ND4 and ND5. Amplified products were separated by 2% E-Gel® EX Agarose Gels (Life Technologies, Grand Island, NY, USA) and purified by Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, CA, USA). The purified DNA fragments were sent for Sanger sequencing in the sequencing core facility at UT Southwestern Medical Center. Sequencing results were manually confirmed by visualizing the traces in FinchTV v1.4 and then compared with the assembled mitogenome.

**Table 3** Composition and skewness of *Papilio glaucus* mitogenome regions.

| Nucleotides | Whole genome | PCGs | tRNAs | rRNAs | A + T rich region | Intergenic spacer region |
|-------------|--------------|------|-------|-------|-------------------|-------------------------|
| A%          | 39.96        | 39.52| 40.06 | 40.76 | 47.13             | 39.36                   |
| T%          | 40.46        | 39.50| 40.96 | 43.33 | 47.54             | 45.74                   |
| G%          | 7.62         | 8.33 | 8.32  | 4.90  | 1.84              | 4.26                    |
| C%          | 11.95        | 12.65| 10.67 | 11.00 | 3.48              | 10.64                   |
| A + T%      | 80.43        | 79.02| 81.01 | 84.10 | 94.67             | 85.11                   |
| G + C%      | 19.57        | 20.98| 18.99 | 15.90 | 5.33              | 14.89                   |
| AT-skew⁎    | -0.0062      | 0.0003| -0.0111| -0.0306| -0.0043            | -0.0750                 |
| GC-skew⁎⁎   | -0.2210      | -0.2061| -0.1241| -0.3832| -0.3077            | -0.4286                 |

* AT-skew = \( \frac{|A - T|}{A + T} \).

** GC-skew = \( \frac{|G - C|}{G + C} \) (Perna and Kocher, 1995).
Fig. 2. Map of genes in the *Papilio glaucus* mitochondrial genome. PCGs are colored in yellow, tRNA-coding genes are in purple, *rrnL* and *rrnS* are in green. Each gene is shown as an arrow indicating the transcription direction. The black line in the middle shows the coordinate of each gene in the mitogenome. The arrows on top of the line correspond to genes coded on the majority strand, and those below show genes on the minority strand.
Phylogenetic analysis

104 complete, non-redundant butterfly mitogenomes that are currently available were downloaded from GenBank (Table 1). Moth mitogenomes from Thitarodes renzhiensis (NC_018094.1), Bombyx mori (KM279431.1), and Egystia hippochaelous (KC831443.1) were also downloaded and used as outgroups. DNA sequences of the 37 protein- and RNA-coding genes were aligned by MAFFT. We manually checked the alignments of each gene, corrected sequences from some species with annotation errors based on consensus, replaced mitogenomes of poor quality (for example, sequences with frame-shift mutation that causes premature ending of proteins) with alternative mitogenome from the same species, and removed positions with large gaps and their surrounding regions with uncertain alignment.

The processed alignments were analyzed for phylogeny with Bayesian Inference and Maximum likelihood methods using MrBayes v3.2 (Huelsenbeck and Ronquist, 2001) and RaxML v8.1 (Stamatakis, 2006). We built trees with a different partitioning of the data sets (unpartitioned; partitioned by genes, PCG codon positions, and the exclusion of 3rd codon positions of PCGs). In addition, the translated protein sequences were aligned with MAFFT and analyzed with MrBayes. For analyses based on DNA alignments, the most suitable nucleotide substitution model (GTR + I + G) selected by jModelTest v2.1.7 (Darriba et al., 2012) was used, and for the protein-based analyses, we used a mixed model (Poisson, Jones, Dayhoff, Mtrev, Mtmam, Wag, Rtrev).

Table 4
Summary of the Papilio glaucus mitogenome.

| Gene   | Direction | From | To   | Size  | Intergenic nucleotides | Anticodon | Start codon | Stop codon |
|--------|-----------|------|------|-------|------------------------|-----------|-------------|------------|
| ND2    | F         | 1    | 1014 | 1014  | -2                     | -         | ATT         | TAA        |
| trnW   | F         | 1013 | 1077 | 65    | -8                     | TCA       | -           | -          |
| trnC   | R         | 1070 | 1131 | 62    | 3                      | GCA       | -           | -          |
| trnY   | R         | 1135 | 1199 | 65    | 2                      | GTA       | -           | -          |
| COX1   | F         | 1202 | 2732 | 1531  | 0                      | -         | CGA         | T          |
| trnL2  | F         | 2733 | 2800 | 68    | 0                      | TAA       | -           | -          |
| COX2   | F         | 2801 | 3482 | 682   | 0                      | -         | AGT         | T          |
| trnK   | F         | 3483 | 3552 | 70    | 0                      | TTT       | -           | -          |
| trnD   | F         | 3553 | 3619 | 67    | 0                      | GTC       | -           | -          |
| ATP8   | F         | 3620 | 3787 | 168   | -7                     | -         | ATT         | TAA        |
| ATP6   | F         | 3781 | 4458 | 678   | 7                      | -         | AGT         | TAA        |
| COX3   | F         | 4466 | 5254 | 789   | 2                      | -         | AGT         | TAA        |
| trnG   | F         | 5257 | 5322 | 66    | -3                     | TCC       | -           | -          |
| ND3    | F         | 5320 | 5676 | 357   | -2                     | -         | ATA         | TAG        |
| trnA   | F         | 5675 | 5738 | 64    | -1                     | TGG       | -           | -          |
| trnR   | F         | 5738 | 5801 | 64    | -1                     | TCG       | -           | -          |
| trnN   | F         | 5801 | 5866 | 66    | 0                      | GTT       | -           | -          |
| trnS1  | F         | 5867 | 5927 | 61    | 1                      | ACT       | -           | -          |
| trnE   | F         | 5929 | 5995 | 67    | -2                     | TTC       | -           | -          |
| trnF   | R         | 5994 | 6060 | 67    | 0                      | GAA       | -           | -          |
| ND5    | R         | 6061 | 7794 | 1734  | 0                      | -         | ATA         | TAA        |
| trnH   | R         | 7795 | 7860 | 66    | 0                      | GTG       | -           | -          |
| ND4    | R         | 7861 | 9202 | 1342  | -4                     | -         | ATA         | T          |
| ND4L   | R         | 9199 | 9489 | 291   | 6                      | -         | AGT         | TAA        |
| trnT   | F         | 9496 | 9559 | 64    | 0                      | TGT       | -           | -          |
| trnP   | R         | 9560 | 9623 | 64    | 2                      | TGG       | -           | -          |
| ND6    | F         | 9626 | 10,159 | 534  | 4                      | -         | ATT         | TAA        |
| CYTB   | F         | 10,164 | 11,318 | 1155  | 2                      | -         | ATA         | TAA        |
| trnS2  | F         | 11,321 | 11,385 | 65   | 16                     | TGA       | -           | -          |
| ND1    | R         | 11,402 | 12,340 | 939  | 1                      | -         | ATG         | TAG        |
| trnL1  | R         | 12,342 | 12,409 | 68   | 0                      | TAG       | -           | -          |
| rnl    | R         | 12,410 | 13,728 | 1319  | 0                      | -         | -           | -          |
| trnV   | R         | 13,729 | 13,791 | 63   | 0                      | TAC       | -           | -          |
| trnS   | R         | 13,792 | 14,572 | 781  | 0                      | -         | -           | -          |
| A + T rich region | 14,573 | 15,060 | 488   | 0                      | -         | -           | -          |
| trnM   | F         | 15,061 | 15,129 | 69   | 0                      | CAT       | -           | -          |
| trnI   | F         | 15,130 | 15,193 | 64   | -3                     | CAT       | -           | -          |
| trnQ   | R         | 15,191 | 15,258 | 68   | 48                     | TTG       | -           | -          |
Fig. 3. Secondary structure of 22 tRNA-coding genes of Papilio glaucus mitogenome predicted by the MITOS web server. The tRNAs are labeled by their corresponding amino acids in abbreviations.
Cprev, Vt, and Blosum) provided by the MrBayes program. The resulting phylogenetic trees were visualized in FigTree v1.4.2.

Results

Coverage of the mitogenome assembly by the reads

The coverage of the assembled *P. glaucus* mitogenome was high by the sequencing reads, with about 6000 fold mean coverage at base pair level and 500 fold minimal coverage (Fig. 1A and B). As shown in Fig. 1A, the regions with the lowest coverage were the beginning, the end and the D-loop regions. The low coverage at the beginning and the end of the mitogenome was primarily an artifact from limiting each read to map to only one most likely position: the circular DNA was represented as a linear sequence, and reads that should map partly to the beginning and partly to the end were restricted to map to either the beginning or the end.

The lower coverage in the D-loop might indicate potential errors in that region. However, independent Sanger sequencing of both the D-loop region (from the end of rrnS to the beginning of ND2) and another arbitrarily selected region (from the end of ND4 to the beginning of ND5), completely matched our assembly, indicating its high quality. In addition, the mitogenome sequence also agreed with the partial mitogenome (2291 bp) sequence of *P. glaucus* (accession: AF044013) in the database. Our sequence showed 0.3% sequence divergence from the previous sequence (only 6 out of 2291 positions are different), which likely corresponded to sequence variations between different individuals of the same species.

Instead, the low coverage of the D-loop region is probably related to its AT-rich composition, which tends to break during library preparation and is thus underrepresented in the sequencing libraries (Benjamini and Speed, 2012). Indeed, we observed that the percentage of A and T in a 50 bp window from the mitogenome is negatively correlated with the coverage for that region by the reads (Fig. 1C).

Base composition and genome structure

The *P. glaucus* mitogenome is a closed circular DNA of 15,306 bp. The nucleotide composition of the majority strand is A = 6117 (39.96%), T = 6193 (40.46%), G = 1167 (7.62%), and C = 1829 (11.95%), which is highly biased towards A and T. The majority strand has a negative AT-skew (−0.0062) and GC-skew (−0.2210) (Table 3), indicating a higher occurrence of T over A, C over G nucleotides on this strand. The *P. glaucus* mitogenome retains the typical insect mitogenome gene set, including 13 PCGs (ND1-6, COX1-3, ND4L, ATP8, ATP6, and CYTB), 22 tRNA genes (two for serine and leucine and one for each of the remaining amino acids), 2 ribosomal RNAs (rrnL and rrnS), and an A + T rich D-loop control region (Fig. 2 and Table 4).

Annotation of the mitogenome

The annotation of the mitogenome is illustrated in Fig. 2 and summarized in Table 4. Nine protein-coding genes (ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND6, and CYTB) are coded on the majority strand. COX1 uses start codon CGA, which is consistent with many other insect mitogenomes (Kim et al., 2009). A recent study using an expressed sequence tag from a Lepidopteran species confirmed the presence of COXI transcripts starting from CGA (Margam et al., 2011). Each of the rest of the genes starts with the typical ATN.

COX1, COX2 and ND4 use an incomplete stop codon T (Ojala et al., 1981), and a complete TAA codon will likely be formed during mRNA maturation (Boore, 1999; Ojala et al., 1981). The 13 PCGs have a total length of 11,214 bp (Table 4).

14 out of the 22 tRNA-coding genes are encoded on the majority strand. The tRNAs have a total length of 1443 bp, and their individual lengths range from 61 bp to 70 bp (Table 4). Secondary structures predicted by MITOS suggest that all tRNA genes adopt a typical cloverleaf structure except for trnS1 (Fig. 3). The dihydrouridine (DHU) arm of trnS1 does not form a stable stem-loop structure, which is very common in butterfly mitogenomes (Kim et al., 2014; Lu et al., 2013). The two rRNA genes, rrnL and rrnS, are located on the minority strand, and their lengths are 1319 bp and 781 bp, respectively (Table 4).

A 488 bp A + T rich region (A + T content: 94.7%) connects rrnS and trnM. This region contains an “ATAGA” motif located 19 bp downstream from rrnS and followed by 14 bp of poly-T stretch that is consistent with a gene regulation element commonly found in Lepidoptera (Lu et al., 2013; Salvato et al., 2008).
addition to this A + T rich region, there are 94 bp non-coding nucleotides that make up 12 intergenic spacer sequences, ranging from 1 bp to 48 bp in length. The longest 48 bp spacer is located between trnQ and ND2, and a 16 bp spacer is located between trnS2 and ND1 (Table 4). In addition, there are 33 bp of overlapping sequences at 10 locations. The longest 8 bp overlap is between trnW and trnC. There is a 7 bp overlap between ATP8 and ATP6 (Table 4), and this is a common feature for Lepidopteran mitogenomes (Lu et al., 2013).

Phylogenetic relationships

We phylogenetically analyzed 105 butterfly species from 6 families: Papilionidae, Hesperiidae, Pieridae, Lycaenidae, Riodinidae, and Nymphalidae, and used 3 species of moths as outgroups. Maximum Likelihood analysis of the DNA alignments (Fig. 4A) and Bayesian Inference of DNA (Fig. 4B and Supplemental S2) and protein sequences (Fig. 4C) correctly partitioned butterflies into 6 families and suggested the same tree topology at the family level: (Papilionidae + (Hesperiidae + (Pieridae + ((Riodinidae + Lycaenidae) + Nymphalidae))))..

Tree topologies between different methods were very similar. The positions of several species vary between trees obtained with different methods, such as Carterocephalus silvicola, Hebomoia glaucippe, and Ypthima akragas. However, all the trees consistently place P. glaucus, the only available mitogenome from the subgenus Perourus, as a sister of P. maraho, the only available mitogenome from the subgenus Agehana (Fig. 4 and Supplemental S2).

Discussion

The traditional method to obtain the mitogenome is through Sanger sequencing of a couple of overlapping segments. Here, we describe our protocol of assembling mitogenomes from Next Generation Sequencing reads for whole genome sequencing, and report the mitogenome of Eastern tiger swallowtail, P. glaucus. We used MITObim, a published tool designed for this task. MITObim has a reference-guided mode: it finds the conserved regions of a mitogenome using related reference species and extends these regions by baiting reads that overlap with the assembled regions until no gaps are left in between. Another mode of MITObim works without a reference mitogenome: it starts with a short COI sequence and extends by baiting reads with overlaps, till the N- and C-termini of the sequence can be mapped to the same reads, indicating that a circular mitogenome has been assembled (Hahn et al., 2013). Compared with the traditional PCR method, this method of mitogenome assembly does not require multiple primer designing and optimization, especially for species with limited knowledge available for primers’ design.

However, MITObim could make mistakes due to (1) ambiguity in mapping and aligning the reads, (2) presence of sequencing error, and (3) difficulty in distinguishing mitochondrial DNA reads from nuclear DNA reads, especially those from nuclear copies of mitochondrial DNA or low-complexity regions. Therefore, taking the consensus of different MITObim runs with different modes and references, and careful validation of the result are needed. We produced a highly accurate assembly by integrating assemblies made by different MITObim modes (using other mitogenomes as references or the COI barcode sequence as seed). Even for the D-loop region that contains multiple short repeats, sequences obtained from Sanger sequencing showed no differences from our assembly.

Our phylogenetic analysis yielded a detailed tree of butterflies with available complete mitogenomes. Traditionally, Hesperiidae were considered to be at the root of the phylogenetic tree for all butterflies due to their similarity in morphology to moths (Kristensen and Skalski, 1998). However, several recent studies with molecular evidence have suggested a different evolutionary history of butterflies: (Papilionidae + (Hesperiidae + (Pieridae + (Riodinidae + Lycaenidae) + Nymphalidae)))) (Heikkila et al., 2012; Kim et al., 2014; Regier et al., 2013). Our phylogenetic analysis contained many more taxa compared with previous analyses and supported the same result in placing Papilionidae at the base. The reason for this apparent

Fig. 4. Phylogeny of butterflies. (A), Phylogenetic tree obtained by RaxML with the data set of PCG and RNA genes, partitioned into 13 PCGs, 1 tRNA, and 2 rRNA groups. (B), Phylogenetic tree obtained by MrBayes based on PCG and RNA genes, partitioned into 13 PCGs, 1 tRNA, and 2 rRNA groups. (C), Phylogenetic tree obtained by MrBayes based on 13 protein sequences, unpartitioned. Number at each node shows confidence of that group by bootstrap in (A), or posterior probability in (B) and (C). Thitarodes renzhiensis (NC_018094.1), Bombyx mori (KM279431.1), and Eugystia hippoclaoecus (KCR31443.1) were used as outgroup.
contradiction between the traditional morphology and recent molecular-based phylogeny is still poorly understood and requires future analysis.

It is notable that *P. glaucus*, a swallowtail native of eastern North America, is found to be confidently grouped with *P. maraho* based on our phylogeny analysis. *P. maraho* is a threatened swallowtail endemic to Taiwan in Asia (Bailie and Groombridge, 1996). Morphology, behavior and phylogenetic studies based on COI barcode suggested that *P. maraho* is very close to *Papilio elwesi* (Igarashi, 1979; Lu et al., 2009). Both *P. elwesi* and *P. maraho* are frequently attributed to the subgenus *Agehana*, which in some studies is included in subgenus *Chilasa* (Hancock, 1983; Zakharov et al., 2004). Despite a wide geographic separation, *Agehana* and *Chilasa* native to Asia are likely to be the closest relatives to the subgenus *Pterourus* (Zakharov et al., 2004), which is native to America. We speculate that these butterflies might have migrated between Asia (*Agehana*) and North America (*Pterourus*). Swallowtail butterflies have dispersed between continents (Condamine et al., 2012, 2013). A recent report had found that *Polyommatus* blue butterflies traveled from Asia to North America via the Bering Strait, ultimately migrating to South America (Vila et al., 2011). Several other reports found that other butterfly species, animals, and plants followed the same route to the New World (Donoghue and Smith, 2004; Enghoff, 1995; Mullen, 2006; Peña et al., 2010).

Although our phylogenetic analyses produced similar phylogenetic trees, despite being performed with different methods, on different subsets of data (without or without RNA-coding sequences, DNA and translated proteins), and with different partition schemes, minor variations in positions of some taxa were observed. These inconsistencies might be caused by the available mitogenomes not carrying sufficient information to deduce accurate phylogeny, available software not being fully adequate for the task, or our position selection strategy requiring improvement. The current sample of taxa did not provide a fully resolved, consistent and accurate phylogeny. Additional analyses, mitogenomes, and possibly information from nuclear genomes are needed to address the question of the applicability of mitogenomes to infer a more accurate phylogenetic tree of butterflies.

Supplementary data to this article can be found online at [http://dx.doi.org/10.1016/j.mgene.2015.05.002](http://dx.doi.org/10.1016/j.mgene.2015.05.002).

**Acknowledgment**

This work was supported by the National Institutes of Health (GM094575 to NVG) and the Welch Foundation (I-1505 to NVG). Qian Cong is a Howard Hughes Medical Institute International Student Research fellow. We acknowledge Texas Parks and Wildlife Department (Natural Resources Program Director David H. Riskind) for the permit #08-02Rev that makes our research in Texas State Parks possible. And we thank Lisa N. Kinch and Dustin Schaeffer for critical suggestions and corrections to the manuscript.

**References**

Baillie, J., Groombridge, B., 1996. 1996 IUCN Red List of Threatened Animals. IUCN, Cland, Switzerland and Cambridge, UK.

Benjaminii, Y., Speed, T.P., 2012. Summarizing and correcting the GC content bias in high-throughput sequencing. Nucleic Acids Res. 40, e72.

Bernet, M., Donath, A., Juhling, F., Externbrink, F., Florentz, C., Fritzsch, G., Putz, J., Mittendorf, M., Studier, P.F., 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69, 313–319.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.

C.A. Bridges, 1988. Catalogue of Papilionidae & Pieridae (Lepidoptera: Rhopalocera). C.A. Bridges, Urbana, IL.

Cally, S., Luuillier, E., Iribar, A., Garzon-Orduna, I., Coissac, E., Murienne, J., 2014. Shotgun assembly of the complete mitochondrial genome of the neotropical cracker butterfly Hamadryas epinome. Mitochondrial DNA 1–3.

Cameron, S.L., 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. 59, 95–117.

Cao, Y.Q., Ma, C., Chen, J.Y., Yang, D.R., 2012. The complete mitochondrial genomes of two ghost moths, *Thitarodes renzhiensis* and *T. yunnanensis*: the ancestral gene arrangement in Lepidoptera. BMC Genomics 13, 276.

Cao, T.-W., Wang, J.-P., Xiao, S.-B., Zhang, M., Guo, Y.-P., Ma, E.-B., 2013. Analysis of complete mitochondrial genome of *Timelaea masura* (Lepidoptera: Nympalidae). Zool. Syst. 38, 468–475.

Caterino, M.S., Sperling, F.A., 1999. Papilio phylogeny based on mitochondrial cytochrome oxidase I and II genes. Mol. Phylogenet. Evol. 11, 122–137.

Chen, M., Tian, L.L., Shi, Q.H., Cao, T.W., Hao, J.S., 2012. Complete mitogenome of the Lesser Purple Emperor *Apatura ilia* (Lepidoptera: Nympalidae: Apaturinae) and comparison with other nympalid butterflies. Dongwu Xuebao 33, 191–201.

Chen, Y.-h., Huang, D.-y., Wang, Y.-l., Zhu, C.-d., Hao, J.-s., 2014a. The complete mitochondrial genome of the endangered Apollo butterfly, *Parnassius apollo* (Lepidoptera: Papilionidae) and its comparison to other Papilionidae species. J. Asia Pac. Entomol. 17, 663–671.

Chen, Y., Can, S., Shao, L., Cheng, C., Hao, J., 2014b. The complete mitochondrial genome of the Painted lady (Lepidoptera: Papilionidae: Papilioninae). Mitochondrial DNA [http://dx.doi.org/10.3109/19401736.2014.905843](http://dx.doi.org/10.3109/19401736.2014.905843).
Chen, Y., Gan, S., Wang, Y., Wang, Y., Zuo, N., Hao, J., 2014c. The complete mitochondrial genome of the Byasa alcinous (Lepidoptera: Papilionidae: Papilioninae). Mitochondrial DNA 1–2.

Chervaux, B., Wetter, T., Suhai, S., 1999. Genome sequence assembly using trace signals and additional sequence information. Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB).

Clarke, C.A., Sheppard, P.M., 1972. The genetics of the mimetic butterfly Papilio polytes L. Philos. Trans. R. Soc. B 263, 431–458.

Condaminne, F.L., Sperling, F.A.H., Wahlberg, N., Rasplus, J.-Y., Kergoat, G.J., 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. Ecol. Lett. 15, 267–277.

Condaminne, F.L., Sperling, F.A.H., Kergoat, G.J., 2013. Global biogeographical pattern of swallowtail diversification demonstrates alternative colonization routes in the Northern and Southern hemispheres. J. Biogeogr. 40, 9–23.

Cong, Q., Borek, D., Otwinowski, Z., Grishin, Nick V., 2015. Tiger swallowtail genome reveals mechanisms for speciation and caterpillar chemical defense. Cell Rep. 10.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Dong, Y., Zhu, L.X., Wu, Y.F., Wu, X.B., 2013. The complete mitochondrial genome of the Alpine black swallowtail, Papilio maackii (Insecta: Lepidoptera: Papilionidae). Mitochondrial DNA 24, 635–641.

Dong, Y., Zhu, L.X., Ding, M.J., Wang, J.J., Luo, L.G., Liu, Y., Ou, Y.Y., 2014. Complete mitochondrial genome of Papilio syfanus (Lepidoptera: Papilionidae). Mitochondrial DNA http://dx.doi.org/10.3109/19401736.2014.898278.

Donoghue, M.J., Smith, S.A., 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. Philos. Trans. R. Soc. B 359, 1633–1644.

Enghoff, H., 1995. Historical biogeography of the holarctic: area relationships, ancestral areas, and dispersal of non-marine animals. Cladistics 11, 223–263.

Engstorfa, F., Sangkot, U., Chatgeat, W., Satasook, C., 2014. Molecular evolution of the odorant and gustatory receptor genes in lepidopteran insects: implications for their adaptation and speciation. J. Mol. Evol. 79, 21–39.

Gan, Sun, X.Y., Gai, Y.H., Hao, J.S., 2014a. The complete mitochondrial genome of Danaus chrysippus (Lepidoptera: Nymphalidae: Danainae). Mitochondrial DNA http://dx.doi.org/10.3109/19401736.2013.855909.

Gan, S., Chen, Y., Zuo, N., Zhang, W., Hao, J., 2014b. The complete mitochondrial genome of Ideopsis similis (Lepidoptera: Nymphalidae: Danainae). Mitochondrial DNA http://dx.doi.org/10.3109/19401736.2014.903842.

Gan, S., Chen, Y., Zuo, N., Xia, C., Hao, J., 2014c. The complete mitochondrial genome of Tirumala limnaius (Lepidoptera: Nymphalidae: Danainae). Mitochondrial DNA 1–3.

Gong, Y.J., Wu, Q.L., Wei, S.J., 2014. The complete mitochondrial genome of the complete nucleotide sequence of the mitochondrial genome of the orange tip butterfly Eogystia hippophaecolus (Lepidoptera: Pieridae). Acta Biochim. Biophys. Sin. 41, 446–450.

Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads–a baiting and iterative mapping approach. Nucleic Acids Res. 41, e129.

Hancock, D.L., 1983. Classification of the Papilionidae (Lepidoptera): a phylogenetic approach. Smithsonian 2, 1–48.

Hao, J., Sun, Q., Zhao, H., Sun, X., Gai, Y., Yang, Q., 2012. The complete mitochondrial genome of Chlosyne phlaeas (Lepidoptera: Papilionidae: Papilioninae): Pyrginae and its phylogenetic implication. Comp. Funct. Genomics 2012, 328049.

Hao, J.J., Wang, Y.L., Sun, X.Y., Zhang, L.L., Hao, J.S., Yang, Q., 2013a. The complete mitochondrial genome of Helbomia glaucippe (Lepidoptera: Pieridae). Mitochondrial DNA 24, 608–670.

Hao, J., Sun, M., Shi, Q., Sun, X., Shao, S., Wang, Y., Yang, Q., 2013b. Complete mitogenomes of Euploea mulciber (Nymphalidae: Danainae) and Libythea celis (Nymphalidae: Libytheinae) and their phylogenetic implications. SSN Genomics 2013, 14.

Hao, J.J., Hao, J.S., Sun, X.Y., Zhang, L.L., Yang, Q., 2014. The complete mitochondrial genomes of the Fenton’s wood white, Leptidea morsei, and the lemon emigrant, Catopsilia pomona. J. Insect Sci. 14, 130.

Heikila, M., Kaila, L., Mutanen, M., Pena, C., Wahlberg, N., 2012. Cretaceous origin and repeated tertiary diversification of the redefined butterflies. Proc. Biol. Sci. R. Soc. 279, 1093–1099.

Heliconius Genome, C., 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487, 94–98.

Hines, H.M., Papa, R., Ruiz, M., Papanicolaou, A., Wang, C., Nijhout, H.F., McMillan, W.O., Reed, R.D., 2012. Transcriptome analysis reveals novel patterning and pigmentation genes underlying Heliconius butterfly wing pattern variation. BMC Genomics 13, 288.

Hong, C., Jiang, S., Yu, M., Yang, Y., Li, F., Xue, F., Wei, Z., 2009. The complete nucleotide sequence of the mitochondrial genome of the cabbage butterfly, Argea melitea (Lepidoptera: Pieridae). Acta Biochim. Biophys. Sin. 41, 446–450. http://hannonlab.cshl.edu/fastx_toolkit/.

Hu, J., Zhang, D., Hao, J., Huang, D., Cameron, S., Zhu, C., 2010. The complete mitochondrial genome of the yellow coater, Azarae isoria (Lepidoptera: Nymphalidae: Helicoinae: Acrainini): sequence, gene organization and a unique tRNA translocation event. Mol. Biol. Rep. 37, 3431–3438.

Huang, Z.H., Dai, P.F., Zhao, C.G., 2014a. The complete mitochondrial genome of Heliconius pachinus (Insecta: Lepidoptera: Nymphalidae). Mitochondrial DNA 1–2.

Huang, D., Hao, J., Zhang, W., Su, T., Wang, Y., Xu, X., 2014b. The complete mitochondrial genome of Melanargia aestiva (Lepidoptera: Nymphalidae: Satyrinae). Mitochondrial DNA 1–3.

Huenenbeck, J.P., Ronquist, F., 2001. MIRAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Igarashi, S., 1979. Papilionidae and Their Early Stages. Kodansha, Tokyo.

Igarashi, S., 1979. Papilionidae and Their Early Stages. Kodansha, Tokyo.

Ji, L.-W., Hao, J.-S., Wang, Y., Zuo, N., Zhang, W., Huang, D.-Y., Zhao, J.-L., Zhu, C.-D., 2012. The complete mitochondrial genome of the dragon swallowtail, Sericinus montela (Lepidoptera: Papilionidae) and its phylogenetic implication. Acta Entomol. Sin. 55, 91–100.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kelley, D.R., Schatz, M.C., Salzberg, S.L., 2010. Quake: quality-aware detection and correction of sequencing errors. Genome Biol. 11, R116.

Kim, M.J., Kim, I., 2014. Complete mitochondrial genome of the Mormon metalmark butterfly, Apodemia mormo (Lepidoptera: Riodinidae). Mitochondrial DNA 1–3.

Kim, I., Lee, E.M., Seol, K.Y., Yun, E.Y., Lee, Y.B., Huang, J.S., Jin, B.R., 2006. The mitochondrial genome of the Korean hairstreak, Corema raphelis (Lepidoptera: Lycaenidae). Insect Mol. Biol. 15, 217–225.
Stamatakis, A., 2006. RaxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Sun, X., Shao, L., Peng, C., Hao, J., Yang, Q., 2014. The complete mitochondrial genome of Eurema hecabe (Lepidoptera: Pieridae: Coliadinae). Mitochondrial DNA http://dx.doi.org/10.3109/19401736.2013.855751.

Surridge, A.K., Lopez-Gomollon, S., Moxon, S., Maroja, L.S., Rathjen, T., Nadeau, N.J., Dalmay, T., Jiggins, C.D., 2011. Characterization and expression of microRNAs in developing wings of the neotropical butterfly Heliconius melpomene. BMC Genomics 12, 62.

Teixeira da Costa, L.F., 2014. The complete mitochondrial genome of Parage aegeria (Insecta: Lepidoptera: Papilionidae). Mitochondrial DNA http://dx.doi.org/10.3109/19401736.2014.905853.

Tian, L.L., Sun, X.Y., Chen, M., Gai, Y.H., Hao, J.S., Yang, Q., 2012. Complete mitochondrial genome of the five-dot sergeant Parathyma sulphita (Nymphalidae: Limenitidinae) and its phylogenetic implications. Dongwuxue Yanjiu 33, 133–143.

Vila, R., Bell, C.D., Macniven, R., Goldman-Huertas, B., Ree, R.H., Marshall, C.R., Bálint, Z., Johnson, K., Benyamini, D., Pierce, N.E., 2011. Phylogenomy and palaeoecology of Polyommatus blue butterflies show Beringia was a climate-regulated gateway to the New World. Proc. R. Soc. Lond. B Biol. Sci. 278 (1719), 2737–2744.

Wang, X.C., Sun, X.Y., Sun, Q.Q., Zhang, D.X., Hu, J., Yang, Q., Hao, J.S., 2014a. Complete mitochondrial genome of the laced fritillary Argyres hyperius (Lepidoptera: Nymphalidae). Dongwuxue Yanjiu 32, 465–475.

Wang, J.-P., Nie, X.-P., Cao, T.-W., Zhang, M., Li, T., Zhang, X., Guo, Y.-P., Ma, E.-B., Zhang, X.-N., 2012. Characterization of complete mitochondrial genome of Apatura metis (Lepidoptera: Nymphalidae). BMC Genomics 15, 468.

Wang, X.C., Sun, X.Y., Sun, Q.Q., Zhang, D.X., Hu, J., Yang, Q., Hao, J.S., 2014b. Mitochondrial DNA 24, 475–484.

Wang, J.-P., Nie, X.-P., Cao, T.-W., Zhang, M., Li, T., Zhang, X., Guo, Y.-P., Ma, E.-B., Zhang, X.-N., 2012. The complete mitochondrial genome of Sasakia charonda coreana (Lepidoptera: Hesperiidae). Mitochondrial DNA 526, 277–283.

Wang, K., Hao, J., Zhao, H., 2013a. Characterization of complete mitochondrial genome of the skipper butterfly, Celaenorrhinus maculosus (Lepidoptera: Hesperioidea). Mitochondrial DNA http://dx.doi.org/10.3109/19401736.2013.840610.

Wang, J.P., Cao, T.W., Xuan, S.B., Wang, H., Zhang, M., Ma, E., 2013b. The complete mitogenome of the common Mormon Papilio polytes (Insecta: Lepidoptera: Papilionoidea). Mitochondrial DNA 1–2.

Wang, Y., Chen, Y., Xia, C., Xia, X., Chen, X., Hao, J., 2014b. The complete mitochondrial genome of Parmassi imperator (Lepidoptera: Papilionidae: Parmassiinae). Mitochondrial DNA 1–2.

Wang, A.R., Jeong, H.C., Han, Y.S., Kim, I., 2014c. The complete mitochondrial genome of the mountainous duskywing, Parage aegeria (Insecta: Lepidoptera: Coliadinae). Mitochondrial DNA 25, 93–94.

Wu, L.W., Lin, L.H., Lees, D.C., Hsu, Y.F., 2014. Mitogenomic sequences effectively recover relationships within brush-footed butterflies (Lepidoptera: Nymphalidae). BMC Genomics 15, 468.

Xia, J., Hu, J., Zhu, G.-P., Zhu, C.-D., Hao, J.-S., 2011. Sequencing and analysis of the complete mitochondrial genome of Calinaga davidis Cherthar (Lepidoptera: Nymphalidae). Acta Entomol. Sin. 54, 555–565.

Zakharov, E.V., Caterino, M.S., Sperling, F.A., 2004. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus Papilio (Lepidoptera: Papilionidae). Syst. Biol. 53, 193–215.

Zhan, S., Zhang, W., Nie, X., Cao, T., Wang, J., Li, T., Zhang, X., Guo, Y., Ma, E., Zhong, Y., 2012. The complete mitochondrial genome of the skipper butterfly Apatura melits (Lepidoptera: Nymphalidae). Mol. Biol. Rep. 39, 6529–6536.

Zhang, W., Zhan, S., Zhang, W., Nie, X., Nie, X.-P., Wang, J.-P., Wang, C.-Z., Wang, J., 2014. Genomic-wide characterization of adaptation and speciation in tiger swallowtail butterflies using de novo transcriptome assemblies. Genome Biol. Evol. 5, 1233–1245.

Zhang, W., Zhan, S., Zhang, W., Nie, X., Nie, X.-P., Wang, J.-P., Wang, C.-Z., Wang, J., 2014. The complete mitochondrial genome of Cupido argiades (Lepidoptera: Lycanidae). Mitochondrial DNA 24, 475–477.

Zhang, L., Wang, J., Zhao, H., Sun, Y.Y., Xuan, S.B., 2013b. Characterization of complete mitochondrial genome of the bright sunbeam Curetis fulvus and the small copper Lycaena phlaeas (Lepidoptera: Lycaenidae) and their phylogenetic implications. Genet. Mol. Res. 12, 4434–4445.

Zhang, W., Can, S., Zuo, N., Chen, C., Wang, Y., Hao, J., 2014a. Complete mitochondrial genome of Triphysa phyma (Lepidoptera: Nymphalidae: Satyrinae). Mitochondrial DNA http://dx.doi.org/10.3109/19401736.2014.900673.

Zhang, H., Li, F., Zuo, N., Meng, Z., 2014b. The complete mitochondrial genome of Bombyx mori strain Baiyun (Lepidoptera: Bombycidae). Mitochondrial DNA 1–2.