Mitochondrial Genomes of Hestina persimilis and Hestinalis nama (Lepidoptera, Nymphalidae): Genome Description and Phylogenetic Implications

Yupeng Wu 1,2,*, Hui Fang 1, Jiping Wen 2,3, Juping Wang 2, Tianwen Cao 2,4* and Bo He 4

Abstract: In this study, the complete mitochondrial genomes (mitogenomes) of Hestina persimilis and Hestinalis nama were acquired. The mitogenomes of H. persimilis and H. nama are 15,252 bp and 15,208 bp in length, respectively. These two mitogenomes have the typical composition, including 37 genes and a control region. The start codons of the protein-coding genes (PCGs) in the two mitogenomes are the typical codon pattern ATN, except CGA in the cox1 gene. Twenty-one tRNA genes show a typical clover leaf structure, however, trnS(AGN) lacks the dihydrouridine (DHU) stem. The secondary structures of rrnL and rrnS of two species were predicted, and there are several new stem loops near the 5′ of rrnL secondary structure. Based on comparative genomic analysis, four similar conservative structures can be found in the control regions of these two mitogenomes. The phylogenetic analyses were performed on mitogenomes of Nymphalidae. The phylogenetic trees show that the relationships among Nymphalidae are generally identical to previous studies, as follows: Libytheinae|Danainae + ((Calinaginae + Satyrinae) + Danainae|Libytheinae + ((Heliconiinae + Limenitidinae) + (Nymphalinae + (Apaturinae + Biblidinae))))). Hestina nama is apart from Hestina, and closely related to Apatura, forming monophyletic clade.

Keywords: lepidoptera; Nymphalidae; mitochondrial genome; phylogeny

1. Introduction

Hestina persimilis and Hestinalis nama belong to the lepidopteran family Nymphalidae, Apaturinae and mainly distribute in the Palaeartic and Oriental region. Their adults inhabit mountainous broad-leaved forests and present the habit of sipping tree juice and water in wetlands. The larvae were reported as a kind of agriculture pest of the host plants, distributed in China. Hestina assimilis, and closely related to Apatura, forming monophyletic clade.
external genitalia and wing veins. Morphologically, including genitalic structure, *Hestina* and *Hestinalis* are easily separable [4]. In recent studies, *Hestinalis* was treated as a distinct genus [5]. Although most modern literature chooses to separate them, some literature, Wu and Hsu still treats them as one [6]. In this paper, phylogenetic analysis shows that *Hestinalis nama* is apart from *Hestina*. Therefore, we also separate them apart.

Mitogenome fragments have been extensively used in phylogenetic analysis for butterflies and moths, particularly for the *cox1* gene which was primarily used as a DNA barcoding for animals [7–11]. In the BOLD system [12], lepidopteran insects consist of the largest amount of data being sequenced. However, the phylogenetic relationships at different taxonomic levels are still controversial [13–15]. It has been proposed that mitochondrial genomes might provide more genetic information than a single gene fragment [16–18]. Therefore, sequencing more mitogenomes might improve our understanding of evolution and phylogeny at different taxonomic levels in Lepidoptera. Furthermore, the mitogenome has been widely used in the areas of population genetic structure, gene drift and phylogenetics, because of its characteristics of maternal inheritance, small genome size (15–20 kb in length) and rapid rate of evolution [19,20]. To date, only one complete mitogenome (*H. assimilis*) has been sequenced from the genus *Hestina*; other species and *Hestinalis nama* were all not sequenced, which is quite limited and will restrict our understanding of evolution in Nymphalidae at the genomic level. In this study, the mitogenomes of *H. persimilis* and *H. nama* were obtained, with the aim of: (1) providing a comparative analysis of Apaturinae mitogenomes, including nucleotide composition, codon usage, gene arrangement, prediction of tRNA and rRNA secondary structures and novel features of the control region, and (2) reconstructing the phylogenetic relationships among subfamilies in Nymphalidae based on mitogenomes data.

2. Materials and Methods

2.1. Sampling and DNA Sequencing

Specimens of *H. persimilis* and *H. nama* were collected from the Sichuan and Yunnan Provinces of China in 2010. Specimens were being made, followed by morphological identification. One side of the hindfoot for each sample was preserved in absolute ethanol and stored in −20 °C freezer in College of Plant Protection, Shanxi Agricultural University, Taiyuan, China.

The DNA extraction kit and primers [21] (Table S1) were produced by Shanghai Major Biomedical Technology Co., Ltd. (Shanghai, China). The reaction systems of PCR amplifications were 25 μL, including upstream and downstream primers 0.5 μL, respectively, PCR Master Mix 12.5 μL, DNA template 3 μL, and ddH₂O 8.5 μL. The amplification reaction conditions were as follows: initial denaturation at 94 °C for 2 min; 35 cycles of denaturation at 94 °C for 1 min, annealing at 53 °C for 45 s, extension at 72 °C for 1 min, and a final extension step at 72 °C for 4 min. PCR products were detected by 1% agarose gel electrophoresis. All the gene fragments were sent to Shanghai Major Biomedical Technology for sequencing.

2.2. Annotation and Analysis of Mitochondrial DNA

The original sequence fragments were assembled with SeqMan (Steve ShearDown, 1998–2001 version reserved by DNASTAR Inc., Madison, WI, USA) to get a complete mitogenome. The secondary structure of tRNA genes were determined by tRNAscan-SE Search Server (http://lowelab.ucsc.edu/tRNAscan-SE/; accessed on 28 June 2021) [22]. Putative tRNA genes, including *trnH* and *trnS1*(AGN), which could not be found by tRNAscanSE, were confirmed by comparison with the homologous genes of other Apaturinae species. PCGs and rRNA genes were identified by the MITOS webserver with invertebrate genetic code [23]. The nucleotide composition and codon usage of PCGs were calculated with MEGA-X [24]. Determination of tandem repeat sequences in control regions were performed using the Tandem Repeats Finder online software (http://tandem.bu.edu/trf/trf.html; ac-
cessed on 10 May 2020) [25]. The mitogenomes of *H. persimilis* and *H. nama* were uploaded to GenBank, with the accession numbers of MT110153 and MT110154.

2.3. Phylogenetic Analysis

Phylogenetic analysis was performed on the dataset of 13 PCGs from 54 complete or nearly complete mitogenomes of Nymphalidae, with two Papilionidae species selected as outgroups (Table S2). All assembled PCGs of 56 mitogenomes were aligned through MEGA-X. The optimal partition tactics and substitution models were selected by PartitionFinder v2 (Tables S3 and S4) [26–28]. The maximum likelihood (ML) and Bayesian inference (BI) analyses were conducted through the online CIPRES Science Gateway [29]. The ML analysis was performed with RAxML-HPC2 on XSEDE [30], with GTRGAMMA model applied to all partitions. Bootstrap values were estimated with 1000 replicates. The BI analyses were carried out through two independent Markov chain Monte Carlo (MCMC) chains, which were set for 1,000,000 generations, with sampling per 1000 generations.

3. Results and Discussion

3.1. Mitogenomes Organization

The complete mitogenomes of *H. persimilis* and *H. nama* are 15,252 and 15,208 bp. They share the consistent gene organization, order and arrangement with most of other lepidopterans, including 13 PCGs, 22 tRNAs and 2 rRNAs (rrnL and rrnS) (Figures 1 and 2). The mitogenome is circular with two strands. The heavy strand (H-strand) encodes most of the genes (9 PCGs and 14 tRNAs), while the light strand (L-strand) contains the remaining reverse complementary genes (four PCGs, eight tRNAs and two rRNAs), as shown in Tables 1 and 2. In addition, the nucleotide composition of the two species are both AT-biased, similar to other lepidopterans. The AT contents of the mitogenomes of *H. persimilis* and *H. nama* are 80.9 and 79.2%, respectively (Tables 3 and 4). The obvious AT-biased (Table S5) is generally believed to be related to the evolution of mitochondrial origin [31].

3.2. Protein Coding Genes and Codon Usage

Orthologs from the two *Hestina* mitogenomes present similar start and stop codons. Most PCGs start with the typical initial codon ATN, but *cox1* initiates with CGA. In particular, the putative start codon CGA in the *cox1* gene is a common feature of most sequenced lepidopterans, but a few species start with codon ATG, ATT, ATA or TTG. While most PCGs end with the stop codon TAA or TAG, truncated codon T is also detected in *cox2* and *nad4*. It has been proposed that truncated stop codons can be completed by polyadenylation, which was also found in other insectan mitogenomes [32].

Relative synonymous codon usage (RSCU) can directly reflect the preference of codon usage (Table S6). The total number of codons of 13 PCGs are 3703 for *H. persimilis* and 3709 for *H. nama*. The RSCU of twelve Apaturinae species show the same codon preference pattern (Figure 3). The mainly used codon families are *Leu1* (CUN), Ile, Phe and Met. There are at least 75 codons (CDs) per thousand CDs in each of them (Figure 4), among which *Leu1* has the highest utilization rate. Relative synonymous codon usage (RSCU) of Apaturinae show that degeneration codons are biased to use more A/T than G/C. The six most prevalent codons in Apaturinae, including AAU (I), AUA (M), AUU (F), UUA (L) and UAU (Y), are all composed of A and/or T. Conversely, some GC-rich codons are seldom utilized in the Apaturinae species. For example, the codon UCG, CCG are not used in *H. persimilis*, while CUG, GUC and CCG are absent in *Sasakia charonda*. This phenomenon is common among lepidopterans [33,34], which indicates that the GC content of genes is closely related to codon preference [35,36].
Figure 1. Linear map of the mitogenome of *Hestina persimilis*. The J-strand is located on the linear map, and the N-strand is under the linear map.
Figure 2. Linear map of the mitogenome of *Hestinalis nama*. The J-strand is located on the linear map, and the N-strand is under the linear map.

Table 1. Annotation of *Hestina persimilis* mitogenome.

| Gene   | Direction | Location | Size | Anticodon     | Start Codon | Stop Codon | Intergenic Nucleotides |
|--------|-----------|----------|------|---------------|-------------|------------|------------------------|
| trnM   | F         | 1–68     | 68   | CAT 32–34     |             |            | 0                      |
| trnI   | F         | 69–134   | 66   | GAT 98–100    |             |            |                        |
| trnQ   | R         | 132–200  | 69   | TTG 159–161   |             |            |                        |
| nad2   | F         | 292–1305 | 1014 |               | ATT         | TAA 91     |                        |
| trnW   | F         | 1304–1371| 68   | TCA 1335–1337 |             |            | −2                     |
| trnC   | R         | 1364–1427| 64   | GCA 1397–1399 |             |            | −8                     |
| trnY   | R         | 1428–1492| 65   | GTA 1359–1461 |             |            | 0                      |
| cox1   | F         | 1498–3033| 1536 |               | CGA         | TAA 5      |                        |
| trnL2  | F         | 3029–3095| 67   | TAA 3059–3061 |             |            | −5                     |
| (UUR)  |           |          |      |               |             |            |                        |
| cox2   | F         | 3096–3774| 679  |               | ATG         | T          | 0                      |
Table 1. Cont.

| Gene   | Direction | Location | Size | Anticodon | Start Codon | Stop Codon | Intergenic Nucleotides |
|--------|-----------|----------|------|-----------|-------------|------------|------------------------|
| trnK   | F         | 3772–3842| 71   | CTT       | 3802–3804   | −3         | 6                      |
| trnD   | F         | 3842–3907| 66   | GTC       | 3872–3874   | −1         | 6                      |
| atp8   | F         | 3948–4009| 62   | ATC       | TAA         | −7         | 6                      |
| atp6   | F         | 4063–4737| 675  | ATG       | TAA         | −1         | 6                      |
| cox3   | F         | 4737–5525| 789  | TGC       | 5976–5978   | 2          | 6                      |
| trnG   | F         | 5528–5594| 67   | TCC       | 5558–5560   | 2          | 6                      |
| nad3   | F         | 5595–5948| 354  | ATT       | TAG         | 0          | 6                      |
| trnA   | F         | 5947–6014| 68   | TGC       | 5976–5978   | 2          | 6                      |
| trnR   | F         | 6014–6077| 64   | TGC       | 6040–6042   | 1          | 6                      |
| trnN   | F         | 6090–6155| 66   | TGC       | 6121–6123   | 2          | 6                      |
| trnS1  | (AGN)     | 6154–6213| 60   | GCT       | 6171–6173   | −2         | 6                      |
| trnE   | F         | 6216–6280| 65   | TTC       | 6245–6247   | 2          | 6                      |
| trnF   | R         | 6281–6342| 69   | GAA       | 6310–6312   | −2         | 6                      |
| nad5   | R         | 6317–8077| 1761 | ATT       | TAA         | −26        | 6                      |
| trnH   | R         | 8075–8141| 67   | GTC       | 8109–8111   | −3         | 6                      |
| nad4   | F         | 8142–9480| 1339 | ATG       | TAA         | −8         | 6                      |
| nad4L  | R         | 9482–9772| 291  | ATT       | TAA         | 1          | 6                      |
| trnT   | F         | 9780–9844| 65   | TGT       | 9811–9813   | 7          | 6                      |
| trnP   | R         | 9845–9908| 64   | TGG       | 9827–9879   | 0          | 6                      |
| nad6   | R         | 9911–10438| 528 | ATT       | TAA         | 2          | 6                      |
| cob    | F         | 10,442–11,593| 1152 | GCT       | 11,617–11,623| 3          | 6                      |
| trnS2  | (UCN)     | 11,594–12,626| 942 | ATG       | TAA         | −2         | 6                      |
| trn11  | (CUN)     | 12,627–12,702| 75  | TAG       | 12,671–12,673| 21         | 6                      |
| trnL1  | (CUN)     | 12,703–14,036| 1334 | GCT       | 14,069–14,071| 0          | 6                      |
| trnL2  | (UCN)     | 14,040–14,876| 776 | GCT       | 14,839–14,841| 0          | 6                      |
| Control region | 14,877–15,252 | 376 | | | | 0 | 6 |

Table 2. Annotation of *Hestinalis nama* mitogenome.

| Gene   | Direction | Location | Size | Anticodon | Start Codon | Stop Codon | Intergenic Nucleotides |
|--------|-----------|----------|------|-----------|-------------|------------|------------------------|
| trnL2  | F         | 3005–3071| 67   | GAT       | 3035–3037   | 5          | 6                      |
| cox2   | F         | 3072–3750| 679  | ATG       | TAA         | 0          | 6                      |
| trnK   | F         | 3748–3818| 71   | GTT       | 3778–3781   | 0          | 6                      |
| trnD   | F         | 3818–3883| 66   | GCT       | 3848–3850   | 0          | 6                      |
| atp8   | F         | 3884–4042| 159  | ATC       | TAA         | 0          | 6                      |
| atp6   | F         | 4036–4713| 678  | ATG       | TAA         | 0          | 6                      |
| cox1   | F         | 4713–5001| 289  | TCC       | 5034–5035   | 0          | 6                      |
| trnG   | F         | 5004–5568| 65   | GAT       | 5534–5535   | 2          | 6                      |
| nad3   | F         | 5566–5922| 357  | GAT       | 5953–5954   | 2          | 6                      |
| trnA   | F         | 5921–5987| 67   | TGC       | 5953–5954   | 2          | 6                      |
| trnF   | R         | 5987–6052| 66   | GAT       | 6034–6036   | 0          | 6                      |
| trnN   | F         | 6053–6118| 66   | GAT       | 6084–6086   | 0          | 6                      |
| trn11  | (AGN)     | 6117–6176| 60   | GAT       | 6134–6136   | 0          | 6                      |
| trnP   | F         | 6180–6243| 64   | GAT       | 6208–6209   | 0          | 6                      |
| nad5   | R         | 6308–8044| 1737 | TAA       | 8071–8073   | 0          | 6                      |
| trnH   | F         | 8018–8106| 88   | TGT       | 8071–8073   | 0          | 6                      |
| nad4   | F         | 8107–9445| 1339 | TAA       | 9447–9731   | 0          | 6                      |
| nad4L  | R         | 9447–9731| 285  | ATG       | TAA         | 1          | 6                      |
| trnL1  | F         | 9744–9807| 64   | TGC       | 9774–9776   | 12         | 6                      |
| trnP   | F         | 9808–9871| 64   | TGC       | 9840–9842   | 0          | 6                      |
| nad6   | F         | 9874–10401| 528  | TAA       | 10,076–10,078| 0          | 6                      |
| cob    | F         | 10,406–11,594| 1149 | ATG       | 11,589–11,591| 0          | 6                      |
| trnL2  | (UCN)     | 11,561–11,624| 64  | TAA       | 11,588–11,591| 13         | 6                      |
| nad1   | R         | 11,638–12,597| 942 | TAA       | 12,623–12,625| 1          | 6                      |
| trnL1  | (CUN)     | 12,703–13,981| 1337 | TAA       | 13,982–14,044| 0          | 6                      |
| rrnL   | R         | 13,982–14,044| 63  | TAA       | 14,014–14,016| 0          | 6                      |
| rrnS   | R         | 14,045–14,818| 774  | TAA       | 14,819–14,820| 0          | 6                      |
| Control region | 14,819–15,208 | 390 | | | | 0 | 6 |
Table 3. Base composition of *Hestina persimilis* mitogenome.

| Size (bp)       | A%  | T%  | G%  | C%  | A + T% | G + C% |
|-----------------|-----|-----|-----|-----|--------|--------|
| mtDNA           | 15,252 | 39.7 | 41.2 | 7.6 | 11.5 | 80.9 | 19.1 |
| PCGs            | 11,222 | 33.8 | 45.9 | 10.4 | 9.9 | 79.7 | 20.3 |
| tRNA            | 1459 | 41.7 | 39.7 | 11.0 | 7.7 | 81.4 | 18.7 |
| rrnL            | 1334 | 44.5 | 39.8 | 10.5 | 5.2 | 84.3 | 15.7 |
| rrnS            | 776 | 43.6 | 41.5 | 10.1 | 4.9 | 85.1 | 15.0 |
| Control region  | 376 | 43.6 | 47.6 | 2.9 | 5.9 | 91.2 | 8.8 |

Table 4. Base composition of *Hestinalis nama* mitogenome.

| Size (bp)       | A%  | T%  | G%  | C%  | A + T% | G + C% |
|-----------------|-----|-----|-----|-----|--------|--------|
| mtDNA           | 15,208 | 39.9 | 39.3 | 7.9 | 12.9 | 79.2 | 20.8 |
| PCGs            | 11,192 | 32.8 | 44.8 | 11.4 | 11.0 | 77.6 | 22.4 |
| tRNA            | 1449 | 42.0 | 39.3 | 10.8 | 7.9 | 81.3 | 18.7 |
| rrnL            | 1327 | 40.2 | 43.4 | 5.3 | 11.1 | 83.6 | 16.4 |
| rrnS            | 774 | 41.2 | 43.9 | 5.2 | 9.7 | 85.1 | 14.9 |
| Control region  | 390 | 44.9 | 43.8 | 2.6 | 8.7 | 88.7 | 11.3 |

Figure 3. Relative Synonymous Codon Usage (RSCU) in the mitogenomes of twelve Apaturinae species. Codon families are provided on the X axis.

Apatura ilia

Apatura metis

Saskaia charonda

Figure 3. Cont.
Figure 3. Relative Synonymous Codon Usage (RSCU) in the mitogenomes of twelve Apaturinae species. Codon families are provided on the X axis.
**Figure 3.** Relative Synonymous Codon Usage (RSCU) in the mitogenomes of twelve Apaturinae species. Codon families are provided on the X axis.

**Apatura ilia**

**Apatura metis**

**Sasakia charonda**

**Figure 4.** Cont.
Figure 4. Cont.
3.3. Transfer RNAs and Ribosomal RNAs

All 22 tRNAs typical of lepidopteran mitogenomes are found in the mitogenomes. Most tRNA genes are in classical clover-leaf secondary structures except for trnS1(AGN), with its DHU arm forming a simple loop, which is considered as a typical feature in metazoan mitogenomes (Figure 5) [37]. Additionally, the anticodon stem of trnS1(AGN) may be shortened as of base mismatch in some insect mitogenomes [38]. Previous studies showed that not only trnS1(AGN), but also some other tRNAs, such as trnS2(UCN) and trnG, lack a DHU or TΨC arm [39]. Missing a DHU arm and base mismatch are thermodynamically unstable, which indicate that a DHU arm might not really exist. Accordingly, this special structure of trnS1(AGN) still needs further investigation. In addition, it has been shown that some isoforms of tRNAs can be found in control regions or some PCGs on the L-strand of mitogenomes. The isoforms of tRNAs can also be folded into cloverleaf structures. However, it is not clear whether their functions are similar to those of tRNAs [40] or not.

The rrlL gene is found between trnL (CUN) and trnV, while the rrsS gene is located between trnV and the control region. The lengths of rrlL genes of H. persimilis and H. nama mitogenomes are 1334 bp (AT content 84.29%) and 1327 bp (AT content 83.44%). The sizes of rrsS genes of H. persimilis and H. nama are 776 bp (AT content 85.03%) and 774 bp (AT content 84.39%). The secondary structure of rrlL genes include six structural domains except for that domain III is absent in arthropods (Figure 6). The rrsS genes include three structural domains (Figure 7). Both the secondary structures of rrlL and rrsS of the two species are roughly similar to other lepidopterans, such as *Amata emma*, *Apis mellifera* [41], *Grapholita molesta* [42], *Manduca sexta* [43], etc. The microsatellite sequence (TA)n is not found in rrlL and rrsS, but exists in other insects (e.g., *Choristoneura longicollina*). There are several new stem loops near the 5′ of rrlL secondary structure, and these loops were not found in other insects.
Figure 5. Predicted secondary cloverleaf structure for the tRNAs of *Hestina persimilis* and *Hestinalis nama*. Dashes (−) and pluses (+) indicate the Watson–Crick base pairings and G-U bonds, respectively.
Figure 6. Predicted rrlL secondary structure of Hestina persimilis and Hestinalis nama mitogenomes.

Dashes, black dots and circles indicate the Watson-Crick base pairings, G-U bonds and U-U, A-A, A-C and A-G bonds, respectively.

3.4. Intergenic and Overlapping Regions

It has been proposed that mitogenomes tend to be highly economized in size by eliminating or reducing intergenic spacers [44]. However, by excluding the control region, 12 intergenic spacers (1 bp 91 bp, 150 bp in total) are found in H. persimilis, and 11 intergenic spacers (1 to 69 bp, 118 bp in total) are found in H. nama. It has been reported that Lepidopteran mitogenomes usually have two typical and relatively conservative intergenic spacers. The longer one is located between trnQ and nad2 genes, with the length of 91 and 69 bp in H. persimilis and H. nama. Previous studies found that the nucleotide sequence of the trnQ-nad2 spacer and the nad2 gene have a highly similarity. It has been inferred that the trnQ-nad2 spacer may come from the nad2 gene [45]. The other shorter spacer is located between trnS2(UCN) and nad1 genes, with the length of 22 and 13 bp in H. persimilis and H. nama, respectively, sharing a conserved sequence of ATACTAA.

Comparing with the intergenic spacers, the overlapping regions are more conservative [46]. Fourteen overlapping spacers (1 to 26 bp, 66 bp in total) are found in H. persimilis, and fourteen overlapping spacers (1 to 8 bp, 42 bp in total) are found in H. nama. ATP8 and ATP6 overlap with the ATGATAA motif in the two mitogenomes, which had also been reported in many other lepidopterans [47].
3.5. Putative Control Regions

The control region, also known as the A + T-region or D-loop, is always the largest intergenic spacer in animal mitogenomes and considered as the initial region for replication [48]. The control regions (376 bp in *H. persimilis* and 390 bp in *H. nama*) in the two mitogenomes are located between *rrnS* and *trnM*. The AT content is also the highest in mitogenomes (91.23% in *H. persimilis* and 88.72% in *H. nama*). Previous studies indicated that the control region is the segment with fastest evolutionary rate and can be used as an important molecular marker for animal population genetics.

There are generally four conserved structures in the control region, including a motif of ATAGA located at downstream of *rrnS* followed by 19 bp Poly-T stretch, a poly-A stretches (9 bp in *H. persimilis* and 6 bp in *H. nama*) at the upstream of *trnM* (Figures 8 and 9), the microsatellite-like repeat regions ((AT)$_{10}$ in *H. persimilis* and (AT)$_{6}$ in *H. nama*), and the repeated sequences (23 bp in *H. persimilis* and 25 bp in *H. nama*). All these characteristics are generally considered to be related to the transcription or replication of...
mitogenomes [49]. Although the location of initial replication region in complete metamorphosis insects (including lepidopterans) are different, they all located after polyT (about 10–20 bp) (Figure 10) [50]. Accordingly, polyT may be involved in the recognition of the initial replication region [51].

**Figure 8.** The control region in *Hestina persimilis.*

**Figure 9.** The control region in *Hestinalis nama.*
3.6. Phylogenetic Analysis

The phylogenetic analyses are performed on concatenated nucleotide sequences of 13 PCGs (data matrix 10,472 bp) derived from 54 available Nymphalidae mitogenomes, with two Papilionidae species serving as outgroups. These 54 sequences represent 9 subfamilies: Apaturinae; Biblidinae; Calinaginae; Danainae; Heliconiinae; Libytheinae; Limenitidinae; Nymphalinae and Satyrinae. The BI (Figure 11) and ML trees (Figure 12) have roughly the same topology, except for Danainae and Libytheinae, which are in different locations. The overall relationship is generally as follows: Libytheinae + ((Calinaginae + Satyrinae) + Danainae + ((Heliconiinae + Limenitidinae) + (Nymphalinae + (Apaturinae + Biblidinae)))). Earlier, Heliconiinae, Limenitidinae, Biblidinae and Apaturinae were considered as “core nymphalids” [52,53]. Subsequently, there are ten widely recognized subfamilies in Nymphalidae, including Apaturinae, Libytheinae, Danainae, Morphinae, Satyrinae, Calinaginae, Eliconiinae, Limenitidinae, Charaxinae and Nymphalinae [54]. Although mitogenomes of the subfamilies Morphinae, Eliconiinae and Charaxinae are not yet available, the topology is generally identical to those of other studies [55–57]. *Hestina* and *Euripus*, *Sasakia* come together to form a branch; *Apatura* and *Hestinalis* form a branch; *Chitoria*, *Timeliae* and *Herona* form a branch. Early on, *Apatura* included *Chitoria* [58], while in this paper, *Chitoria* is distant from *Apatura*. Ohshima et al. (2010) has published an updated work on Apaturinae butterflies, considering that *Chitoria* should been removed from *Apatura* [59], their grouping is concordant to this paper. *H. persimilis*, *H. assimilis* and the genus *Euripus* (represented by *Euripus nyctelius*) form a branch, of which the result is the same as the morphological study. *Hestinalis* is a clearly distinct lineage within the genus *Hestina*, not together with *H. persimilis* and *H. assimilis*, and closely related to Apatura, that is considered to be a result of mimicry [60], which is also consistent with our paper.
Figure 11. Inferred phylogenetic relationships among Lepidoptera based on the concatenated nucleotide sequences of 13 PCGs using BI. Numbers on branches are Bayesian posterior probabilities. *Papilio protenor* (KY272622) and *Lamproptera curius* (KJ141168) are used as outgroups.
4. Conclusions

The mitogenomes of *H. persimilis* and *H. nama* were obtained using sanger sequencing. Comparing them with other mitogenomes of Apaturinae butterflies, the conclusion can be drawn that the mitogenomes are highly conserved, sharing the same gene order, gene location, codon usage, nucleotide composition and AT-biased pattern. The secondary structures of *rrnL* and *rrnS* of two species are roughly similar to other lepidopterans. Although the control regions vary greatly in length, their structure has not changed much, which includes four basic conservative regions. The topology of phylogenetic analyses are generally identical to those of other studies. *Hestinalis nama* is not grouped with *Hestina*, and is closely related to Apatura, which is consistent with early studies.
Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/insects12080754/s1, Figure S1: phylogenetic relationships by BI analysis based on the 13PCGs. Figure S2: phylogenetic relationships by ML analysis based on the 13PCGs. Table S1: sequences of 22 pairs of primers. Table S2: list of species used to construct the phylogenetic tree. Table S3: the starting partitions used to initiate the PartitionFinder analysis. Table S4: evolutionary models from partition strategies start scheme used in the phylogenetic analysis. Table S5: bias of base composition of protein-coding genes in the two mitogenomes. Table S6: codon usage of protein genes of twelve mitogenomes in Apaturinae.

Author Contributions: Conceptualization, Y.W. and T.C.; specimen collection and identification, J.W. (Jiping Wen), J.W. (Juping Wang) and T.C.; methodology, Y.W. and H.F.; software, Y.W., B.H. and H.F.; analysis, J.W. (Juping Wang), Y.W. and H.F.; writing—original draft preparation, Y.W.; writing—review and editing, Y.W. and B.H.; funding acquisition, Y.W. and T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Science and Technology basic condition platform of Shanxi Province (201605D121020), the Innovative Research Project in Shanxi Academy of Agricultural Sciences, China (YGPJY1908), the open fund of Shanxi Key Laboratory of Integrated Pest Management in Agriculture (YHSW2019001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All sequences were deposited in the GenBank under accession numbers MT110153 and MT110154.

Acknowledgments: We sincerely express our gratitude to Tianjuan Su (School of Life Sciences, Jinggangshan University, Jinggangshan, China) for reviewing the manuscript. Thanks to Chunsheng Wu and Qianju Jia (Key Laboratory of Zoological Systematics and Evolution (CAS), Institute of Zoology, Chinese Academy of Sciences, Beijing, China) for the literature inquiry and reviewing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Doubleday, E. List of the Specimens of Lepidopterous Insects in the Collection of the British Museum; Trustees of the British Museum: London, UK, 1844; Volume 1, p. 97.
2. Chou, I. Monograph of Chinese Butterflies; Henan Scientific Technological Publishing House: Zhengzhou, China, 1994; pp. 451–456.
3. Lee, C.L.; Zhu, B.Y. Atlas of Chinese Butterflies; Shanghai Far East Publishers: Shanghai, China, 1992; p. 136.
4. Masui, A.; Bozano, G.C.; Floriani, A. Guide to the Butterflies of the Palearctic Region: Nymphalidae 4: Apaturinae; Omnes Artes: Milan, Italy, 2011; Volume 4, pp. 45–102.
5. Lang, S.Y. The Nymphalidae of China (Lepidoptera, Rhopalocera); Tshikolovets Publications: Pardubice, Czech Republic, 2012; Volume 1, p. 21.
6. Wu, C.S.; Hsu, Y.F. Butterflies of China; The straits Publishing &Distributing Group: Fuzhou, China, 2017; pp. 871–874.
7. Simmons, R.B.; Weller, S.J. Utility evolution of cytochrome b in insects. Mol. Phylogenet. Ecol. 2001, 20, 196–210. [CrossRef]
8. Hebert, P.D.N.; Cywinska, A.; Ball, S.L.; de Waard, J.R. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B Bio. 2003, 270, 313–321. [CrossRef]
9. Regier, J.G.; Mitter, C.; Zwick, A.; Bazinet, A.L.; Cummings, M.P.; Kawahara, A.Y.; Sohn, J.C.; Zwickl, D.J.; Cho, S.; Davis, D.R.; et al. A large-scale, higher–level, molecular phylogenetic study of the insect order Lepidoptera (moths butterflies). PLoS ONE 2013, 8, e58568. [CrossRef] [PubMed]
10. Wu, Y.P.; Zhao, J.L.; Su, T.J.; Luo, A.R.; Zhu, C.D. The complete mitochondrial genome of Choristoneura longicellana (Lepidoptera: Tortricidae) phylogenetic analysis of Lepidoptera. Gene 2016, 591, 161–176. [CrossRef] [PubMed]
11. Dong, W.W.; Dong, S.Y.; Jiang, C.F.; Huang, G.H. Characterization of the complete mitochondrial genome of tea tussock moth, Euproctis pseudoconspersa (Lepidoptera: Lymantridae) its phylogenetic implications. Gene 2016, 577, 37–46. [CrossRef] [PubMed]
12. Ratnasingham, S.; Hebert, P.D.N. Bold: The Barcode of Life Data System. Mol. Ecol. Notes 2007, 7, 355–364. [CrossRef] [PubMed]
13. Lafontaine, J.D.; Figer, M. Revised higher classification of the Noctuoidea (Lepidoptera). Can. Entomol. 2006, 138, 610–635. [CrossRef]
14. Mitter, C.; Davis, D.R.; Cummings, M.P. Phylogeny Evolution of Lepidoptera. Annu. Rev. Entomol. 2017, 62, 265–283. [CrossRef]
15. Yang, M.S.; Song, L.; Shi, Y.X.; Li, J.H.; Zhang, Y.L.; Song, N. The first mitochondrial genome of the family Epicopeiidae higher–level phylogeny of Macroheterocera (Lepidoptera: Ditrysia). Int. J. Biol. Macromol. 2019, 136, 123–132. [CrossRef]
16. Mueller, R.L. Evolutionary rates, divergence dates, the performance of mitochondrial genes in bayesian phylogenetic analysis. *Syst. Biol.* **2006**, *55*, 289–300. [CrossRef]

17. Timmermans, M.J.T.N.; Lees, D.C.; Simonsen, T.J. Towards a mitogenomic phylogeny of Lepidoptera. *Mol. Phylogenet. Evol.* **2014**, *79*, 169–178. [CrossRef]

18. Liu, N.Y.; Li, N.; Yang, P.Y.; Sun, C.Q.; Fang, J.; Wang, S.Y. The complete mitochondrial genome of *Damora sagana* phylogenetic analyses of the family Nymphalidae. *Genes Genom.* **2018**, *40*, 109–122. [CrossRef]

19. Habib, M.; Lakra, W.S.; Mohindra, V.; Khare, P.; Barman, A.S.; Singh, A.; Lal, K.K.; Punia, P.; Khan, A.A. Evaluation of cytochrome b mtDNA sequences in genetic diversity studies of *Chama maruius* (Channidae: Perciformes). *Mol. Biol. Rep.* **2011**, *38*, 841–846. [CrossRef]

20. Chris, S.; Thomas, R.B.; Francesco, F.; James, B.S.; Andrew, T.B. Incorporating molecular evolution into phylogenetic analysis, a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.* **2006**, *37*, 545–579.

21. Simon, C.; Frati, F.; Bekenbach, A.; Crespi, B.; Liu, H.; Flook, P. Evolution, weighting, phylogenetic utility of mitochondrial gene-sequences a compilation of conserved polymerase chain reaction primers. *Am. Entomol. Soc. Am.* **1994**, *87*, 651–701. [CrossRef]

22. Lanfear, R.; Frandsen, P.B.; Wright, A.M.; Senfeld, T.; Calcott, B. PartitionFinder 2: New methods for selecting partitioned models likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef]

23. Miller, M.A.; Pfeiffer, W.T.; Schwartz, T. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees; Gateway Computing Environments Workshop (GCE): New Orleans, LA, USA, 2010; pp. 1–7.

24. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis post–analysis large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]

25. Clary, D.O.; Wolstenholme, D.R. The mitochondrial DNA molecular of *Drosophila yakuba*: Nucleotide sequence, gene organization, genetic code. *J. Mol. Evol.* **1985**, *22*, 252–271. [CrossRef] [PubMed]

26. Lu, H.F.; Su, T.J.; Luo, A.R.; Zhu, C.D.; Wu, C.S. Characterization of the complete mitochondrion genome of Diurnal moth *Amata emma* (Butler) (Lepidoptera: Erebidae) its phylogenetic implications. *Mol. Biol. Evol.* **2000**, *17*, 252–271. [CrossRef]

27. Yuan, M.L.; Zhang, Q.L.; Guo, Z.L.; Wang, J.; Shen, Y.Y. The complete mitochondrial genome of *Corizus tetraspilus* (Hemiptera: Ropalaidae) phylogenetic analysis of Pentatomomorpha. *PloS ONE* **2015**, *10*, e0129003.

28. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [CrossRef]

29. Lanfear, R.; Calcott, B.; Ho, S.Y.W.; Guindon, S. Partitionfinder: Combined selection of partitioning schemes substitution models for phylogenetic analyses. *Mol. Biol. Evol.* **2012**, *29*, 1695–1701. [CrossRef]

30. Lanfear, R.; Frandsen, P.B.; Wright, A.M.; Senfeld, T.; Calcott, B. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular morphological phylogenetic analyses. *Mol. Biol. Evol.* **2017**, *34*, 772–773. [CrossRef]

31. Clary, D.O.; Wolstenholme, D.R. The mitochondrial DNA molecular of *Drosophila yakuba*: Nucleotide sequence, gene organization, genetic code. *J. Mol. Evol.* **1985**, *22*, 252–271. [CrossRef] [PubMed]

32. Yuan, M.L.; Zhang, Q.L.; Guo, Z.L.; Wang, J.; Shen, Y.Y. The complete mitochondrial genome of *Corizus tetraspilus* (Hemiptera: Ropalaidae) phylogenetic analysis of Pentatomomorpha. *PloS ONE* **2015**, *10*, e0129003.

33. Hershberg, R.; Petrov, D.A. Selection on codon bias. *Annu. Rev. Genet.* **2008**, *42*, 287–299. [CrossRef] [PubMed]

34. Rodriguez–Trelles, F.; Tarrio, R.; Ayala, F.J. Fluctuating mutation bias the evolution of base composition in Drosophila. *J. Mol. Evol.* **2000**, *50*, 1–10. [CrossRef] [PubMed]

35. Rodriguez–Trelles, F.; Tarrio, R.; Ayala, F.J. Fluctuating mutation bias the evolution of base composition in Drosophila. *J. Mol. Evol.* **2000**, *50*, 1–10. [CrossRef] [PubMed]
44. Sheffield, N.C.; Song, H.; Cameron, S.L.; Whiting, M.F. A comparative analysis of mitochondrial genomes in coleoptera genome
descriptions of six new beetles. Mol. Biol. Evol. 2008, 25, 2499–2509. [CrossRef] [PubMed]
45. Wu, Y.P.; Zhao, J.L.; Su, T.J.; Li, J.; Yu, F.; Chesters, D.; Fan, R.J.; Chen, M.C.; Wu, C.S.; Zhu, C.D. The Complete mitochondrial
genome of Leucoptera malifoliella Costa (Lepidoptera: Lyconetiidae). DNA Cell Biol. 2012, 31, 1508–1522. [CrossRef]
46. Cao, Y.Q.; Ma, C.; Chen, J.Y.; Yang, D.R. The complete mitochondrial genomes of two ghost moths, Thitarodes renzhensis and
Thitarodes yunnanensis: The ancestral gene arrangement in Lepidoptera. BMC Genom. 2012, 13, 276. [CrossRef]
47. Yang, M.; Shi, S.; Dai, P.; Song, L.; Liu, X. Complete mitochondrial genome of Palpita hypohomalia (Lepidoptera: Pyraloidea:
Crambidae) and its phylogenetic implications. Eur. J. Entomol. 2018, 115, 708–717. [CrossRef]
48. Tayler, M.F.; McKechnie, S.W.; Pierce, N.; Kreitman, M. The lepidopteran mitochondrial control region: Structure evolution. Mol.
Biol. Evol. 1993, 10, 1259–1272.
49. Clayton, D.A. Transcription replication of animal mitochondrial DNAs. Int. Rev. Cyt. 1992, 141, 217–232.
50. Saito, S.; Tamura, K.; Aotsuka, T. Replication origin of mitochondrial DNA in insects. Genetics 2005, 171, 1695–1705. [CrossRef]
51. Ye, W.; Dang, J.P.; Xie, L.D.; Huang, Y. Complete mitochondrial genome of Teleogryllus emma (Orthoptera: Gryllidae) with a new
gene order in Orthoptera. Zool. Res. 2008, 29, 236–244. [CrossRef]
52. Ehrlich, P.R. The comparative morphology, phylogeny higher classification of the butterfly (Lepidoptera: Papilionoidea).
Univ. Kans. Sci. Bull. 1958, 39, 305–370.
53. Scott, J.A. The phylogeny of butterflies (Papilionoidea Hesperoidea). J. Res. Lepid. 1985, 23, 24–281.
54. Kristensen, N.P.; Scoble, M.J.; Karsholt, O. Lepidoptera phylogeny systematic: The state of inventorying moth butterfly diversity.
Zootaxa 2007, 1668, 699–747. [CrossRef]
55. Wahlberg, N.; Weingartner, E.; Nylin, S. Towards a better understanding of the higher systematics of Nymphalidae (Lepidoptera:
Papilionoidea). Mol. Phylogenet. Evol. 2003, 28, 473–484. [CrossRef]
56. Wahlberg, N.; Wheat, C.W. Genomic outposts serve the phylogenomic pioneers: Designing novel nuclear markers for genomic
DNA extractions of Lepidoptera. Syst. Biol. 2008, 57, 231–242. [CrossRef] [PubMed]
57. Wahlberg, N.; Leneveu, J.; Kodandaramaiah, U.; Peña, C.; Nylin, S.; Freitas, A.V.L.; Brower, A.V.Z. Nymphalid butterflies diversify
following near demise at the Cretaceous/Tertiary boundary. Proceedings of the Royal Society of London Series B Biological
Sciences. Proc. Biol. Sci. 2009, 276, 4295–4302. [PubMed]
58. Freitas, A.V.L.; Brown, K.S. Phylogeny of the Nymphalidae (Lepidoptera). Syst. Biol. 2004, 53, 363–383. [CrossRef]
59. Zhang, M.; Cao, T.W.; Zhang, R.; Guo, Y.P.; Duan, Y.H.; Ma, E.B. Phylogeny of Apaturinae Butterflies (Lepidoptera: Nymphalidae)
based on mitochondrial cytochrome oxidase I gene. J. Genet Genom. 2007, 34, 812–823. [CrossRef]
60. Ohshima, I.; Tanikawadodo, Y.; Saigusa, T.; Nishiyama, T.; Kitani, M.; Hasebe, M.; Mohri, H. Phylogeny, biogeography, host-plant
association in the subfamily Apaturinae (Insecta: Lepidoptera: Nymphalidae) inferred from eight nuclear seven mitochondrial
genes. Mol. Phylogenet. Evol. 2010, 57, 1026–1036. [CrossRef] [PubMed]