

The complete mitogenome of *Euschemon rafflesia* (Lepidoptera: HesperIIDae)

Jing Zhang^a, Qian Cong^a, Jinhui Shen^a, Xiao-Ling Fan^b, Min Wang^b and Nick V. Grishin^{a,c}

^aDepartments of Biophysics and Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX, USA; ^bDepartment of Entomology, South China Agricultural University, Guangzhou, China; ^cHoward Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA

ABSTRACT

We assembled a complete mitochondrial genome of a unique Australian skipper butterfly *Euschemon rafflesia* (HesperIIDae) from next generation sequencing reads. The 15,447 bp mitogenome covers 13 protein-coding genes (PCGs), 22 transfer RNA genes (tRNAs), 2 ribosomal RNA genes (rRNAs), and an A + T-rich region. Its gene order is typical for mitogenomes of Lepidoptera. Phylogenetic analysis places *Euschemon rafflesia* as a sister to the rest of HesperIIDae except Coeliadinae.

ARTICLE HISTORY

Received 28 January 2017
Accepted 5 February 2017

KEYWORDS

Next-generation sequencing; phylogeny; Euschemoninae; frenulum; Coeliadinae



The Regent Skipper (*Euschemon rafflesia*) is an Australian endemic and the only member in the subfamily Euschemoninae (family HesperIIDae). It is a showy large butterfly with black, yellow-spotted wings and red abdomen tip, over 5 cm in wingspan (Figure 1). Like no other butterfly, it possesses a frenulum and retinaculum (in males only) (Braby 2000), frequently present in moths to couple the wings. Phylogenetic affinities of *Euschemon* have been debated for years, and it has even been considered to be a moth by earlier authors (Watson 1893). Later, it has been placed in or near HesperIIDae, possibly related to *Celaenorrhinus* (Braby 2000). Recently, *Euschemon* was given a subfamily status based on the combination of DNA and morphological evidence (Warren et al. 2008; Warren et al. 2009), the result corroborated in an expanded phylogenetic study (Sahoo et al. 2016).

To better understand the phylogeny of HesperIIDae and clarify the phylogenetic position of the Regent Skipper, we sequenced, assembled and annotated the complete mitogenome of *Euschemon rafflesia rafflesia* from the male voucher NVG-15103B05 in the National Museum of Natural History collection (Smithsonian Institution, Washington DC) from Australia: Queensland, Southbrook, collected around 1946. The specimen is illustrated in Figure 1. A single leg was used for DNA extraction. Methods for genomic DNA extraction, library construction, next-generation sequencing, and computational procedures have been reported by us previously (Shen et al. 2015; Cong & Grishin 2016; Cong et al. 2016a, 2016b; Shen et al. 2016). The mitogenome of *Lobocla bifasciatus* (Kim et al. 2014) was used as a reference to search for ('bait') similar sequence reads using BWA (Li & Durbin 2009). Nearly 0.8% (159,718 out of 20,472,582) of *E. rafflesia* total

genomic reads were extracted by BWA for mitogenome assembly (Hahn et al. 2013). The complete mitogenome of *E. rafflesia* was assembled *de novo* using Platanus (Kajitani et al. 2014) followed by a manual gap-closing procedure.

The complete mitogenome of *Euschemon rafflesia* is 15,447 bp in length (Genbank: KY513288) and is AT-rich, with a base composition of 39.1% A, 42.3% T, 7.4% G, and 11.2% C. It retains the typical insect mitogenome gene set, including 13 PCGs (ND1-6, COX1-3, ND4L, ATP8, ATP6, and CYTB), 22 tRNA genes (two for serine and leucine and one for each of the rest amino acids), 2 ribosomal RNAs (rrnL and rrnS), and an A + T-rich D-loop control region. As in many Lepidoptera mitogenomes, the exact start of COX1 gene is unknown, but is probably the codon TTG (Kim et al. 2009). The typical start codon ATN is used in other genes. COX1, COX2, ND4 and ND5 genes have an incomplete stop codon T, and a complete TAA codon is likely formed during mRNA maturation (Ojala et al. 1981; Boore 1999). The length of tRNAs ranges from 59 to 71 bp. The size of the two rRNAs are 1351 and 775 bp, respectively. A 453 bp A + T-rich region connects rrnS and tRNA-Met.

To phylogenetically place *Euschemon rafflesia* within HesperIIDae with available mitogenomes (Hao et al. 2012; Kim et al. 2014; Shen et al. 2015; Shao et al. 2015; Wang et al. 2013, 2014, 2015; Shen et al. 2016; Cong & Grishin 2016), we constructed RaxML (Stamatakis, 2006) maximum likelihood tree rooted with *Pterourus glaucus* (Papilionidae) mitogenome (Shen et al. 2015) (Figure 1). The placement of *Euschemon* as a sister to the rest of HesperIIDae except Coeliadinae is strongly supported and is in agreement with the previous results (Warren et al. 2008, 2009; Sahoo et al. 2016). The tree topology

CONTACT Nick V. Grishin  grishin@chop.swmed.edu  Departments of Biophysics and Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX, USA

© 2017 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.

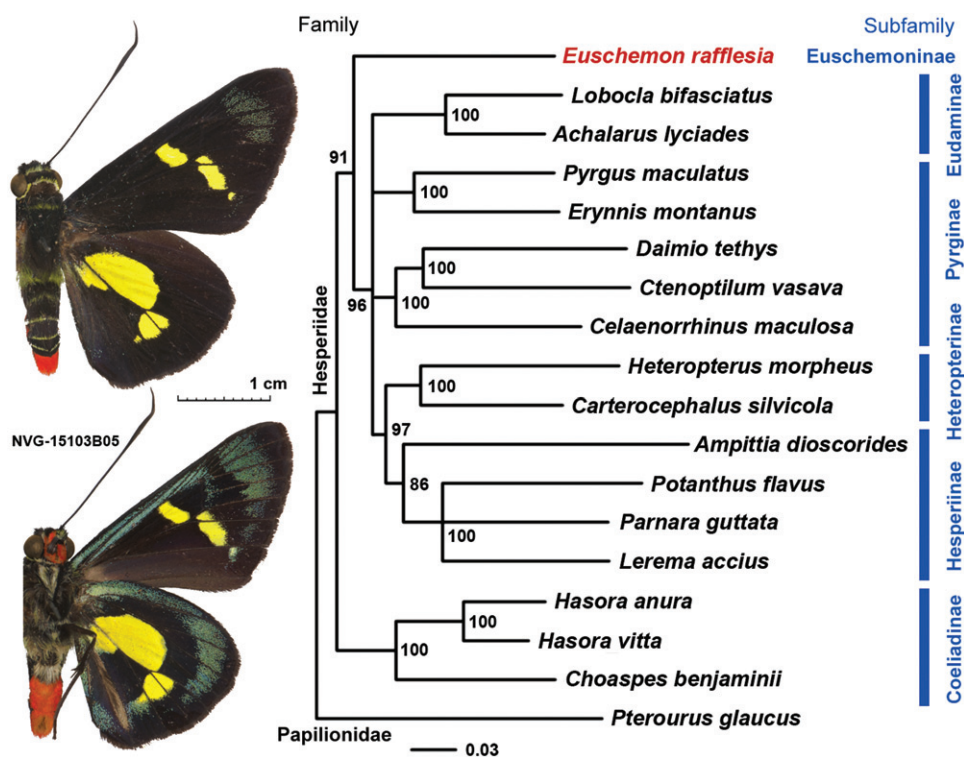


Figure 1. Maximum likelihood tree of complete mitogenomes of 17 Hesperidae species rooted with *Pterourus glaucus* (Papilionidae). *Euschemon rafflesia* with mitogenome reported here is shown first and the specimen sequenced (voucher NVG-15103B05) is pictured on the left, dorsal and ventral sides above and below, respectively. Numbers by the nodes show bootstrap support values and branches with bootstrap less than 60% are collapsed. GenBank accessions for sequences are: *Acharus lyciades* NC_030602.1; *Ampittia dioscorides* KM102732.1; *Celaenorrhinus maculosa* NC_022853.1; *Daimio tethys* NC_024648.1; *Euschemon rafflesia* KY513288; *Erynnis montanus* NC_021427.1; *Hasora anura* NC_027263.1; *Hasora vitta* NC_027170.1; *Heteropterus morpheus* NC_028506.1; *Choaspes benjaminii* NC_024647.1; *Lerema accius* NC_029826.1; *Lobocla bifasciatus* NC_024649.1; *Carterocephalus silvicola* NC_024646.1; *Potanthus flavus* NC_024650.1; *Parnara guttata* NC_029136.1; *Pyrgus maculatus* NC_030192.1; *Ctenoptilum vasava* NC_016704.1; *Papilio glaucus* NC_027252.

is consistent with previous phylogenetic studies (Warren et al. 2008, 2009; Sahoo et al. 2016): Coeliadinae are the sister to all other Hesperidae; relationship between Eudaminae and Pyrginae is unresolved; Heteropterinae are the sister to Hesperinae within which Aeromachini (represented by *Ampittia*) is the sister to the rest. As seen previously (Shen et al. 2016), bootstrap on mitogenomes is insufficient to support monophyly of Pyrginae, and the topology within the crown Hesperinae group needs to be investigated further. In conclusion, the complete mitogenome of *Euschemon rafflesia* clarifies its phylogenetic position and strengthens the evidence for this unique skipper to be treated as a sole representative of the subfamily Euschemoninae.

Acknowledgements

We are grateful to Robert K. Robbins, John M. Burns, and Brian Harris (National Museum of Natural History, Smithsonian Institution, Washington, DC) for granting access to the collections under their care and stimulating discussions.

Disclosure statement

The authors declare no conflicts of interest.

Funding

This work was supported by the National Institutes of Health (NVG) and the Welch Foundation (I-1505 to NVG).

References

- Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* 27:1767–1780.
- Braby MF, Upton MS, Australian National Insect Collection & CSIRO Entomology. 2000. *The butterflies of Australia: their identification, biology and distribution.* Melbourne: CSIRO publishing.
- Cong Q, Grishin NV. 2016. The complete mitochondrial genome of *Lerema accius* and its phylogenetic implications. *Peer J.* 4:e1546.
- Cong Q, Shen J, Borek D, Robbins RK, Otwinowski Z, Grishin NV. 2016a. Complete genomes of Hairstreak butterflies, their speciation, and nucleo-mitochondrial incongruence. *Sci Rep.* 6:24863.
- Cong Q, Shen J, Warren AD, Borek D, Otwinowski Z, Grishin NV. 2016b. Speciation in cloudless sulphurs gleaned from complete genomes. *Genome Biol Evol.* 8:915–931.
- Hahn C, Bachmann L, Chevreaux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* 41:e129.
- Hao J, Sun Q, Zhao H, Sun X, Gai Y, Yang Q. 2012. The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperidae: Pyrginae) and its phylogenetic implication. *Comp Funct Genomics.* 2012:328049.
- Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama HN, Kohara Y, et al. 2014. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res.* 24:1384–1395.
- Kim MI, Baek JY, Kim MJ, Jeong HC, Kim KG, Bae CH, Han YS, Jin BR, Kim I. 2009. Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects. *Mol Cells.* 28:347–363.
- Kim MJ, Wang AR, Park JS, Kim I. 2014. Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperidae) and phylogenetic reconstruction of Lepidoptera. *Gene.* 549:97–112.

- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25:1754–1760.
- Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*. 290:470–474.
- Sahoo RK, Warren AD, Wahlberg N, Brower AVZ, Lukhtanov VA, Kodandaramaiah U. 2016. Ten genes and two topologies: an exploration of higher relationships in skipper butterflies (Hesperiidae). *Peer J*. 4:e2653.
- Shao L, Sun Q, Hao J. 2015. The complete mitochondrial genome of *Parara guttata* (Lepidoptera: Hesperidae). *Mitochondrial DNA*. 26:724–725.
- Shen J, Cong Q, Grishin NV. 2015. The complete mitochondrial genome of *Papilio glaucus* and its phylogenetic implications. *Meta Gene*. 5:68–83.
- Shen J, Cong Q, Grishin NV. 2016. The complete mitogenome of *Achalarus lyciades* (Lepidoptera: Hesperidae). *Mitochondrial DNA B Res*. 1:581–583.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 22:2688–2690.
- Wang AR, Jeong HC, Han YS, Kim I. 2014. The complete mitochondrial genome of the mountainous duskywing, *Erynnis montanus* (Lepidoptera: Hesperidae): a new gene arrangement in Lepidoptera. *Mitochondrial DNA*. 25:93–94.
- Wang J, James John Y, Xuan S, Cao T, Yuan X. 2015. The complete mitochondrial genome of the butterfly *Hasora anura* (Lepidoptera: Hesperidae). *Mitochondrial DNA*. 27:4401–4402.
- Wang K, Hao J, Zhao H. 2013. Characterization of complete mitochondrial genome of the skipper butterfly, *Celaenorrhinus maculosus* (Lepidoptera: Hesperidae). *Mitochondrial DNA*. 26:690–691.
- Warren AD, Ogawa JR, Brower AVZ. 2008. Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperidae (Lepidoptera: Hesperioidea). *Cladistics*. 24:642–676.
- Warren AD, Ogawa JR, Brower AVZ. 2009. Revised classification of the family Hesperidae (Lepidoptera: Hesperioidea) based on combined molecular and morphological data. *Syst Entomol*. 34:467–523.
- Watson EY. 1893. A proposed classification of the Hesperidae, with a revision of the genera. *Proc Zool Soc Lond*. 1893:3–132.