

## The complete mitochondrial genome of *Orthaga achatina* (Lepidoptera: Pyralidae)

Bing Liu<sup>a</sup>, Hui Sun<sup>a</sup>, Qingbin Zhan<sup>b</sup> and Yunpeng Gai<sup>c</sup>

<sup>a</sup>Collaborative Innovation Center of Sustainable Forestry in Southern China, College of Forestry, Nanjing Forestry University, Nanjing, China;

<sup>b</sup>Nanjing Forest Police College, Nanjing, China; <sup>c</sup>Key Lab of Molecular Biology of Crop Pathogens and Insects, Ministry of Agriculture, Institute of Biotechnology, Zhejiang University, Hangzhou, China

### ABSTRACT

*Orthaga achatina* Butler is an important pest of camphor trees in Asia. The complete mitochondrial genome of *O. achatina* was sequenced in this study, which was 15,150 bp in size and comprised of 13 protein-coding genes, 22 transfer RNA genes, 2 ribosomal RNA genes, and a control region. Besides, we used a phylogenomic approach to infer evolutionary relationships of *O. achatina* and 23 Lepidoptera species based on 13 conserved protein sequences of the mitochondrial genome. Our results underline the potential importance of mitochondrial genomes in comparative genomic analyses of Lepidoptera species and provide a robust evolutionary insight across the tree of Lepidoptera insects.

### ARTICLE HISTORY

Received 15 September 2020  
Accepted 10 November 2020

### KEYWORDS

*Orthaga achatina*; complete mitochondrial genome; phylogenomic tree; Lepidoptera; *Cinnamomum camphora*



Pyralidae is the largest family in Lepidoptera, with more than 25,000 species in the world, some of which are pests of agricultural and forestry plants, such as *Orthaga* (Yang et al. 2020). *Orthaga achatina* Butler (Lepidoptera: Pyralidae) is the most serious pest of camphor trees (*Cinnamomum camphora*) in China, Korea, Japan, and Malaysia (Wu 2006). *Orthaga achatina* can also feed on other Lauraceae plants, such as *Lindera glauca* and *Cinnamomum cassia*, causing serious defoliation (Long et al. 2017). Mitochondrial genome has high mutation rate and low DNA recombination rate, which is of great significance in insect phylogenetic analysis (Gai et al. 2020; Yang et al. 2020). However, the mitochondrial genome of *O. achatina* has not been publicly reported. Therefore, we used *de novo* sequencing technology to sequence the whole mitochondrial genome of *O. achatina* to understand its mitogenomic background and genetic evolution relationship.

In the present study, samples of *O. achatina* were collected from camphor trees (N31.16568°, E120.62638°) in Suzhou, Jiangsu Province, China in July 2020. Some of these samples were immediately frozen at –80 °C for sequencing analysis and others were preserved in the Entomological Lab of Nanjing Forestry University, and their specimen code is 2020NJEM1855-1860. The genomic DNA was extracted from *O. achatina* using CTAB (cetyltrimethylammonium Ammonium Bromide) method (Huanca-Mamani et al. 2015). The quality and quantity of DNA sample were evaluated using an Agilent Bioanalyzer (Agilent Technologies, CA, USA) and the Qubit 3.0 Fluorometer (Life Technologies, CA, USA).

The sequencing libraries were sequenced on the Illumina HiSeq platform in a paired-end 2 × 150 bp mode (Illumina, San Diego, CA). The raw data was processed to remove adapters and low-quality bases using Trimmomatic v0.36 with default parameters (Bolger et al. 2014). *De novo* assembly was performed by SPAdes version 3.14 with the parameter ‘-k 21,33,55,77,99,127 –careful’ (Bankevich et al. 2012). The complete mitochondrial genomes were annotated by MITOS WebServer (<http://mitos.bioinf.uni-leipzig.de/index.py>) (Bernt et al. 2013) and submitted to NCBI GenBank (GenBank accession number: MT916176).

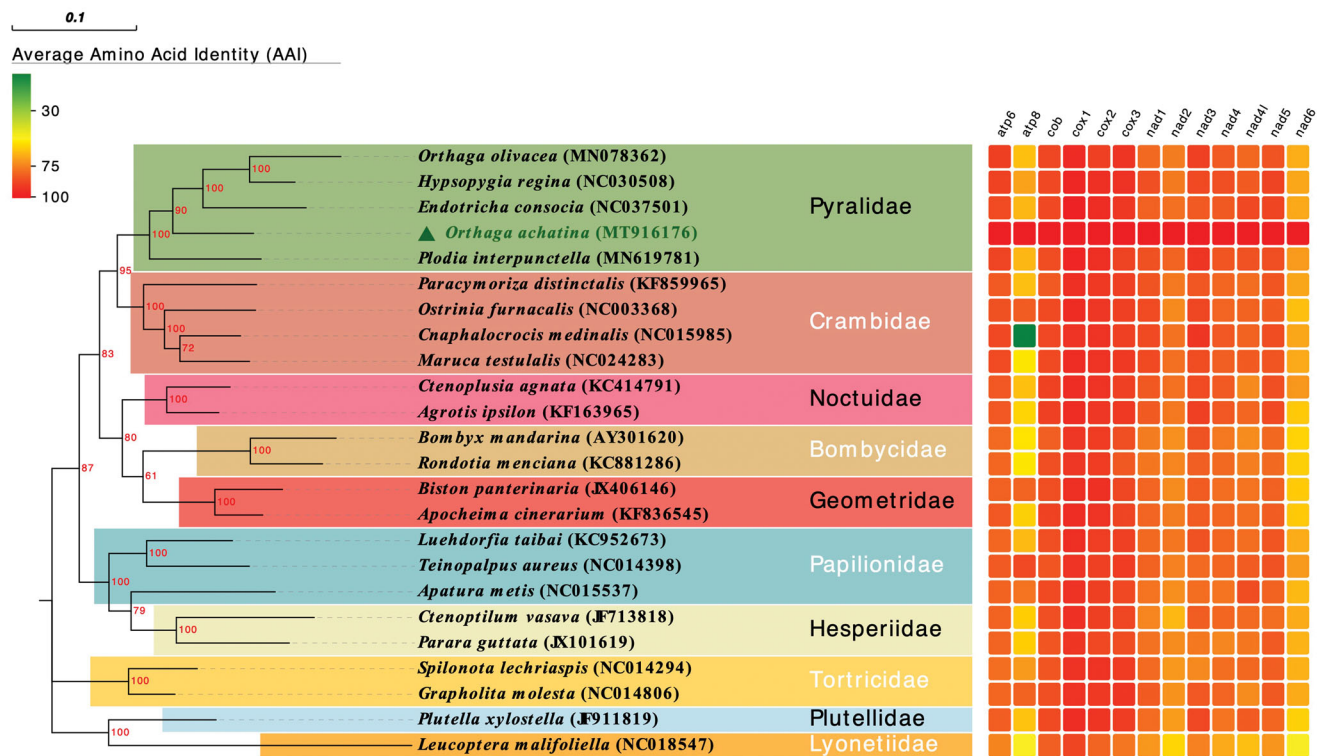
The mitochondrial genome of *O. achatina* is 15,150 bp, with a nucleotide composition of 38.9% A, 41.8% T, 11.4% C and 7.9% G. The mitochondrial genome of *O. achatina* comprised the entire set of 37 typical invertebrate mitochondrial genes consisting of 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), 2 ribosomal RNA genes (rRNAs), and a control region (D-loop). The majority coding strand encoded 23 genes (9 PCGs and 14 tRNAs), whereas the minority coding strand encoded 14 genes (4 PCGs, 8 tRNAs, and 2 rRNAs). The sequence and arrangement of the genes are highly conserved, indicating that they are similar to the typical characteristics of Lepidoptera mitochondrial genome (Wu et al. 2016, 2020; Liu et al. 2018). A total of 44 overlapping nucleotides were found in 6 loci, ranging from 2 to 25 bp, and 857 bp intergenic nucleotides in 22 locations, ranging from 4 to 297 bp in length.

All protein-coding genes (PCGs) are initiated with ATN as the starting codon except the *cox1*, which is no justification

**CONTACT** Yunpeng Gai  [gaiyunpeng@gmail.com](mailto:gaiyunpeng@gmail.com)  Key Lab of Molecular Biology of Crop Pathogens and Insects, Ministry of Agriculture, Institute of Biotechnology, Zhejiang University, Hangzhou, China

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Phylogenetic relationship of the nine Lepidoptera families, which were inferred from phylogenetic analysis of the 13 proteins by using the maximum likelihood method of MEGA 7.0. The number above in each node indicates the bootstrap support values with 1000 replicates.

for continued speculation about polynucleotide start codon similar to other Lepidoptera insects (Singh et al. 2017; Liu et al. 2018; Yang et al. 2020). Ten PCGs had canonical stop codons TAA or TAG, while three had incomplete termination codons single T (*cox3* and *atp6*) or TA (*nad4L*). There were 22 tRNA genes with a length between 63 and 70 bp. All tRNA genes exhibited a typical clover-leaf secondary structure, except for tRNA-Ser(AGN) lacking the dihydrouridine (DHU) arm, which is common in Lepidoptera insects (Garey and Wolstenholme 1989). The lengths of lrRNA and srRNA were 1362 bp and 780 bp, respectively. The control region was located between srRNA and tRNA-Met with a total length of 298 bp.

The mitochondrial genomes of 23 Lepidoptera species previously sequenced were downloaded from GenBank and compared with *O. achatina* in this study. To identify the orthogroups, the BLAST-based ortholog detector OrthoFinder v2.2.7 (Emms and Steven 2019) with default parameter values were used to identify ortholog among all the protein sequences of the 24 mitochondrial genomes. The phylogenetic relationship of *O. achatina* and 23 Lepidoptera species was inferred from phylogenetic analysis of the 13 protein coding genes using MEGA7.0 software with maximum likelihood method and 1000 replicate sets on bootstrap analysis. The amino acid identity (AAI) of the 13 protein coding genes of each Lepidoptera species and *O. achatina* were calculated by NCBI BLASTP. The phylogenetic tree and AAI heatmap of each protein was visualized using EVOLVIEW version 2 (<https://evolgenius.info/evolview-v2>) (He et al. 2016). Phylogenetic analyses showed similar relationships among sampled families as shown in Yang et al. (2020). Each clade

showed a monophyletic cluster and the following clades were highly supported (Figure 1): (1) Pyralidae + Crambidae; and (2) (Pyralidae + Crambidae) + (Noctuidae + (Bombycidae + Geometridae)). We also found that *O. achatina* strains had the closest relationship with the genus *Hypsopygia* and *Endotricha*, which were located in a clade in the clade of Pyralidae. This study can provide a useful resource for the genetic evolution of *O. achatina* and underline the potential importance of mitochondrial genomes in comparative genomic analyses of Lepidoptera species.

## Acknowledgements

Authors are grateful to Dr. Bin Li and Dr. Lingjuan Han for their valuable advice.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This study was supported by the research funding for Postgraduate Research & Practice Innovation Program of Jiangsu Province [KYCX19\_1076] and the research funding for Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions.

## Data availability statement

The data that support the findings of this study are openly available in the National Center for Biotechnology Information under the BioProject

accession number PRJNA669357 (<https://www.ncbi.nlm.nih.gov/bioproject/669357>). The figure is available on the figshare repository: <https://doi.org/10.6084/m9.figshare.13154066>. The complete mitochondrial genome and its annotation file are available in the Zenodo repository: <http://doi.org/10.5281/zenodo.4159192>.

## References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsche G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol.* 69(2): 313–319.
- Bolger AM, Marc L, Bjoern U. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30(15):2114–2120.
- Emms DM, Steven K. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20(1):14.
- Gai Y, Pan R, Peng X. 2020. A phylogenomic tree of fungi: evolutionary relationships among *Calonectria ilicicola* and 586 fungal mitochondrial genomes. *Mitochondrial DNA Part B.* 5(2):1709–1711.
- Garey JR, Wolstenholme DR. 1989. Platyhelminth mitochondrial DNA: Evidence for early evolutionary origin of a tRNA ser AGN that contains a dihydrouridine arm replacement loop, and of serine-specifying AGA and AGG codons. *J Mol Evol.* 28(5):374–387.
- He Z, Zhang H, Gao S, Lercher MJ, Chen WH, Hu S. 2016. Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.* 44(W1): W236–W241.
- Huanca-Mamani W, Rivera-Cabello D, Maita-Maita J. 2015. A simple, fast, and inexpensive CTAB-PVP-Silica based method for genomic DNA isolation from single, small insect larvae and pupae. *Genet Mol Res.* 14(3):8000–8001.
- Liu Q, Jiang X, Hou X, Yang H, Chen W. 2018. The mitochondrial genome of *Ephesia elutella* (Insecta: Lepidoptera: Pyralidae). *Mitochondrial DNA Part B.* 3(1):189–190.
- Long YB, Zhao DY, Qin CS. 2017. Occurrence status and control of *O. achatina*. *Forest Environ Sci.* 33:107–110.
- Singh D, Kabiraj D, Sharma P, Chetia H, Mosahari PV, Neog K, Bora U. 2017. The mitochondrial genome of Muga silkworm (*Antheraea assamensis*) and its comparative analysis with other Lepidopteran insects. *PLOS One.* 12(11):e0188077
- Wu XF. 2006. The Occurrence and Integrated Control of *O. achatina*. *Forest Sci Technol.* 31(4):25–27.
- Wu Y-P, Lu J-J, Yang J, Fan R-J. 2020. Complete mitochondrial genome of *Dioryctria yiai* (Lepidoptera: Pyralidae). *Mitochondrial DNA Part B.* 5(1): 1062–1064.
- Wu Y-P, Zhao J-L, Su T-J, Luo A-R, Zhu C-D. 2016. The complete mitochondrial genome of *Choristoneura longicellana* (Lepidoptera: Tortricidae) and phylogenetic analysis of Lepidoptera. *Gene.* 591(1): 161–176.
- Yang L, Dai J, Gao Q, Yuan G, Liu J, Sun Y, Sun Y, Wang L, Qian C, Zhu B, et al. 2020. Characterization of the complete mitochondrial genome of *Orthaga olivacea* Warre (Lepidoptera Pyralidae) and comparison with other Lepidopteran insects. *PLOS One.* 15(3):e0227831.