




Three mitochondrial genomes of early-winged insects (Ephemeroptera: Baetidae and Leptophlebiidae)

Sereina Rutschmann^{a,b,c} , Ping Chen^d, Changfa Zhou^d  and Michael T. Monaghan^{a,b} 

^aLeibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany; ^bBerlin Center for Genomics in Biodiversity Research, Berlin, Germany; ^cUniversity of Potsdam, Institute of Biochemistry and Biology, Potsdam, Germany; ^dCollege of Life Sciences, The Key Laboratory of Jiangsu Biodiversity and Biotechnology, Nanjing Normal University, Nanjing, China

ABSTRACT

Mayflies (Ephemeroptera) are a semi-aquatic insect order with comparatively few genomic data available despite their phylogenetic position at the root of the winged-insects and possession of ancestral traits. Here, we provide three mitochondrial genomes (mtgenomes) from representatives of the two most species-rich families, *Baetis rutilocylindratus* and *Cloeon dipterum* (Baetidae), and *Habrophlebiodes zijinensis* (Leptophlebiidae). All mtgenomes had a complete set of 13 protein-coding genes and a conserved orientation except for two inverted tRNAs in *H. zijinensis*. Phylogenetic reconstructions using 21 mayfly mtgenomes and representatives of seven additional orders recovered both Baetidae and Leptophlebiidae as well supported monophyletic clades, with Ephemeroptera as the sister-taxon to all other winged insects (i.e. Odonata and Neoptera).

ARTICLE HISTORY

Received 8 May 2021
Accepted 21 August 2021



KEYWORDS

Baetis; *Cloeon*;
Habrophlebiodes; mayfly;
mitochondrial phylogeny

The number of published molecular markers and genomes of the semi-aquatic insect order Ephemeroptera (mayflies) is low compared to the evolutionary diversity that the order represents (Almudi et al. 2020; Hotaling et al. 2020; Xu et al. 2020). The mayflies comprise >3000 species in around 42 families and >400 genera, and fossil records date back to the lower Carboniferous (Barber-James et al. 2008; Ogden et al. 2009). The monophyly of the order is well established (Hovmöller et al. 2002; Ogden et al. 2009), although its relationship to Odonata and all other winged insects remains a long-standing controversy (Hovmöller et al. 2002; Ogden and Whiting 2003; Misof et al. 2014). Within the mayflies, the relationships of several major clades remain unresolved (e.g. Miller et al. 2018; Xu et al. 2020). The availability of additional molecular data will help to shed light on these unresolved issues.

Here, we newly sequenced three mitochondrial genomes (mtgenomes) of the two most diverse Ephemeroptera families: *Baetis rutilocylindratus* Wang, Qin, Chen & Zhou, 2011 and *Cloeon dipterum* L. 1761 from the Baetidae, and *Habrophlebiodes zijinensis* Gui, Zhang & Wu, 1996 from the Leptophlebiidae. Specimens of *H. zijinensis* and *B. rutilocylindratus* were collected in Zijin Hill, Nanjing, China (32.055 latitude, 118.859 longitude). Their DNA was extracted using the DNeasy[®] Blood & Tissue (Qiagen, Hilden, Germany) kit. Initially, four fragments of the mtgenome were amplified via standard PCR using universal primers (Simon et al. 1994) and

subsequently the sequences were extended by newly designed primers (see [Supplementary Information Table S1](#)). All purified amplification products were sequenced successively in both directions using a 3130 xl (Applied Biosystems, Waltham, MA) capillary sequencer. The mtgenome of *C. dipterum* was obtained from DNA of laboratory-reared subimago specimens (full siblings). The laboratory-founder individual was collected in the United States of America (39.865 latitude, -75.818 longitude). Analysis of nuclear and mitochondrial markers established the species as one of six putative species of *C. dipterum* s.l. with a geographical distribution in North America and Europe (Rutschmann et al. 2017). Specimens and DNA of the species are stored at The Key Laboratory of Jiangsu Biodiversity and Biotechnology (Nanjing, China, voucher numbers NNU-20110910, NNU-2010710) and the Museum of Zoology (Lausanne, Switzerland, voucher number US). We extracted the DNA, prepared a shotgun library, and sequenced four lanes on a 454 GS FLX pyrosequencer (Roche Diagnostics, Rotkreuz, Switzerland). Sequence reads were trimmed and *de novo* assembled as described in Rutschmann et al. (2017). In order to extract the mtgenome, we performed BLASTN searches (Altschul 1997) using all assembled contigs to query the NCBI database. We mapped all matching reads back to the mtgenome with BWA (Li and Durbin 2009), using the BWA-SW algorithm (Li and Durbin 2010). All mtgenomes were

CONTACT Michael T. Monaghan  monaghan@igb-berlin.de  Leibniz Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 301, 12587 Berlin, Germany

 Supplemental data for this article is available online at <https://doi.org/10.1080/23802359.2021.1974966>.

© 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.

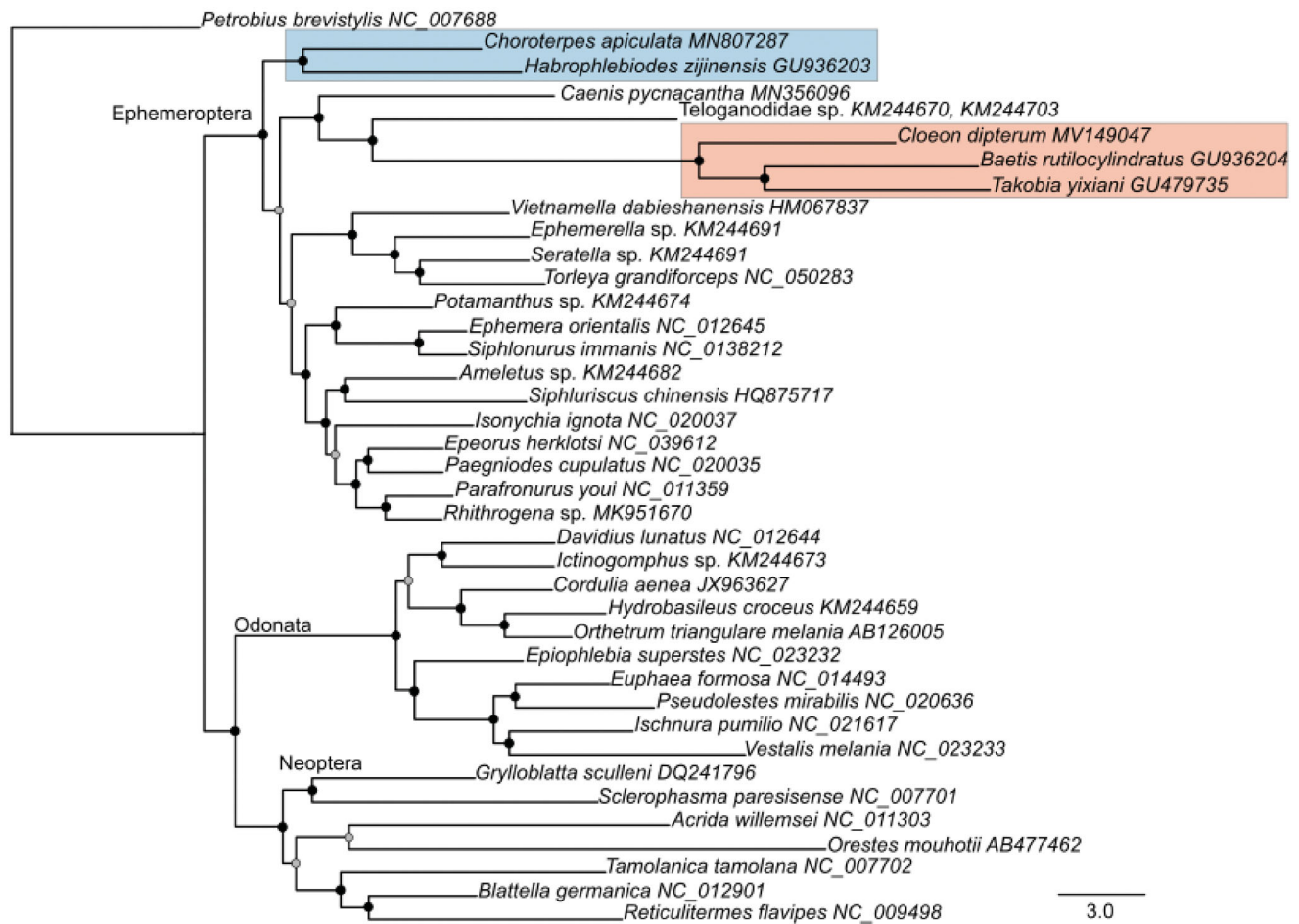


Figure 1. Phylogenetic relationships of insect major orders using Bayesian inference reconstruction based on the nucleotide dataset of the 13 mitochondrial protein-coding genes. Filled circles indicate supported nodes; whereby black circles represent Bayesian posterior probability (BPP) ≥ 0.99 , and shaded circles BPP ≥ 0.95 . The two families of Baetidae and Leptophlebiidae are highlighted. The GenBank accession numbers of all specimens are provided in the tip labels.

annotated using MITOS (Bernt et al. 2013) and tRNAscan-SE v.1.21 (Lowe and Eddy 1997).

The sequences are deposited in GenBank with the accession numbers GU936204, MW149047, and GU936203. They were 14,883 bp (*B. rutilocylindratus*), 15,407 bp (*C. dipterum*), and 14,355 bp (*H. zijinensis*) long, whereby the one of *B. rutilocylindratus* was complete (i.e. circular). All three contained the entire set of 13 protein-coding genes (PCGs), two rRNAs, and 22 (*B. rutilocylindratus* and *C. dipterum*) or 19 tRNAs (*H. zijinensis*). Gene order was conserved with the exception of two inverted tRNAs (arginine and alanine) in *H. zijinensis*.

Bayesian phylogenetic trees were reconstructed based on the nucleotide dataset of the 13 PCGs using MrBayes v.3.2.7 (Ronquist et al. 2012). We aligned each PCG separately using MAFFT v.7.407 (L-INS-I algorithm with default settings, Katoh and Standley 2013) and inferred the tree based on the concatenated alignment, applying the best-fit evolutionary models and partitioning-schemes as inferred with PartitionFinder v.2 (<https://github.com/brettc/partitionfinder>). *Petrobius brevistylis* (Archaeognatha, NC_007688) was specified as outgroup. The recovered phylogeny (Figure 1) was highly supported, including all nodes with Bayesian posterior probability (BPP ≥ 0.95). The Ephemeroptera were clustered as sister group to the remaining winged-insects, comprising the

Odonata and Neoptera. Within mayflies, the Leptophlebiidae were recovered as sister group to all other representatives. The Caenidae, Teloganodidae, and Baetidae formed a highly supported monophyletic clade.

Acknowledgements

We are grateful to D. H. Funk from the Stroud Water Research Center for providing the *Cloëon dipterum* specimens, S. Mbedi, K. Preuß, and L. Wächter for their great help with laboratory work, and G. Glöckner and C. J. Mazzoni for constructive comments on the analysis.

Disclosure statement

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

Funding

The work was supported by the Leibniz Association PAKT für Forschung und Innovation project 'FREDIE' [SAW-2011-ZFMK-3].

ORCID

Sereina Rutschmann  <http://orcid.org/0000-0002-7391-0836>
 Changfa Zhou  <http://orcid.org/0000-0001-8785-5228>
 Michael T. Monaghan  <http://orcid.org/0000-0001-6200-2376>

Data availability statement

The data that were newly sequenced as part of this study are available in NCBI (<https://www.ncbi.nlm.nih.gov/>, accession numbers GU936204, MW149047, and GU936203). The data are associated with BioSample SAMN03202660 and BioProject PRJNA268073, and the raw sequencing reads are available from the Sequence Read Archive SRP050093.

References

- Almudi I, Vizueta J, Wyatt CDR, de Mendoza A, Marlétaz F, Firbas PN, Feuda R, Masiero G, Medina P, Alcaina-Caro A, et al. 2020. Genomic adaptations to aquatic and aerial life in mayflies and the origin of insect wings. *Nat Commun.* 11(1):2631.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389–3402.
- Barber-James HM, Gattolliat JL, Sartori M, Hubbard MD. 2008. Global diversity of mayflies (Ephemeroptera, Insecta) in freshwater. *Hydrobiologia.* 595(1):339–350.
- 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. *Mol Phylogenet Evol.* 69(2): 313–319.
- Hotaling S, Kelley JL, Frandsen PB. 2020. Aquatic insects are dramatically underrepresented in genomic research. *Insects.* 11(9):601.
- Hovmöller R, Pape T, Källersjö M. 2002. The Paleoptera problem: basal pterygote phylogeny inferred from 18S and 28S rDNA sequences. *Cladistics.* 18(3):313–323.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 25(14):1754–1760.
- Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics.* 26(5):589–595.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25(5): 955–964.
- Miller DB, Bartlett S, Sartori M, Breinholt JW, Ogden TH. 2018. Anchored phylogenomics of burrowing mayflies (Ephemeroptera) and the evolution of tusks. *Syst Entomol.* 43(4):692–701.
- Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutel RG, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science.* 346(6210):763–767.
- Ogden TH, Gattolliat JL, Sartori M, Staniczek AH, Soldán T, Whiting MF. 2009. Towards a new paradigm in mayfly phylogeny (Ephemeroptera): combined analysis of morphological and molecular data. *Syst Entomol.* 34(4):616–634.
- Ogden HT, Whiting MF. 2003. The problem with “the Paleoptera Problem:” sense and sensitivity. *Cladistics.* 19(5):432–442.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 61(3):539–542.
- Rutschmann S, Detering H, Simon S, Funk DH, Gattolliat J-L, Hughes SJ, Raposeiro PM, DeSalle R, Sartori M, Monaghan MT. 2017. Colonization and diversification of aquatic insects on three Macaronesian archipelagos using 59 nuclear loci derived from a draft genome. *Mol Phylogenet Evol.* 107:27–38.
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am.* 87(6):651–701.
- Xu XD, Jia YY, Cao SS, Zhang ZY, Storey KB, Yu DN, Zhang JY. 2020. Six complete mitochondrial genomes of mayflies from three genera of Ephemerellidae (Insecta: Ephemeroptera) with inversion and translocation of trnl rearrangement and their phylogenetic relationships. *PeerJ.* 8:e9740.