Comparative Mitochondrial Analysis of *Cnaphalocrocis exigua* (Lepidoptera: Crambidae) and Its Close Relative *C. medinalis*

Kai-Jun Zhang,* Xia Rong,* Zhao-Peng Lv, Lei Liu, Guo-Yang Li, and Huai Liu

Key Laboratory of Entomology and Pest Control Engineering, College of Plant Protection, Southwest University, Beibei District, 400715, Chongqing, China and 1Corresponding author, e-mail: liuhuai@swu.edu.cn

*These authors contributed equally to this work.

Subject Editor: Margaret Allen

Received 7 May 2020; Editorial decision 4 September 2020

Abstract

Rice leafrollers are important pests on rice in Asia, Oceania, and Africa, causing serious loss to rice production. There are two main rice leafrollers in China, namely *Cnaphalocrocis medinalis* (Guenee) and *C. exigua* (Butler) with the former having the ability of long-distance migration. To reveal the differences in the mitochondrial genomes (mitogenome) between them, we compared the completed mitogenome of *C. exigua* with three *C. medinalis* individuals. Although phylogenetic analysis based on the mitogenomic data strongly supported the close relationship between these two species, many differences were still being revealed. The results showed that the mitogenome of *C. exigua* was shorter in length (15,262 bp) and slight lower in AT content than that of *C. medinalis*. Except for the different start codons of nad3 and nad6 genes, we also found the cox1 gene had a typical start codon ‘ATG’ which suggested that the starting position of this gene must be reconsidered in the entire superfamily Pyraloidea. All tRNAs have a typical clover-leaf structure, except for the dihydrouridine (DHU) stem losing of trnS1, which has the atypical anticondon ‘TCT’ instead of ‘GCT’ in *C. medinalis* and most Pyraloidea species. Two intergenic regions (between trnY and cox1, nad3 and trnA) featured by AT repeats were only found in *C. medinalis* and even rarely appeared in reported Pyraloidea species. Furthermore, regardless of interspecific comparison or intraspecific comparison of these two species, protein coding genes, especially the *atp8* genes, had quite different evolutionary rates.

Key words: *Cnaphalocrocis exigua*, *Cnaphalocrocis medinalis*, Mitogenome, Pyraloidea

Two main rice leafrollers, *Cnaphalocrocis exigua* and *C. medinalis* (Lepidoptera: Pyraloidea: Crambidae) are widely distributed in the tropical and temperate areas of Asia, Oceania, and Africa, which cause heavy loss on the rice production. They create longitudinal white and transparent streaks on the leaves by attaching the leaf margins together with silk strands and feeding inside (Khan et al. 1988; Park et al. 2010, 2014; Padmavathi et al. 2013; Yang et al. 2015). Because the two species have similar morphology and habits, they are generally confused in the field in China (Pan 1984, 1985). The rice leafroller *C. medinalis* had drawn much attention from researchers worldwide for its ability to migrate and re-migrate over long distances (Gao et al. 2008, Huang et al. 2010, Wang et al. 2010). To date, only a few studies were proceeded on *C. exigua*, and there was no evidence that support *C. exigua* can migrate yet (Feng et al. 2017, Liao et al. 2017). Though *C. exigua* is not as widely distributed as *C. medinalis* in China, but it has dominated paddy fields in many places of Sichuan Basin in southwest China which suggests it become a principal rice pest (Pan 1984, Gao et al. 2008, Yang et al. 2015).

The mitogenome has been widely applied as an useful molecular marker for diverse evolutionary studies among species including phylogenetic, population genetics, and comparative and evolutionary genomics (Harrison 1989, Boore 1999, Babbucci et al. 2014, Cameron 2014). Mitochondria are the energy-producing organelles in eukaryotic cells (Scott et al. 2011). The excellent flight ability of *C. medinalis* suggests it is a qualifier with developed energy metabolic level. Thus, we infer the mitogenome of *C. medinalis* differs from that of *C. exigua* to some extent.

The Pyraloidea, with more than 15,570 species described worldwide, is one of the largest superfamilies in Lepidoptera and shows the most diverse life history adaptations (Nieukerken et al. 2011, Regier et al. 2012). Up to now, only about 40 complete or nearly complete mitogenomes from Pyraloidea species have been sequenced. A better understanding of Pyraloidea mitogenomes also requires an expansion of the taxon and genome samplings.

In this study, we report the complete mitogenome of *C. exigua* and compare with that of its close relative *C. medinalis*. Detailed
mitogenomic information of these two important rice pests may help us further understanding the mechanisms of migration of *C. medinalis* and creating new ways to control these pests. Moreover, phylogenetic analysis was also conducted to further study the phylogenetic relationships within Pyraloidea species.

**Materials and Methods**

**DNA Sample Extraction**

The overwintering *C. exigua* pupae were collected from rice stubble fields in Qianwei county (Leshan, Sichuan Province, China; E103° 93′, N29° 21′) in March 2016 and reared in an incubator (LAC-250HPY-2, Shanghai Longyue Instruments, Shanghai, China) under constant conditions (26 ± 1°C, 80 ± 5% RH, and a photoperiod of 14:10 (L:D) h) for two generations (Liao et al. 2017). The newly emerged adult was kept in ethanol (100%) and stored at −80°C till the isolation of DNA. Total genomic DNA was isolated from an individual of *C. exigua* using the MiniBEST Universal Genomic DNA Extraction Kit (Ver. 5.0 TaKaRa, Japan), according to the manufacturer’s protocol. The quality of isolated DNA was examined by 0.7% (w/v) agarose gel electrophoresis and then the DNA was used to amplify the entire mitogenome of *C. exigua*.

**PCR Amplification, Cloning, and Sequencing**

To obtained the whole genome sequence, two-pair of primers (Ce12Sr: GAA AGC GAC GGG CAA ATC TAT GT and CeCO1r722: TAA ACT TCT GGA TTA GAT ACC CTA TTA T (Simon et al. 1994) and CeCOH14: AYT CWA CAA ATC ATA AAG TTA TG G) were designed to amplify one long fragment cox1-rrnS and one short fragment *rrns-cox1*, respectively. PCR reaction mixture (50 µl in total) included 10 µl 5× PrimeSTAR GXL Buffer, 5 µl dNTP mixture (2.5 mM each), 1 µl PrimeSTAR GXL DNA Polymerase (1.25 U/µl, TaKaRa, Japan), 1 µl primer (10 µM) each, 2 µl template DNA, and 30 µl double distilled water. The amplification program for this two fragments was performed in Easycycler Gradient 96 (Analytik Jena AG, Germany) under the following conditions: an initial denaturation at 98°C for 10 s, followed by 32 cycles, at 98°C for 10 s, 56°C for 15 s, 68°C for 8 min (5 min for the short fragment *rrns-cox1*), and a final extension one cycle at 68°C for 5 min. The PCR products were subjected to 0.7% (w/v) agarose gel electrophoresis and purified using agarose Gel DNA Extraction Kit (MiniBEST Ver. 4.0, TaKaRa, Japan). Then, these two purified fragments were directly sequenced by step through BGI Genome-sequencing firm. To minimize in vitro amplification errors, each PCR had four replications. The complete mitochondrial genome sequence of *C. exigua* obtained in this study was deposited in GenBank under the accession numbers MN877384.

**Genome Annotation and Secondary Structure Prediction**

Protein-coding genes (PCGs) and rRNA genes were identified based on homologous regions of previously sequenced Pyraloidea mitogenomes using the Clustal X version 2.0 (Larkin et al. 2007), and the GeneDoc version 2.6 software. The base composition and the GeneDoc version 2.6 software. The base composition and sequence features of being capable of folding into the typical cloverleaf secondary structure with legitimate anticodon. The Map of the mitogenome was drawn by using the GView Server (http://stothard.afns.ualberta.ca/cgview_server/).

**Phylogenetic Analysis**

In addition to the sequenced mitogenome of *C. exigua*, 39 mitogenome sequences of 37 Pyraloidea species were downloaded from the NCBI database (Supp Table 1 [online only]). Two Tortricoidae species, *Cydia pimoniella* (GenBank JX407107) and *Adoxophyes orana* (GenBank JX872403) were served as outgroups. The amino acid and nucleotide sequences of each of the 13 PCGs were aligned by Muscle through MEGA 7.0 software (Kumar et al. 2016), and then the concatenated set of nucleotide sequences from all 13 PCGs was used for phylogenetic analyses using Bayesian inference (BI) and maximum likelihood (ML) methods. The nucleotide substitution model test performed in MEGA version 7.0 showed that the GTR+I+G was the best-fitting model for the concatenated data. The BI analyses were implemented in MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001) with four MCMC chains running for 2,000,000 generations and sampling every 2,000 generations. After discarding the first 25% samples as burn-in, Bayesian posterior probability values were calculated in a consensus tree. The ML analyses were performed using MEGA version 7.0 with 500 bootstrap replicates.

**Results and Discussion**

**General Features of the *C. exigua* Mitogenome**

The mitogenome of *C. exigua* was a closed-circular molecule of 15,262 in length (Table 1), which was within the range observed in the sequenced other Pyraloidea mitogenomes available with the size ranging from 15,110 bp in *Maruca testulalis* (Zou et al. 2016) to 15,594 bp in *Ephestia kuehniella* (Zhu et al. 2018). Annotation of the *C. exigua* mitogenome indicated that the structure and orientation of the 13 PCGs, 22 tRNA genes, 2 rRNA genes, and a noncoding region (A+T-rich control region) were typical of and similar to the other closely related pyraloid species (Yang et al. 2018, Zhu et al. 2018). Nine PCGs (nad2, nad3, nad6, cox1-3, atp6, atp8, cyt b), 14 tRNAs, and the control region were located on the major J-strand. Fourteen genes (4 PCGs: nad5, nad4, nad4l, and nad1; both rRNA genes; and 8 tRNA genes) were located on the minor
| Feature            | Strand | Start–end | Size  | Intergenic nucleotides | Anti/start codon | Stop codon |
|--------------------|--------|-----------|-------|------------------------|------------------|------------|
| tRNA-Met (M)       | J      | 1–68      | 68    |                        | CAT              | CAT        |
| tRNA-Ile (I)       | J      | 69–133    | 65    |                        | GAT              | GAT        |
| tRNA-Gln (Q)       | N      | 133–201   | 69    | -3                     | TTG              | CAT        |
| ND2                | J      | 268–1,281 | 1,014 | 1,017                  | ATT              | TAA        |
| tRNA-Trp (W)       | J      | 1,288–1,355 | 68   | 5                      | TCA              | TCA        |
| tRNA-Cys (C)       | N      | 1,348–1,414 | 67   | -8                     | GCA              | GCA        |
| tRNA-Tyr (Y)       | N      | 1,424–1,489 | 66   | 9                      | GTA              | T-         |
| ND2                | J      | 1,494–3,027 | 1,534| 1,531                  | ATG              | TAA        |
| tRNA-Leu (UUR)     | J      | 3,028–3,094 | 67   | 0                      | ATT              | CAT        |
| COX1               | J      | 3,095–3,776 | 682  | 682                    | ATG              | T-         |
| tRNA-Lys (K)       | J      | 3,777–3,847 | 71   | 71                     | CTT              | CTT        |
| tRNA-Asp (D)       | J      | 3,856–3,921 | 66   | 11                     | GTC              | GTC        |
| ATP6               | J      | 4,074–4,748 | 675  | 7                      | ATT              | TAA        |
| COX3               | J      | 4,748–5,536 | 789  | 789                    | ATG              | TAA        |
| tRNA-Gly (G)       | J      | 5,539–5,603 | 65   | 65                     | TCC              | TCC        |
| ND3                | J      | 5,604–5,957 | 354  | 354                    | ATC              | CAT        |
| tRNA-Ala (A)       | J      | 5,967–6,032 | 66   | 66                     | TGC              | TGC        |
| tRNA-Agg (R)       | J      | 6,032–6,098 | 67   | 67                     | TCG              | TCG        |
| tRNA-Asn (N)       | J      | 6,099–6,164 | 66   | 7                      | GGT              | GGT        |
| tRNA-Ser (AGN)     | J      | 6,165–6,230 | 66   | 2                      | TCT              | TCT        |
| tRNA-Glu (E)       | J      | 6,235–6,300 | 66   | 4                      | TTC              | TTC        |
| tRNA-Phe (F)       | N      | 6,299–6,646 | 68   | 69                     | GAA              | GAA        |
| ND5                | J      | 6,367–8,101 | 1,735| 1,735                  | ATT              | T-         |
| tRNA-His (H)       | N      | 8,102–8,169 | 68   | 68                     | GTC              | GTC        |
| ND4                | J      | 8,315–9,803 | 294  | 294                    | TGG              | TGG        |
| tRNA-Thr (T)       | J      | 9,806–9,870 | 65   | 65                     | TGT              | TGT        |
| tRNA-Pro (P)       | J      | 9,971–9,937 | 67   | 0                      | TGG              | TGG        |
| ND6                | J      | 9,940–10,476 | 537  | 2                       | ATC              | TAA        |
| tRNA-Leu (UCN)     | J      | 10,481–11,629 | 1,149| 1,149                  | ATG              | TAA        |
| ND1                | N      | 11,709–12,467 | 939  | 939                    | ATM              | TAA        |
| tRNA-Leu (CN)      | N      | 12,649–12,717 | 69   | 69                     | TAG              | TAG        |
| l-rRNA             | N      | 12,718–14,074 | 1,357| 1,348                  | TAC              | TAC        |
| tRNA-Val (V)       | N      | 14,075–14,140 | 66   | 66                     | TAG              | TAG        |
| s-rRNA             | N      | 14,141–14,922 | 782  | 781                    | TAC              | TAC        |
| Control Region     | J      | 14,923–15,262 | 340  | 340                    | TAC              | TAC        |

*Ce* indicates GenBank JN246082; *Cm1* indicates GenBank JQ305693; *Cm3* indicates GenBank JQ647917.
N-strand (Fig. 1). Distinct to the predicted ancestral gene order for insects, trnM-trnL-trnQ instead of trnL-trnQ-trnM was recognized in C. exigua, which was consistent to most Ditrysia species of Lepidoptera (Timmermans et al. 2014).

The commonest start codon of PCGs was the ATG (in 8 PCGs: cox1, cox2, atp6, cox3, nad4, nad4l, cob, nad4), followed by three for ATT (nad2, atp8, nad5) and two for ATC (nad3, nad6) (Table 1). Nine PCGs had TAA stop codon, the remaining genes had incomplete stop codons (TA in nad4; T in cox1, cox2, and nad5).

Similar to other pyraloid species, the C. exigua mitogenome nucleotide composition was bias toward adenine and thymine (accounting for 81.6%: A = 40.5%, T = 41.1%, C = 10.9% G = 7.5%) (Yang et al. 2018, Zhu et al. 2018). The nucleotide bias was also reflected in the codon usage. Base composition at each codon position of concatenated 13 PCGs showed that the A+T content of the third codon positions were significantly higher than the first and second positions. In particular, T in each codon position of PCGs was over represented (Supp Table 2 [online only]).

All the 22 identified tRNA genes in C. exigua mitogenome were folded into the typical clover-leaf structure of mitochondrial tRNAs, except for trnS1 (AGN) in which the dihydrouridine (DHU) arm failed to form a stable stem-loop structure (Supp Table 3 [online only]). The exception trnS1 had been observed in many Pyraloidea and other insect mitogenomes (Zhang et al. 2016, Zhang and Ye 2017, Yang et al. 2018, Zhu et al. 2018). The length and base composition of these tRNAs were similar to other pyralids (Yang et al. 2018, Zhu et al. 2018). The number of base pairs in the DHU-stem ranged from 3 to 4. All the TWC-stem of C. exigua had 5 base pairs except 3 bp in trnT, 6 bp in trnS1, and 4 bp in trnG, trnN and trnF (Supp Table 3 [online only]). Six unmatched base pairs including five U-U and one A-C pairs were identified in the stem region of tRNAs. The mismatch of base pairs existed widely in Pyraloidea mitogenomes (Chai et al. 2012, Yang et al. 2018).

Two ribosomal genes, lrRNA and srRNA, were 1357 bp and 782 bp, located between trnL1 (CUN) and trnV, and between trnV and the A+T-rich region, respectively (Table 1). Both the AT-skew and GC-skew of tRNA genes and rRNA genes were slightly positive in the two leaffolders (Supp Table 2 [online only]). The A+T-rich region of C. exigua mitogenome extended over 340 bp with extremely high AT content (94.4%) and was located between the rrnS and trnM genes (Table 1; Supp Table 2 [online only]).

### Comparative Analysis of C. exigua and C. medinalis Mitogenomes

In order to reveal the difference between the two leaffolders, we compared the mitogenome of C. exigua with that of three individuals of C. medinalis. In general, the analysis results showed that differences between mitochondrial genomes exist not only between the two leaffolders but also within species of C. medinalis.

The PCGs

Except for cox1, nad3, and nad6 gene, the start codons of other genes were the same in the mitogenomes of the two leaffolders (Table 1). Including C. medinalis, most of the pyralis had ATT start codon for both nad3 and nad6 genes (Yang et al. 2018, Zhu et al. 2018). Besides C. exigua, the ATC start codon of nad3 could be found in Dichocrocis punctiferalis, Pseudargyria interruptella, and Scirpophaga incertulas, and the ATC start codon of nad6 was only found in Spoladea recurvalis (Wu et al. 2013, He et al. 2015, Song et al. 2016). Interestingly, for the cox1 gene, we found that it has the common start codon ‘ATG,’ which is different from the previously reported that ‘CGA’ as the start codon of this gene of most Pyraloidea species. We reanalyzed the starting position of the cox1 genes of all reported Pyraloidea species and found that other four species in family Crambidae (GenBank JX144892, KJ739310, KM453724, KC493629) also have the common start codon ‘ATG.’ In addition, although the three bases ‘TAG’ before the start codon ‘CGA’ appeared in many species of Pyraloidea, the ‘TTG’ in the same position could also be found in

---

**Fig. 1.** Circular mitochondrial genome map of C. exigua and C. medinalis. Genes encoded on the heavy strand or light strand are represented inside or outside of the circular mitochondrial genome map. The tRNA genes are abbreviated by triple letter, with Leu1 = CUN, Leu2 = UUR, Ser1 = AGN, and Ser2 = UCN. The second circle shows the GC content and the third shows GC skew calculated as (G-C)/(G+C). The GC content and GC skew are plotted as the deviation from the average value of the entire sequence. The circular mitochondrial genome map of C. medinalis was drawn based on the sequence GenBank JN246082.
Fig. 2. The codon distribution (A) and the relative synonymous codon usage (RSCU) (B) of *C. exigua* (*Ce*) and *C. medinalis* (*Cm*) mitochondrial genomes. CDspT = codons per thousand codons. Codon families are plotted on the x-axis. Codons represented above the bar are not found in the mitogenomes. *Cm1* indicates GenBank JN246082; *Cm2* indicates GenBank JQ305693; *Cm3* indicates GenBank JQ647917.
six species (including one of three individuals of *C. medinalis*) of family Crambidae (GenBank KP347977, KM244688, JF339041, KJ174087, KF859965, JQ305693) and two species of Gallerinae in family Pyralidae (GenBank HQ897685, KT750964). In view of the above reasons, we suggest that when annotating mitogenome of species in Pyraloidea, it is necessary to reconsider the starting position of this gene.

The distribution of codon families in two leaffolders were very similar, and both exhibited that leucine (Leu), isoleucine (Ile), and phenylalanine (Phe) were the three most abundant codon families, whereas codon family Cysteine (Cys), arginine (Arg), and glutamate (Glu) were the three least abundant (Fig. 2A). This situation was in accordance with other Pyraloidea species (Dai et al. 2018, Zhu et al. 2018). Interestingly, the Phe amount of *Cm1* was higher than that of *Ce, Cm2*, and *Cm3*. The RSCU analysis showed that total 5 and 7 codons could not be identified in the mitogenomes of *C. exigua* and *C. medinalis*, respectively (Fig. 2B). As reported before, missing codons ranged from 1 to 8 were found in most known Pyraloidea mitogenomes (Yang et al. 2018, Zhu et al. 2018). Particularly, the codons GCG and AGG were not presented in *C. medinalis* mitogenome but could be found in *C. exigua* mitogenome (Fig. 2B). Therefore, *C. medinalis* had more missing codons than *C. exigua*, and these missing codons showed high G/C content as the previously reported (Zhu et al. 2018), which may be one of the reasons that the A+T content of PCGs of *C. medinalis* is slightly higher than that of *C. exigua*.

The genetic distance analysis of mitochondrial genes of *C. exigua* and *C. medinalis* revealed that regardless of 13 PCGs, 2 tRNA genes, and 22 tRNA genes, the difference was greater than that of within species. In both interspecific and intraspecific comparisons, the *rrnL* gene showed greater variability than *rrnS* gene and tRNA genes. Among 13 PCGs, the genetic distance of each gene between species was greater than 0.03, ranging from 0.0317 (*nad4l*) to 0.1973 (*atp8*), while the genetic distance within *C. medinalis* ranged from 0 (*cox3*, *nad3*) to 0.1139 (*atp8*) (Table 2). Notably, mitochondrial genes (amino acid sequences) of *Cm2* and *Cm3* were more homologous than that of *Cm1* and *Cm2* (8 identical PCGs versus 3 identical PCGs). Since *Cm1* and *Cm2* were both collected from China, while *Cm3* was collected from Korea, it is hard to explain why *Cm2* and *Cm3* are more similar in mitochondrial genes. It can only speculate that there may be significantly different individuals in the natural population of *C. medinalis*.

We also did evolutionary force analysis of each PCG of *C. exigua* and *C. medinalis* through calculating the rate of nonsynonymous (Ka), the rate of synonymous (Ks), and the ratio of Ka/Ks, respectively. In the comparison of the two species, the values of Ka/Ks of PCGs were all less than 1, which indicated that these genes were evolving under purifying selection. However, the ratio of Ka/Ks of *atp8* gene was much higher than other PCGs (Supp Table 4 [online only]). Notably, the *atp8* gene sequence of *Cm1* showed too different from that of *Cm2* and *Cm3*, there was only 1 amino acid distinction between *Cm2* and *Cm3*, but exist 13 and 14 amino acid variations between *Cm1–Cm2* pair and *Cm1–Cm3* pair, respectively. Furthermore, the ratios of Ka/Ks for *atp8* in *Cm1–Cm2* pair (5.58) and *Cm1–Cm3* pair (5.943) were even much higher than that of *Cm1–Cm2* pairs (0.993 for *Cm1–Cm1* pair, 0.814 for *Cm2–Cm2* pair, and 0.745 for *Cm3–Cm3* pair), and it implied that this gene might had suffered some kind of positive selection in partial populations of *C. medinalis*.

Since the *atp8* gene exhibits significantly different characteristics, we further compared the *atp8* gene of 38 species of Pyraloidea. As the result, the *atp8* gene showed polymorphic in size not only within *Cnaphalocrocis*, but also within superfamily Pyraloidea (ranged

---

**Table 2. Intraspecific and interspecific genetic distances of mitochondrial genes in Cnaphalocrocis exigua (Ga) and C. medinalis (Cm) (%)**

| Sequence type | Compared group | Nucleotide | Amino acid |
|---------------|----------------|------------|------------|
|               |                |            |            |
|               |                | Ce-Cm1     | 0.0982     | 0.1365     |
|               |                | Ce-Cm2     | 0.0748     | 0.0969     |
|               |                | Ce-Cm3     | 0.0748     | 0.0669     |
|               |                | Cm1-Cm2    | 0.0043     | 0.0043     |
|               |                | Cm1-Cm3    | 0.0020     | 0.0020     |
|               |                | Cm2-Cm3    | 0.0020     | 0.0020     |
|               |                | Cox1       | 0.0669     | 0.0669     |
|               |                | Cox2       | 0.0669     | 0.0669     |
|               |                | Cox3       | 0.0669     | 0.0669     |
|               |                | Atp8       | 0.1973     | 0.1973     |
|               |                | Atp6       | 0.0969     | 0.0969     |
|               |                | Cox1       | 0.0669     | 0.0669     |
|               |                | Cox2       | 0.0669     | 0.0669     |
|               |                | Cox3       | 0.0669     | 0.0669     |
|               |                | Atp8       | 0.1973     | 0.1973     |
|               |                | Atp6       | 0.0969     | 0.0969     |
|               |                | Cox1       | 0.0669     | 0.0669     |
|               |                | Cox2       | 0.0669     | 0.0669     |
|               |                | Cox3       | 0.0669     | 0.0669     |

---

Ce indicates GenBank JF246082; Cm indicates GenBank JQ305693; Ga indicates GenBank JQ47917.
from 156 to 168 bp) (Supp Fig. 1 [online only]). The A+T content of atp8 gene of Cm1 (93.9%), Cm2 (93.2%), and Cm3 (93.2%) was higher than C. exigua (90.4%) and the average level (91.0%) of 38 Pyraloidea species. Furthermore, Ile and Asn was the most used amino acid in atp8 of C. exigua and C. medinalis mitogenomes, respectively. Interestingly, the amino acid composition analysis show that the cysteine (Cys, C) and Glutamic acid (Glu, E) were specifically existed in atp8 of family Crambidae (Supp Fig. 1 [online only]).

As an inner membrane polypeptide of the F0 component, atp8 is responsible for the correct assembly of ATP synthase holoenzyme (Tzagoloff et al. 2004). Previous studies had found that genetic diversity of atp8 is associated with high-altitude adaptation in yak (Wang et al. 2017). At present, it is unclear whether the atp8 gene plays a key role in insect migration habits, and different individuals of C. medinalis and C. exigua have different energy production capabilities. Given that the atp8 gene plays an important role in energy metabolism, we believe that further investigation and research are still needed in this regard.

The tRNA Genes
In general, some tRNA genes showed highly conservative properties between these two species and between different C. medinalis individuals, while some tRNA genes showed certain variability. For instance, the trnM and trnW genes were not only conserved in different individuals of C. medinalis but also conserved in both leaffolders. Of the three individuals of C. medinalis, only 10 tRNA genes were identical and the conservative property of 22 tRNAs in the Cm2–Cm3 pair was higher than that of other pairs, which corresponds to Cm2 and Cm3 having 19 identical tRNA genes (Supp Table 3 [online only]). There were 11 and 12 differential tRNAs with base variation between Cm1 and Cm2, Cm1, and Cm3, respectively, and the differential bases were mainly found in the loop structures of the dihydrooridine (DHU) arm and the ΨC arm of tRNA. This also indicated that the variation of tRNA in the same species was mainly derived from the variation of the bases on the loop. The variation of tRNA in C. medinalis also indicated that there may be a large intraspecific genetic divergence in the natural populations of this species.

It was worth mentioning that the anticodon of trnS1 (AGN) were ‘TCT’ and ‘GCT’ in C. exigua and C. medinalis mitogenomes, respectively. According to the known mitogenomes, most pyralids contained a ‘GCT’ anticodon of trnS1, except for Paracymoriza distinctalis (TCT) and Eudonia angustea (ACT) (Timmermans et al. 2014, Ye and You 2016, Yang et al. 2018, Zhu et al. 2018). Furthermore, 11 other known pyralids had the same trnS1 gene sequence with that of C. medinalis and 8 of them belonged to subfamily Spilomelinae of Crambidae, thus implying the trnS1 gene may be conserved in more unreported moth species.

Overlap and Intergenic Regions
The overlaps of C. exigua and C. medinalis were in seven regions (24 bp in total) and six regions (22 bp in Cm1, 23 bp in Cm2 and Cm3), respectively (Table 1). In both leaffolders, the largest overlap was between trnW and trnC genes and contained a same motif ‘AAGCCTTA,’ the second large overlap was between atp8 and atp6 genes and shared a common motif ‘ATGATAA’.

The mitogenomes of C. exigua contained 132 bp of intergenic spacer sequences, whereas Cm1, Cm2, and Cm3 had 216 bp, 223 bp, and 239 bp of intergenic spacer sequences, respectively.
There were four major intergenic spacers in *C. exigua* and *C. medinalis* mitogenomes (Fig. 3). The spacer between *trnS2* and *nad1* contained a conserved motif ‘ATACTAW,’ which was consistent with other Pyralids (Yang et al. 2018, Zhu et al. 2018). Interestingly, spacers between *trnY* and *cox1* genes, *nad3* and *trnA* genes were much longer in *C. medinalis* than in *C. exigua*, which were featured by AT repeats (Fig. 3). It was worth mentioning that this feature rarely appeared in the superfamily Pyraloidea, especially the former only found in the *C. medinalis*.

### The A+T-rich Region

The A+T-rich region of *C. medinalis* also had a slightly higher A+T content than that of *C. exigua* (95.9 vs 94.4%) (Supp Table 2 [online only]). There was a conserved structure that consisted of the motif ‘ATAGW’ and a poly-T downstream of the *rrnS* gene in most known lepidopteran mitogenomes (Chai et al. 2012, Wan et al. 2013, Zhu et al. 2018). However, we found another style in the A+T-rich region of *C. exigua* at this site: the tetranucleotide ‘CTAG’ followed by poly-T sets. A different pattern occurred in *C. medinalis*, where the sequence ‘ATAG’ followed by a poly-T stretch (Fig. 4). Furthermore, a microsatellite-like (AT)₉ was observed in *C. exigua*, while a microsatellite-like (AT)₉ region preceded by motif ‘ATTTA’ was found in the 3’ end of *C. medinalis* (Cm2, Cm3) control region. A former study on 187 specimens of *C. medinalis* has found that the A+T-rich region of *C. medinalis* was 339–348 bp in length, 95.1–96.0% in A+T content, and characterized by a conserved motif ‘ATAG; a poly-T stretch (10–16 bp), a microsatellite-like AT repeat (10–14 bp), and a 5-bp long-motif ‘ATTTA’ (Wan et al. 2011). Since *C. exigua* and *C. medinalis* show distinctions in the structure of A+T-rich region, the function of these conserved elements needs to be further studied.

As a whole, it was important to note that these two species of *Cnaphalocrocis* were very close morphologically and hard to distinguish in the field. Therefore, it was surprising to find out that their mitochondrial genomes were so divergent.
Phylogenetic Analysis

Phylogenetic trees of Pyraloidea were reconstructed based on 13 PCGs of mitogenomes from 38 species with two species of Tortricoida as outgroup. The results strongly supported the close sister relationship between C. exigua and C. medinalis, and both of them belonged to the subfamily Spilomelinae (Crambidae). To some extent, this results also provided an evidence to explain why these two major rice leaffolders were not only similar in morphological feature but also similar in damage characteristics of rice.

Both BI and ML phylogenetic analyses strongly supported the division of the superfamly Pyraloidea into two families (Fig. 5), Crambidae and Pyralidae. In the family Pyralidae, relationships between the four subfamilies received strongly supported and were consistent with previous research results (Regier et al. 2012, Yang et al. 2018, Zhu et al. 2018). In the family Crambidae, the sister subfamilies Pyraustinae and Spilomelinae was well supported. However, there are still some doubts about the genetic relationship of other subfamilies in this families. Considering that some subfamilies do not have any species with mitogenome being sequenced, and some subfamilies have only sequenced one or two species, we suggest that in order to improve or even solve these problems, more mitochondrial data are needed.

Supplementary Data

Supplementary data are available at Journal of Insect Science online.

Acknowledgments

We thank Professor Ying-Hong Liu and Miss Qiu-ju Liao (College of Plant Protection, Southwest University, Chongqing, China) for providing the rice leaffolder C. exigua. We also thank Professor Xi-Cui Du (College of Plant Protection, Southwest University, Chongqing, China) for insightful comments and suggestions on the early version of manuscripts. This research was supported by Fundamental Research Funds for the Central Universities (XDJK2017C068, XDJK2014C097) and Chongqing Postdoctoral Science Special Foundation (Xm2014047).

Author Contributions

XR, KZ, and HL: conceived and designed the experiments. KZ, XR, ZL, GL, and LL: performed the experiments. KZ and XR: carried out the data analysis and drafted the manuscript.

References Cited

Babbucci, M., A. Basso, A. Scupola, T. Patarnello, and E. Negrisolo. 2014. Is it an ant or a butterfly? Convergent evolution in the mitochondrial gene order of Hymenoptera and Lepidoptera. Genome Biol. Evol. 6: 3326–3343.

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

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

Chai, H. N., Y. Z. Du, and B. P. Zhai. 2012. Characterization of the complete mitochondrial genomes of Cnaphalocrocis medinalis and Chilo suppressalis (Lepidoptera: Pyralidae). Int. J. Biol. Sci. 8: 561–579.

Dai, L. S., X. D. Zhou, S. Kausar, M. N. Abbas, L. Wu, and H. L. Zhou. 2018. Mitochondrial genome of Diabrotica undecimpunctata (Saunders) (Lepidoptera: Pyraloidea) and implications for its phylogeny. Int. J. Biol. Macromol. 108: 981–989.

Feng, B., Y. Yin, C. H. Peng, X. M. Peng, H. P. Jiang, R. Guo, and Y. J. Du. 2017. Morphological characteristics of Cnaphalocrocis exigua in comparison with C. medinalis (Lepidoptera: Pyralidae). Acta entomologica Sinica. 60: 95–103.

Gao, Y. B., X. Chen, Y. X. Bao, and R. M. Yang. 2008. Dynamic analysis on the migration of the rice leaf roller Cnaphalocrocis medinalis (Lepidoptera: Pyralidae) by Doppler Insect Monitoring Radar and numerical simulation. Acta ecologica Sinica 37: 487–496.

Harrison, R. G. 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4: 6–11.

He, S. L., Y. Zou, L. F. Zhang, W. Q. Ma, X. Y. Zhang, and B. S. Yue. 2015. The complete mitochondrial genome of the beet webworm, Spoldaea rerumaulis (Lepidoptera: Crambidae) and its phylogenetic implications. PloS One. 10: e0129355.

Huang, X. F., X. X. Zhang, and B. P. Zhai. 2010. Effect of copulation on flight capacity and remigration capacity of Cnaphalocrocis medinalis (Gueneé). J. Nanjing Agricult. Univ. 33: 23–28.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 17: 754–755.

Khan, Z. R., A. T. Barrion, J. A. Litsinger, N. P. Castilla, and R. C. Joshi. 1988. A bibliography of rice leaffolders (Lepidoptera: Pyralidae). Insect Sci. Appl. 9: 129–174.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33: 1870–1874.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentín, I. M. Wallace, A. Wilm, R. Lopez, et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23: 2947–2948.

Liao, Q. J., Y. J. Yang, J. Wang, X. P. Pang, C. M. Xu, C. L. Peng, Z. X. Lu, and Y. H. Liu. 2017. Temperature-dependent development and reproduction of rice leaffolder, Marasmius exigua (Lepidoptera: Pyralidae). Plos One. 12: e0187972.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25: 1451–1452.

Lowe, T. M., and P. P. Chan. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. 44: W54–W57.

Nieuerken, E. J., L. Kaila, I. J. Kitching, N. P. Kristensen, D. C. Lee, J. Minet, C. Mitter, M. Mutanen, J. C. Regier, T. J. Simonsen, et al. 2011. Order Lepidoptera Linnaeus, 1758, pp. 212–221. In: Z.-Q. Zhang (ed.), Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness. Magnolia Press, Auckland, New Zealand.

Padmavathi, C., G. Katti, A. P. Padmakumari, S. R. Veleti, and L. V. S. Rao. 2013. The effect of leaffolder Cnaphalocrocis medinalis (Gueneé) [Lepidoptera: Pyralidae] injury on the plant physiology and yield loss in rice. J. Appl. Entomol. 137: 249–256.

Pan, X. X. 1984. Research on the occurrence regularity of Marasmius exigua. Entomol. Knowl. 3: 003.

Pan, X. X. 1985. Studies on migratory law and control strategies of rice leaf roller (Cnaphalocrocis medinalis Gueneé) in Sichuan Basin. J. Nanjing Agricult. Univ. (China) 3: 32–40.

Park, H., C. Jumrae, P. Chanyangg, K. Kwanggo, G. Hyoungwan, and L. Sangsuei. 2010. The occurrence of rice leaf-folder, Cnaphalocrocis medinalis (Lepidoptera: Crambidae) in Suwon and its responses to insecticides. Korean J. Appl. Entomol. 49: 219–226.

Park, H., J. Ahn, and C. Park. 2014. Temperature-dependent development of Cnaphalocrocis medinalis Gueneé (Lepidoptera: Pyralidae) and their validation in semi-field condition. J. Asia-Pac. Entomol. 17: 83–91.

Perna, N. T., and T. D. Kocher. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41: 353–358.

Regier, J. C., C. Mitter, M. A. Solis, J. E. Hayden, B. Landry, M. Nuss, T. J. Simonsen, S. H. Yen, A. Zwick, and M. P. Cummings. 2012. A molecular phylogeny for the pyraloid moths (Lepidoptera: Pyraloidea) and its implications for higher-level classification. Syst. Entomol. 37: 635–656.

Scott, G. R., P. M. Schulte, S. Egginton, A. L. Scott, J. G. Richards, and W. K. Milom. 2011. Molecular evolution of cytochrome C oxidase underlies high-altitude adaptation in the bar-headed goose. Mol. Biol. Evol. 28: 351–363.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene
sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87: 651–701.

Song, J. R., P. F. Li, and P. You. 2016. The complete mitochondrial genome of *Pseudargyria interruptella* (Lepidoptera: Crambidae). Mitochondrial DNA A. DNA Mapp. Seq. Anal. 27: 3899–3900.

Timmermans, M. J., D. C. Lees, and T. J. Simonsen. 2014. Towards a mitogenomic phylogeny of Lepidoptera. Mol. Phylogen. Evol. 79: 169–178.

Tzagoloff, A., A. Barrientos, W. Neupert, and J. M. Herrmann. 2004. Atp10p assists assembly of Atp6p into the F0 unit of the yeast mitochondrial ATPase. J. Biol. Chem. 279: 19775–19780.

Wan, X., M. J. Kim, and I. Kim. 2013. Description of new mitochondrial genomes (*Spodoptera litura*, Noctuoidea and *Cnaphalocrocis medinalis*, Pyraloidea) and phylogenetic reconstruction of Lepidoptera with the comment on optimization schemes. Mol. Biol. Rep. 40: 6333–6349.

Wang, F., X. Zhang, and B. Zhai. 2010. Flight and re-migration capacity of the rice leaffolder moth, *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae), inferred from the mitochondrial A+T-rich region and nuclear ITS2 sequences. Genet. Mol. Res. 10: 273–294.

Wang, J., Y. Shi, M. A. Elzo, S. Z. Dang, X. B. Jia, and S. J. Lai. 2017. Genetic diversity of ATP8 and ATP6 genes is associated with high-altitude adaptation in yak. Mitochondrial DNA A. DNA Mapp. Seq. Anal. 29: 1–9.

Ye, F., and P. You. 2016. The complete mitochondrial genome of *Paracymoriza distinctalis* (Lepidoptera: Crambidae). Mitochondrial DNA A. DNA Mapp. Seq. Anal. 27: 28–29.

Yin, Y. H., F. J. Qu, Z. W. Yang, X. Y. Zhang, and B. S. Yue. 2014. Structural characteristics and phylogenetic analysis of the mitochondrial genome of the rice leafroller, *Cnaphalocrocis medinalis* (Lepidoptera: Crambidae). Mol. Biol. Rep. 41: 1109–1116.

Zhang, H. L., and F. Ye. 2017. Comparative mitogenomic analyses of praying mantises (Dictyoptera, Mantodea): origin and evolution of unusual intergenic gaps. Int. J. Biol. Sci. 13: 367–382.

Zhang, K. J., L. Liu, X. Rong, G. H. Zhang, H. Liu, and Y. H. Liu. 2016. The complete mitochondrial genome of *Bactrocera diaphora* (Diptera: Tephritidae). Mitochondrial DNA A. DNA Mapp. Seq. Anal. 27: 4314–4315.

Zou, Y., W. Ma, L. Zhang, S. He, X. Zhang, and Z. Tao. 2016. The complete mitochondrial genome of the bean pod borer, *Maruca testulalis* (Lepidoptera: Crambidae: Spilomelinae). Mitochondrial DNA A. DNA Mapp. Seq. Anal. 27: 740–741.