MITOCHONDRIAL DNA PART B: RESOURCES
2025, VOL. 10, NO. 6, 528–531
https://doi.org/10.1080/23802359.2025.2509807

Taylor & Francis
Taylor & Francis Group

MITOGENOME REPORT

OPEN ACCESS Check for updates

Complete mitochondrial genome data and phylogenetic analysis of the Plain Banded Awl, *Hasora vitta* (Lepidoptera: Hesperiidae: Coeliadinae) from Malaysia

Marylin Miga[a], Aqilah Awg Abdul Rahman[b,c], Sivachandran Parimannan[d], Heera Rajandas[d], Frankie Thomas Sitam[e], Lili Tokiman[f], Jai Kemalok[b,c], Mohd Shahir Shamsir[a] and Faezah Mohd Salleh[a,g]

[a]Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Johor Bahru, Malaysia; [b]Department of Technology and Natural Resources, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia (Pagoh Campus), Muar, Malaysia; [c]Faculty of Applied Sciences and Technology (FAST), Centre of Research for Sustainable Uses of Natural Resources (SUNR), Universiti Tun Hussein Onn Malaysia, Pagoh Higher Education Hub, Muar, Malaysia; [d]Faculty of Applied Sciences, Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), AIMST University, Bedong, Malaysia; [e]Department of Wildlife and National Parks (PERHILITAN), National Wildlife Forensic Laboratory (NWFL), Ex-Situ Conservation Division, Kuala Lumpur, Malaysia; [f]Johor National Parks Corporation, Kota Iskandar, Iskandar Puteri, Malaysia; [g]Investigative and Forensic Sciences Research Group, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

ABSTRACT

We report the complete mitogenome of *Hasora vitta* (Plain Banded Awl) sampled from Malaysia. The mitogenome is 15,289 bp long, comprising of 13 protein-coding genes (PCGs), 22 transfer RNAs (tRNAs), two ribosomal RNAs (rRNAs), and a control region. All PCGs were initiated by the typical ATN codon, except for COX1 with a CGA start codon. Two PCGs (COX1 and COX2) were terminated with an incomplete stop codon T. Phylogenetic analysis highly supported the placement of *Hasora vitta* from Malaysia within Coeliadinae and is clustered with two *Hasora vitta* samples from China.

ARTICLE HISTORY
Received 12 September 2024
Accepted 19 May 2025

KEYWORDS
Skipper; complete mitochondrial genome; Endau-Rompin; Johor National Park

Introduction

Hasora vitta (Butler, 1870) or commonly known as the Plain Banded Awl skipper is a species from the subfamily Coeliadinae (Family Hesperiidae) and can be found across the Southeast Asian region (Sun et al. 2021). This species can be identified based on the yellow hyaline spot present on the forewing at space 6 and they are known as policemen skipper due to the shape of their labial palpi and fast-flying ability (Figure 1) (Chiba 2009). To date, two complete mitogenome sequences of *Hasora vitta* were deposited in GenBank which originate from China, but none from Malaysia and this shows that the mitogenomic data for Hesperiidae species are still underrepresented for Malaysia. Therefore, in this study, we have sequenced and analyzed the complete mitochondrial genome of *Hasora vitta* from Malaysia which can be used to support the phylogenetic relationships analysis of this genus across a broader population through mitochondrial genomic research.

Materials and methods

The specimen was collected from Endau-Rompin Johor National Park, Malaysia specifically at the Sungai Kawal area (2.03 N 102.49 E) on 9 May 2023, and the hind legs were stored in a 95% ethanol for molecular analysis. The specimen was curated and deposited in Universiti Teknologi Malaysia (UTM) (https://www.utm.my/, Dr Faezah Mohd Salleh, faezah@utm.my) with the voucher ID S14 (Figure 1). The genomic DNA was extracted from the hind legs using the Qiagen Blood and Tissue Kit (Qiagen, Valencia, CA) following the manufacturer's instructions. As for the library preparation, NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® was used before sequencing by Illumina NovaSeq 6000 system (PE150) similar to Miga et al. (2023). The raw reads obtained were pre-processed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and then trimmed using AdapterRemoval (Schubert et al., 2016) to remove sequencing adapters and low quality reads. The mitogenome was assembled using NOVOPlasty v.4.2 (Dierckxsens et al., 2017) and was ran through a PALEOMIX BAM pipeline (Schubert et al., 2014) with default parameters, to remove reads shorter than 15 bp after trimming (Miga et al., 2023). Then the mitochondrial contigs were annotated using MITOS2 web server (Bernt et al., 2013) available in the UseGalaxy platform (Afgan et al. 2022). To improve the annotation, the predicted proteins were further verified using ORFFinder (https://www.ncbi.nlm.nih.gov/orffinder/), by comparing the alignments of each annotation produced by Mitos2 and ORFFinder. Tablet (Milne et al., 2010) was also used to

CONTACT Faezah Mohd Salleh faezah@utm.my Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Johor Bahru, 81310 Johor, Malaysia

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2025.2509807.

visualize the assembled mitogenome for indels and sequence coverage. A circular mitogenome map of the sequenced *Hasora vitta* in this work was generated using Proksee web server (https://proksee.ca/), which is the updated version of the CGView web server (Grant et al., 2023) (Figure 2).

To investigate the phylogenetic relationships within this subfamily, 13 concatenated protein-coding genes (PCGs) PCGs from 28 Hesperiidae mitogenome sequences were used to reconstruct the maximum-likelihood (ML) tree (Figure 3).

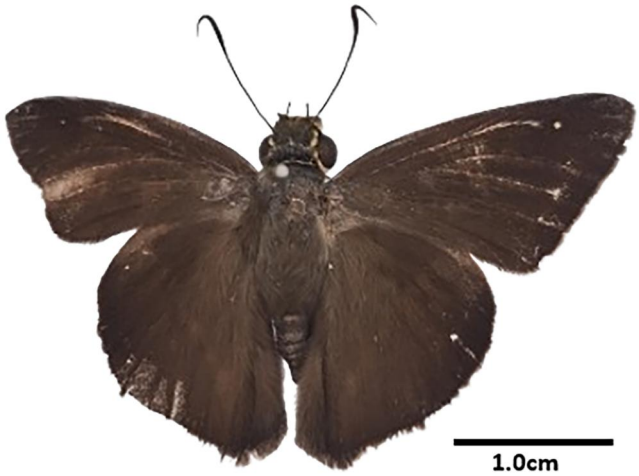

Figure 1. A reference image of the sequenced *Hasora vitta* collected from Endau-Rompin Johor National Park, Malaysia (collected and photographed by Marylin Miga from Universiti Technologi Malaysia, Johor, Malaysia).

The outgroup used in the analysis was *Papilio helenus* (KM244656) (Papilionidae) and *Macrosoma conifera* (MT852025) (Hedylidae). Phylogenetic analysis was performed using IQ-Tree (Nguyen et al., 2015), under an edge-linked partition model with 5000 ultrafast bootstrap replicates. The best partitioning model was determined by ModelFinder (Kalyaanamoorthy et al. 2017).

Results

The complete mitochondrial genome (mitogenome) of *Hasora vitta* (GenBank Accession No. PP789054) is 15,289 bp in length, which encodes 13 PCGs, 22 transfer RNAs (tRNAs), two ribosomal RNAs (rRNAs), and a control region. The mitogenome is A + T biased (79.91%) with the nucleotide composition of A (39.56%), T (40.34%), C (12.26%), and G (7.83%). The mitogenome has a depth coverage of 169×, amounting to a total length of 11,281 bp for PCGs, 1472 bp for tRNAs, 789 bp for 12S rRNA, and 1350 bp for 16S rRNA. As commonly found across the Lepidoptera mitogenome, the COX1 gene of *Hasora vitta* utilizes the CGA start codon, and the rest of the genes were initiated by the typical ATN start codon. In the analysis, out of the 13 PCGs, two PCGs (COX1 and COX2) genes ended with an incomplete stop codon T, while the rest were terminated by the TAA stop codon.

Phylogenetic analysis showed that all subfamilies formed monophyletic clades. The result placed *Hasora vitta* reported here with the *Hasora vitta* sampled from China (KR076553)

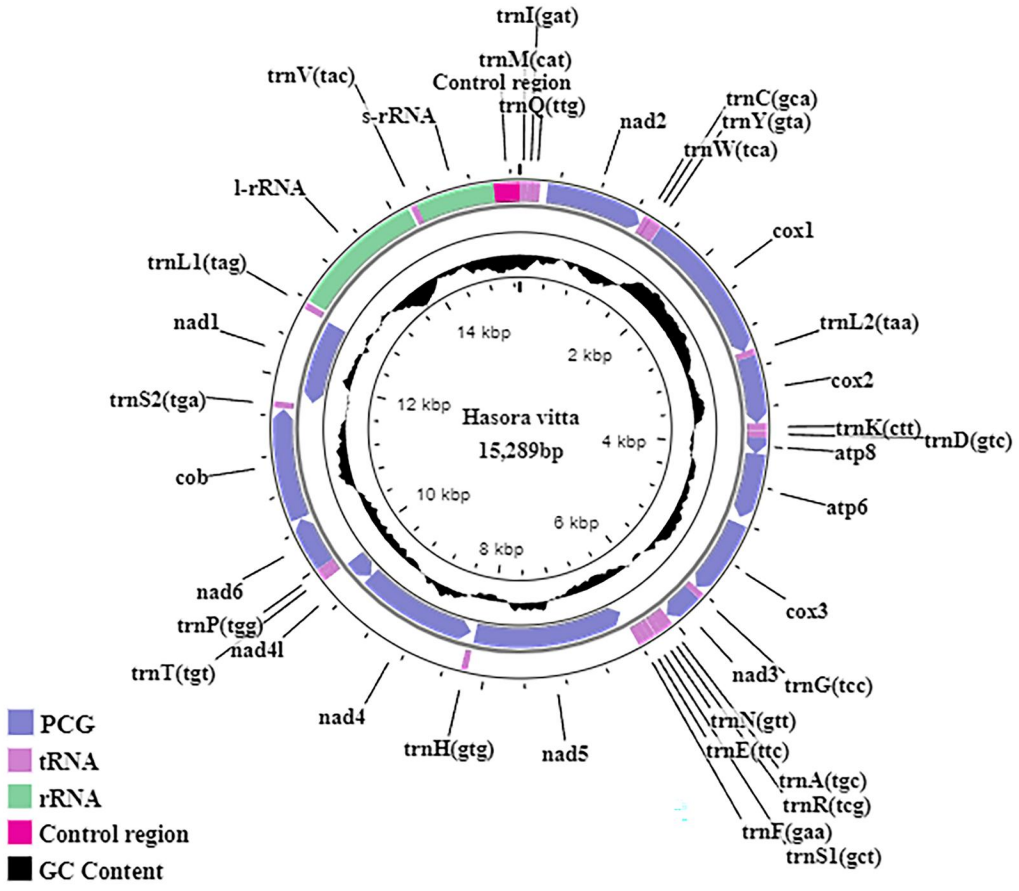

Figure 2. Mitogenome map of *Hasora vitta* generated using Proksee web server.

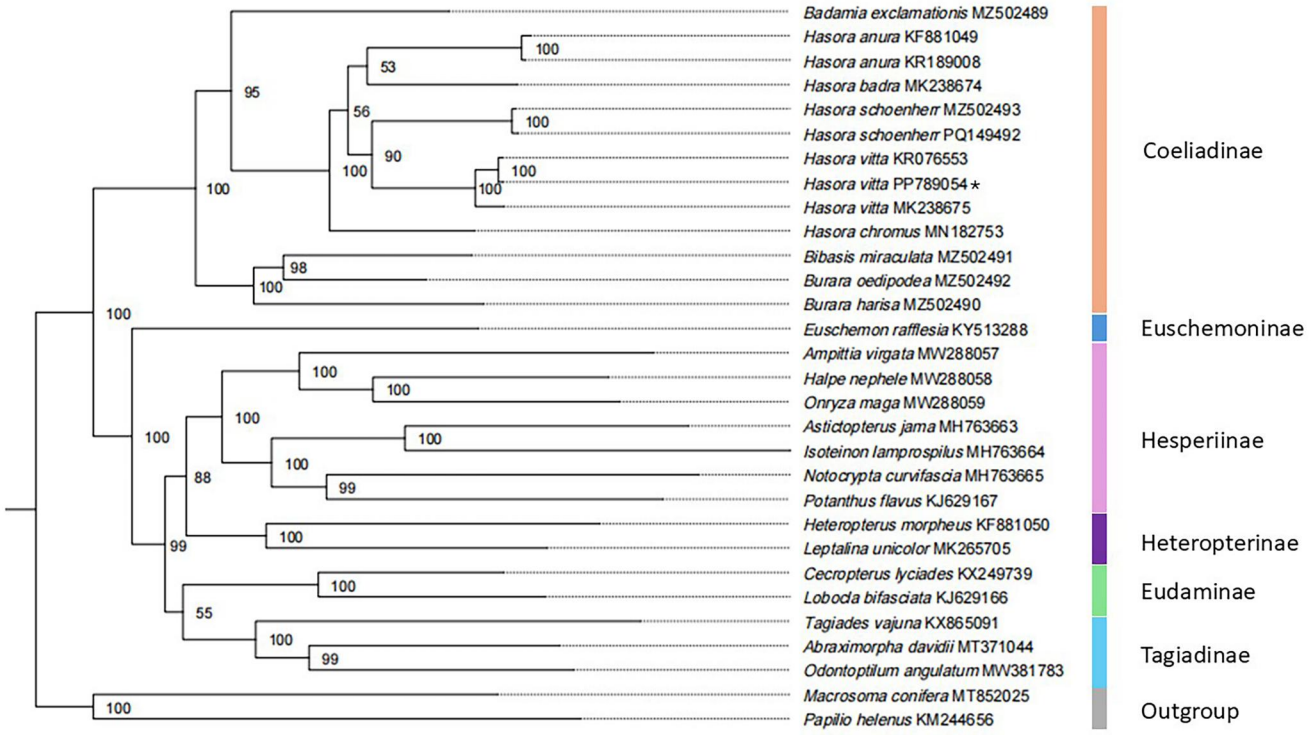

Figure 3. Phylogenetic analysis of *Hasora vitta* sequenced in this study, as indicated by asterisk (*). The node values represent the bootstrap of ultrafast 5000 replicates (BS) and each color indicates the different subfamilies in Hesperiidae and the outgroup. Ingroup: *Hasora vitta* (MK238675) (unpublished), (KR076553) (Cao et al. 2016); *Badamia exclamationis* (MZ502489), *Burara harisa* (MZ502490), *Bibasis miraculata* (MZ502491), *Burara oedipodea* (MZ502492), *Hasora schoenherr* (MZ502493) (Sun et al. 2021), (PQ149492) (unpublished); *Hasora anura* (KR189008),(KF881049) (Wang et al. 2016); *Hasora badra* (MK238674) (unpublished), *Hasora chromus* (MN182753) (unpublished); *Euschemon rafflesia* (KY513288) (Zhang et al. 2017); *Cecropterus lyciades* (KX249739) (unpublished); *Lobocla bifasciata* (KJ629166), *Potanthus flavus* (KJ629167) (Kim et al. 2014); *Ampittia virgata* (MW288057), *Halpe nephele* (MW288058), *Onryza maga* (MW288059) (Hao et al. 2021); *Astictopterus jama* (MH763663), *Isoteinon lamprospilus* (MH763664), *Notocrypta curvifascia* (MH763665) (Ma et al. 2020); *Heteropterus morpheus* (KF881050) (unpublished). Outgroup: *Papilio helenus* (KM244656) (Tang et al. 2014); *Macrosoma conifera* (MT852025) (McCullagh et al. 2020).

(Cao et al. 2016) and (MK238675) (unpublished) with a bootstrap support value of 100% and is closely related to *Hasora schoenherr* (MK502493) (Sun et al. 2021). Additionally, when both sequences were analyzed using BlastN, it shows a 99.15% similarity with each other.

Discussion and conclusions

The newly sequenced *Hasora vitta* reported has similar gene characteristics with other genera within the subfamily Coeliadinae. The utilization of the CGA start codon is also common in Lepidoptera mitogenome and the phenomena of the incomplete stop codon are presumed to be associated with the polyadenylation processes (Chen et al., 2020). When a BlastN analysis was conducted between two sequences of *H. vitta* originating from China (KR076553) and Malaysia (PP789054), the analysis showed a 99.52% similarity. Upon closer inspection of their alignment using Jalview (Waterhouse et al. 2009), a number of base differences were observed across the gene sequence which could explain the percentage similarities between both sequences. Additionally, indels (deletions) were also observed which could contribute to the result of their percentage identity. Therefore, the sequenced mitogenome of *Hasora vitta* reported here provides additional information that can be used to further analyze their phylogeny across a broad range of populations.

Acknowledgements

This work was done with the permission by the Department of Wildlife and National Parks Peninsular Malaysia (DWNP) (Perhilitan) (JPHL.600-6/1/4 JLD2(138)) and Johor National Parks Corporation (PTNJ) (TNJ 700-2/1/1 Bil. (29)). Marylin Miga: conceptualization, methodology, data curation, software, writing – original draft preparation. Aqilah Awg Abdul Rahman: methodology, data curation, and reviewing. Sivachandran Parimannan and Heera Rajandas: formal analysis, resources, reviewing and funding acquisition. Mohd Shahir Shamsir: methodology and reviewing. Frankie Thomas Sitam: methodology and reviewing. Lili Tokiman: methodology. Jai Kemalok: methodology. Faezah Mohd Salleh: conceptualization, methodology, resources, reviewing and editing, supervision and funding acquisition. All authors have read and approved the manuscript.

Author contributions

CRediT: **Marylin Miga**: Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing; **Aqilah Awg Abdul Rahman**: Data curation, Methodology, Supervision, Writing – review & editing; **Sivachandran Parimannan**: Formal analysis, Funding acquisition, Resources, Writing – review & editing; **Heera Rajandas**: Formal analysis, Funding acquisition, Resources, Writing – review & editing; **Frankie Thomas Sitam**: Methodology, Writing – review & editing; **Lili Tokiman**: Methodology; **Jai Kemalok**: Methodology; **Mohd Shahir Shamsir**: Methodology, Writing – review & editing; **Faezah Mohd Salleh**: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Disclosure statement

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

Funding

This work was supported/funded by the Ministry of Higher Education Malaysia under Fundamental Research Grant Scheme FRGS/1/2022/STG03/UTM/02/2.

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/ under the accession number PP789054. The associated BioProject, SRA, and BioSample numbers are PRJNA753627, SRR28674276, and SAMN40961112, respectively.

References

Cao L, Wang J, John YJ, Shiu MY, Yuan X, Liu J, Cao T. 2016. The complete mitochondrial genome of *Hasora vitta* (Butler, 1870) (Lepidoptera: Hesperiidae). Mitochondrial DNA A DNA Mapp Seq Anal. 27(4):3020–3021.

Chen Lu, Wahlberg N, Liao C-Q, Wang C-B, Ma F-Z, Huang G-H. 2020. Fourteen complete mitochondrial genomes of butterflies from the genus Lethe (Lepidoptera, Nymphalidae, Satyrinae) with mitogenome-based phylogenetic analysis. Genomics. 112(6):4435–4441. doi:10.1016/j.ygeno.2020.07.042.

Chiba H. 2009. A revision of the subfamily Coeliadinae (Lepidoptera – Hesperiidae). Bull Kitakyushu Mus Nat Hist Hum Hist Ser A (Nat Hist). 7:1–51.

Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C-Y, Graham M, Van Domselaar G, Stothard P. 2023. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 51(W1): W484–W492. doi:10.1093/nar/gkad326.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4). doi:10.1093/nar/gkw955

Afgan E, Nekrutenko A, Grüning BA, Blankenberg D, Goecks J, Schatz MC, Ostrovsky AE, Mahmoud A, Lonie AJ, Syme A, et al. 2022. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 50(W1):W345–W351.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 69(2):313–319. doi:10.1016/j.ympev.2012.08.023.

Hao X, Liu J, Chiba H, Xiao J, Yuan X. 2021. Complete mitochondrial genomes of three skippers in the tribe Aeromachini (Lepidoptera: Hesperiidae: Hesperiinae) and their phylogenetic implications. Ecol Evol. 11(12):8381–8393. doi:10.1002/ece3.7666.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14(6):587–589. doi:10.1038/nmeth.4285.

Kim MJ, Wang AR, Park JS, Kim I. 2014. Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperiidae) and phylogenetic reconstruction of Lepidoptera. Gene. 549(1):97–112. doi:10.1016/j.gene.2014.07.052.

Ma L, Liu F, Chiba H, Yuan X. 2020. The mitochondrial genomes of three skippers: insights into the evolution of the family Hesperiidae (Lepidoptera). Genomics. 112(1):432–441. doi:10.1016/j.ygeno.2019.03.006.

McCullagh BS, Alexiuk MR, Payment JE, Hamilton RV, Lalonde MML, Marcus JM. 2020. It's a moth! It's a butterfly! It's the complete mitochondrial genome of the American moth-butterfly *Macrosoma conifera* (Warren, 1897) (Insecta: Lepidoptera: Hedylidae)! Mitochondrial DNA B Resour. 5(3):3633–3635. doi:10.1080/23802359.2020.1831991.

Miga M, Yap YZ, Jahari PNS, Parimannan S, Rajandas H, Abu Bakar-Latiff M, Jing Wei Y, Shamsir MS, Mohd Salleh F. 2023. Complete mitochondrial genome data and phylogenetic analysis of the Great Marquis, *Bassarona dunya* (Doubleday, 1848) (Lepidoptera: Nymphalidae: Limenitidinae) from Malaysia. Mitochondrial DNA B Resour. 8(1):167–171. doi:10.1080/23802359.2023.2167476.

Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, Marshall D. 2010. Tablet-next generation sequence assembly visualization. Bioinformatics. 26(3):401–402. doi:10.1093/bioinformatics/btp666.

Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274. doi:10.1093/molbev/msu300.

Schubert M, Ermini L, Der Sarkissian C, Jónsson H, Ginolhac A, Schaefer R, Martin MD, Fernández R, Kircher M, McCue M, et al. 2014. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nat Protoc. 9(5):1056–1082. doi:10.1038/nprot.2014.063.

Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. BMC Research Notes 9(88):1–7. doi:10.1186/s13104-016-1900-2

Sun Q, Yang Y, Hao X, Xiao J, Liu J, Yuan X. 2021. Comparative mitogenomic analysis of five awl skippers (Lepidoptera: Hesperiidae: Coeliadinae) and their phylogenetic implications. Insects. 12(8):757. doi:10.3390/insects12080757.

Tang M, Tan M, Meng G, Yang S, Su X, Liu S, Song W, Li Y, Wu Q, Zhang A, et al. 2014. Multiplex sequencing of pooled mitochondrial genomes – a crucial step toward biodiversity analysis using mito-metagenomics. Nucleic Acids Res. 42(22):e166. doi:10.1093/nar/gku917.

Wang J, James John Y, Xuan S, Cao T, Yuan X. 2016. The complete mitochondrial genome of the butterfly *Hasora anura* (Lepidoptera: Hesperiidae). Mitochondrial DNA A DNA Mapp Seq Anal. 27(6):4401–4402. doi:10.3109/19401736.2015.1089543.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics. 25(9):1189–1191. doi:10.1093/bioinformatics/btp033.

Zhang J, Cong Q, Shen J, Fan XL, Wang M, Grishin NV. 2017. The complete mitogenome of *Euschemon rafflesia* (Lepidoptera: Hesperiidae). Mitochondrial DNA B Resour. 2(1):136–138. doi:10.1080/23802359.2017.1292478.