

Next-generation sequencing yields the complete mitochondrial genome of *Rhodoprasina callantha* (Lepidoptera: Sphingidae) and its evolutionary status

Yin-Feng Meng^a, Guo-Tao Lv^b, Yi-Xin Huang^c, Xu Wang^d and Yong-Ling Wu^a

^aCollege of Biology Pharmacy and Food Engineering, Shangluo University, Shangluo, China; ^bDatong Municipal Bureau of Agriculture and Rural Affairs, Datong, China; ^cCollaborative Innovation Center of Recovery and Reconstruction of Degraded Ecosystem in Wanjiang Basin Co-founded by Anhui Province and Ministry of Education, School of Ecology and Environment, Anhui Normal University, Wuhu, China; ^dAnhui Provincial Key Laboratory of the Conservation and Exploitation of Biological Resources, College of Life Sciences, Anhui Normal University, Wuhu, China

ABSTRACT

In this study, we sequenced and analyzed the complete mitochondrial genome of *Rhodoprasina callantha* Jordan, 1929. The complete mitochondrial genome sequence of *R. callantha* was 15,343 bp in size and encoded 13 protein-coding genes (PCGs), two ribosomal RNA genes (rRNA), 22 transfer RNA genes (tRNA) and one control region (CR). The nucleotide composition of mitogenome was highly biased to A and T. Most protein-coding genes (PCGs) initiate with the standard start codon of ATN and terminate with the typical stop codon TAA/TAG. The phylogeny of Sphingidae based on nucleotide sequences of PCGs recovered the monophyly of subfamilies of Sphingidae with high support values. Langinae was the first subfamily diverged in Sphingidae, which was in accordance with previous study. *R. callantha* was the member of subfamily Smerinthinae, and closely related to the genus *Marumba*.

ARTICLE HISTORY

Received 11 April 2022
Accepted 23 July 2022

KEYWORDS

Mitogenome; hawkmoth;
Bombycoidea;
Yunnan Province

Hawkmoths belong to the family Sphingidae, order Lepidoptera, most of which are well-known flower visitors and significant pollinators, and most adults have well-developed proboscises (Krpac et al. 2019). When a hawkmoth drinks from a flower, the proboscis picks up pollen, which can be spread to flowers farther than 29 km away as the moth travels along its feeding route. Their larvae have cylindrical, medium to large sized bodies, generally with a single caudate scoli, and are known as significant agricultural pests (Nagamine et al. 2019). Among them, *Rhodoprasina callantha* Jordan, 1929 is an endemic species in Asia. It is distributed from northeastern India and Nepal across Bhutan, northern Myanmar/Burma, southwestern China (Yunnan), northern Thailand, northern Laos to northern and central-southern Vietnam. And it is a member of 'Cypa group' of which the relationship is still unclear. Up to date, the mitogenome of *R. callantha* remains unknown and the related molecular research of *R. callantha* is still sparse. Therefore, we documented the first complete mitogenome of *R. callantha* and reconstructed the phylogenetic tree of Sphingidae including *R. callantha*, in order to provide comprehensive data for this species and also to elucidate its phylogenetic position within the family Sphingidae.

The specimens of *Rhodoprasina callantha* Jordan, 1929 (Lepidoptera: Sphingidae) were collected from Menghai, Yunnan, China (21°41'53"N, 100°03'24"E) on 17 April 2021, and were deposited in the Entomological Museum, College

of Life Sciences, Anhui Normal University (<https://www.ahnu.edu.cn/>), YX, Huang, huangyx@ahnu.edu.cn) under the accession no. YN20210417. Animal sampling was performed according to the protocols approved by the Institutional Animal Care and Use Committee of Anhui Normal University (Grant number AHNU-ET2021032). Mitochondrial DNA was extracted from the leg of each male adult specimen. After cluster generation, the library preparations were sequenced on an Illumina platform and 150 bp paired-end reads were generated. The library was prepared from DNA of this species only. Raw data were retrieved and quantified by FastQC (Andrews 2020). Totally, 10 Gb data was sequenced and 3 Gb was randomly generated to assemble with mean coverage of 2000X. NovoPlasty and mitoz was used to assemble and annotate the mitogenome (Dierckxsens et al. 2017; Meng et al. 2019). Protein-coding genes (PCGs) were identified according to open reading frames of *Theretra japonica* (MG655620). The tRNAs were indicated by Mitoz Web Server (Bernt et al. 2013). Geneious was used to verify the results of assembly and annotation (Kearse et al. 2012).

The genome sequence data that support the findings of this study are openly available in the GenBank of NCBI (<https://www.ncbi.nlm.nih.gov/>) under the accession NO. MZ343573. The length of the mitogenome of *R. callantha* was 15,343 bp. The mitogenome was circular and double-stranded. It was composed of 37 genes, including 13 PCGs (protein-coding genes), 22 tRNAs, two rRNAs and one control



Figure 1. Maximum likelihood tree of Sphingidae based on the nucleotide sequences of 13 protein-coding genes dataset. The ultrafast bootstrap approximation (UFboot) value was shown on the nodes (values below 75 were not shown).

region, which were usually found in animal mitogenomes (Cameron 2014). The gene order is identical to the referenced sphingid *Ampelophaga rubiginosa* Bremer & Grey, 1853 and also consistent with other taxon in Sphingidae. The majority strand (J-strand) encoded 23 genes (9 PCGs and 14 tRNAs), and the minority strand (N-strand) encoded 14 genes (4 PCGs, 8 tRNAs and 2 rRNAs). The overall base composition of the mitogenome had a high AT content of 80.8%. Among the PCGs of *R. callantha*, 9 PCGs (*NAD2*, *COX1*, *COX2*, *ATP8*, *ATP6*, *COX3*, *NAD3*, *NAD6*, *CYTB*) were encoded by the J strand, while the rest four were encoded on the N strand. Most PCGs started with ATG or ATT and stop with TAA. However, the *COX1* gene started with CGA and three genes (*COX1*, *COX2*, *NAD5*) used the incomplete stop codon T. The usage of start and stop codon were similar to most insect mitogenomes (Crozier and Crozier 1993; Korkmaz et al. 2015).

All 22 tRNA genes that were usually found in the mitogenomes of insects, resided in that of *R. callantha*. The length of tRNA genes ranged from 62 bp (*trnS1*) to 71 bp (*trnK* and *trnV*), and A + T content ranged from 73.1% (*trnL2*) to 92.4% (*trnE*). The length of two rRNA genes were 1293 bp (*rrnL*) and 770 bp (*rrnS*), and the A + T content were 83.7% (*rrnL*) and 85.3% (*rrnS*), respectively. The A + T-rich region of *R. callantha* was 340 bp and located between the *rrnS* and *trnM*. The A + T content of this region was 93.2%.

To validate the phylogenetic position of *R. callantha*, we selected the mitochondrial DNA sequences of 36 Lepidoptera species (34 Sphingidae as ingroup and two as outgroup). *Biston panterinaria* (Bremer et Grey, 1853) (GenBank accession number KU325533) and *Phthonandria atrilineata* (Butler, 1881) (GenBank accession number EU569764) from Geometridae which were confirmed as the most closely related to the superfamily Bombycoidea, were selected as outgroup (Yang et al. 2009). Every nucleotide sequences of PCGs were aligned by MAFFT (Katoh et al. 2005). Then the aligned sequences were concatenated into a dataset. The alignment of

Sphingidae was deposited in the Science Data Bank, DOI 10.57760/sciencedb.01748. We analyzed the concatenated dataset of PCGs using the maximum-likelihood (ML) on the W-IQ-Tree web server method to reconstruct the phylogenetic relationship of *R. callantha* with other species of Sphingidae under the best-fit model: GTR + F + I + G4 chosen according to BIC selected by W-IQ-Tree web server (Trifinopoulos et al. 2016). The results recovered the monophyly of subfamilies of Sphingidae with high supporting values (Figure 1). Langiinae was the first subfamily diverged in Sphingidae, which was in accordance with the results of Wang et al. (2021). *R. callantha* was the member of subfamily Smerinthinae, and closely related to the genus *Marumba* which was agreed with the morphological features (Kitching 2022).

Authors contribution

Yin-Feng Meng, Guo-Tao Lv, Xu Wang, Yi-Xin Huang and Yong-Ling Wu were involved in the conception and design, analysis and interpretation of the data; Xu Wang and Yi-Xin Huang collected insect; Yin-Feng Meng and Yong-Ling Wu drafted the paper and revised it critically for intellectual content. Yong-Ling Wu gave final approval of the version to be published. All authors agree to be accountable for all aspects of the work.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

The project was supported by grants from the Specialized Research Fund of Education Department of Shaanxi Province [21JK0618], Project of Science and Technology Special of Shangluo [2021-Z-0048], the Scientific Research Foundation of Shangluo University [20SKY009], the Natural

Science Fund of Anhui Province [1908085QC93] and Natural Science Foundation of Universities of Anhui Province [KJ2020A0094].

Data availability statement

The data that support the findings of this study are openly available in GenBank at <https://www.ncbi.nlm.nih.gov/genbank/>, reference number MZ343573. The associated BioProject, Bio-Sample numbers, and SRA are PRJNA735310, SAMN19575429, and SRR14740110, respectively.

References

- Andrews S. 2020. FastQC: a quality control tool for high throughput sequence data. [accessed 2020 July 10]. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation molecular phylogenetics and evolution. *Mol Phylogenet Evol.* 69(2):313–319.
- Cameron SL. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu Rev Entomol.* 59:95–117.
- Crozier RH, Crozier YC. 1993. The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. *Genetics.* 133(1):97–117.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45:e18.
- Katoh K, Kuma KI, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33(2): 511–518.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28(12): 1647–1649.
- Kitching IJ. 2022. Sphingidae taxonomic inventory. [accessed 2022 May 12]. <http://sphingidae.myspecies.info>
- Korkmaz EM, Doğan Ö, Budak M, Başbüyük HH. 2015. Two nearly complete mitogenomes of wheat stem borers, *Cephus pygmeus* (L.) and *Cephus sareptanus* Dognar-Zapolskij (Hymenoptera: Cephidae): an unusual elongation of *rnrS* gene. *Gene.* 558(2):254–264.
- Krpač V, Kućinić M, Abdija X, Beadini N, Krpač M, Darcemont C, Lemonnier – Darcemont M, Krsteska V, Lazarevska S, Černila M. 2019. Contribution to the fauna of the hawk moth family (Lepidoptera, Sphingidae) in the Republic of North Macedonia. *Nat Croat.* 28(1): 107–130.
- Meng GL, Li YY, Yang CT, Liu SL. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Res.* 47(11):e63–e63.
- Nagamine K, Hojoh K, Nagata S, Shintani Y. 2019. Rearing *Theretra oldenlandiae* (Lepidoptera: Sphingidae) larvae on an artificial diet. *J Insect Sci.* 19(3):10.
- Trifinopoulos J, Nguyen LT, Haeseler AV, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44(W1):W232–235. doi:<https://doi.org/10.1371/journal.pone.0005719>.
- Wang X, Zhang H, Kitching I, Xu ZB, Huang YX. 2021. First mitogenome of subfamily Langiinae (Lepidoptera: Sphingidae) with its phylogenetic implications. *Gene.* 789:145667.
- Yang L, Wei ZJ, Hong GY, Jiang ST, Wen LP. 2009. The complete nucleotide sequence of the mitochondrial genome of *Phthonandria atrilineata* Lepidoptera: Geometridae). *Mol Biol Rep.* 36(6):1441–1449.