

## The complete mitochondrial genome of *Semblis atrata* (Trichoptera: Phryganeidae)

Enzhi Liu\*, Junjian Li\*, Siyang Ou, Bingjun Dong, Baotian Yang and Yu Zhou

College of Life Sciences, Shenyang Normal University, Shenyang, China

### ABSTRACT

*Semblis atrata* is one of three *Semblis* species distributed in clean brooks and streams in northern Eurasia. Genomic DNA of an *S. atrata* sample was extracted and sequenced for assembly and annotation of its complete mitogenome. The complete mitochondrial genome of *S. atrata* was 14,909 bp in length and consisted of 13 protein-coding genes, 2 rRNA genes, and 22 tRNA genes. The *S. atrata* COX1 gene features a CGA start codon, and COX1, COX2, ND1, and ND5 exhibit incomplete stop codons that are presumed to be completed by the addition of 3' A residues to the mRNA. The nucleotide composition was highly AT biased, accounting for 77.71% of the whole mitogenome. Phylogenetic analysis placed *Semblis* as sister to *Eubasilissa*. The complete mitochondrial genome will be helpful for further studies on the population genetics of this species and phylogenetic analyses of Trichoptera.

### ARTICLE HISTORY

Received 3 August 2021  
Accepted 15 May 2022



### KEYWORDS

Illumina sequencing;  
mitogenomics; phylogeny;  
*Semblis atrata*

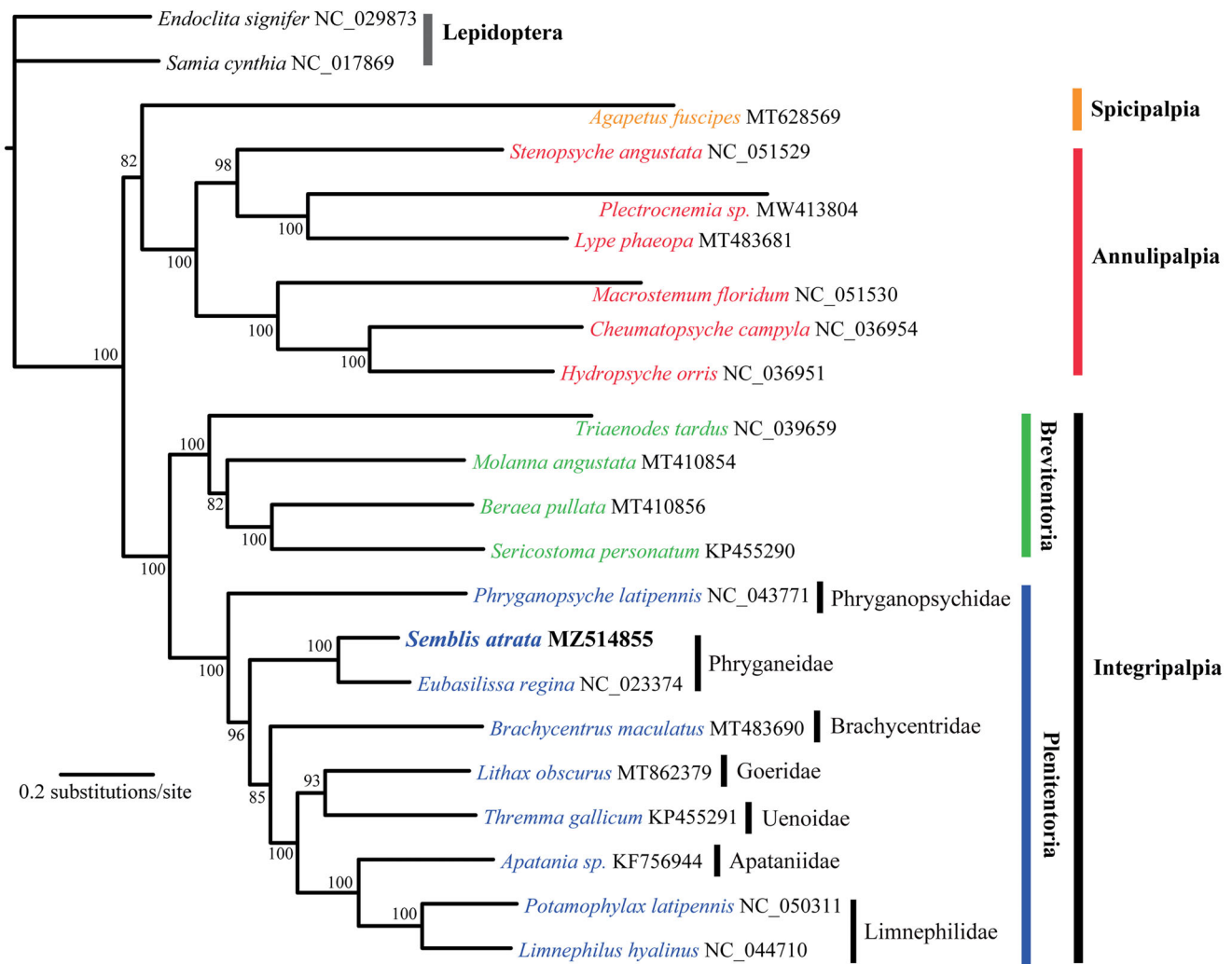
The genus *Semblis* within family Phryganeidae has the closest phylogenetic relationship with *Eubasilissa* according to molecular phylogenetic evidence (Thomas et al. 2020). *Semblis atrata* Gmelin 1789 is an obscure species distributed throughout northern Eurasia. The larvae of *S. atrata* predominantly inhabit slow-flowing brooks and streams with clear, humic, oxygen-rich, slightly acidic water and stony sand bottoms. Adults of *S. atrata* ubiquitously fly around such brooks and streams (Berglind et al. 1999). Currently, the only published mitochondrial DNA sequences of *Semblis* species are partial COX1 gene sequences, which have been used for DNA barcoding or multi-gene phylogenetic research (Zhou et al. 2016; Thomas et al. 2020). The complete mitochondrial genome of *S. atrata* is thus of great importance for shedding light on the diversity and phylogenetic relationships of Trichoptera.

The research materials were collected from Hushan village, Anshan, Liaoning Province, China (123.460°E, 40.787°N) in May 2021. After collection, specimens were immediately preserved in 95% ethanol until DNA extraction. A voucher specimen was deposited at Shenyang Normal University, Shenyang, China (Yu Zhou is the contact person: [zhouyu1988@outlook.com](mailto:zhouyu1988@outlook.com)) under voucher number ZY-2021050501. Total genomic DNA was isolated using the TIANamp Genomic DNA Kit (TIANGEN Biotech) according to the manufacturer's instructions. The sequencing library was prepared following McCullagh and Marcus (2015) and sequenced by the Illumina HiSeq 2500 platform with 150 bp paired-end sequencing (Sangon Biotech, Shanghai, China). The mitochondrial genome was assembled *de novo* using

NOVOPlasty (Dierckxsens et al. 2017) with a partial mitochondrial COX1 sequence of *S. atrata* (GenBank: KX107215) (Zhou et al. 2016) used as the seed sequence. The mitochondrial genome was annotated using the MITOS Webserver (Bernt et al. 2013) with reference to the complete *Eubasilissa regina* mitogenome (GenBank: NC023374). The annotated mitogenome was submitted to GenBank under accession number MZ514855. For each of the 13 protein-coding genes (PCGs), DNA sequences were aligned using PRANK (Löytynoja 2014) with the 'mtttranslate' alignment option. The alignments were further refined using Gblocks 9.1 b (Castresana 2000) with the 'codon' model and other default settings. All refined alignments were then concatenated into the final data set. For the concatenated data set, we manually defined three partitioning strategies: unpartitioned, three partitions (one partition for each codon position), and 13 partitions (one partition for each PCG). Comparisons of the three partitioning strategies and selection of corresponding nucleotide substitution models were conducted with the Bayesian information criterion implemented in PartitionFinder (Lanfear et al. 2012). The 3-partition scheme (one partition for each codon position) was chosen as the best-fitting partitioning strategy, and all three partitions favored the GTR+ $\Gamma$ +I model. The phylogenetic relationships between *Semblis* and 19 other Trichoptera species were reconstructed using maximum likelihood (ML) analysis. Two species from the sister order Lepidoptera were included as outgroups (Figure 1). The ML tree was estimated by using RAxML version 8.0 (Stamatakis 2014) with the 'GTRGAMMAI' model. Support for nodes in

**CONTACT** Yu Zhou  [zhouyu1988@outlook.com](mailto:zhouyu1988@outlook.com)  College of Life Sciences, Shenyang Normal University, 253 Huanghe North Street, Shenyang 110034, China  
\*These authors contributed equally to this work.

© 2022 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.** The maximum likelihood (ML)-based phylogenetic relationships of twenty Trichoptera species based on 13 protein-coding genes.

the ML tree was assessed with a rapid bootstrap analysis (option -f a) with 1,000 replicates.

The circular mitochondrial genome of *S. atrata* is 14,909 bp in length. According to the annotation results, it contains 13 PCGs, 22 tRNA genes, and 2 rRNA genes, which is typical of animal mitogenomes (Boore 1999). The nucleotide composition of this mitogenome is A 39.11%, C 14.11%, G 8.18%, and T 38.60%, with a higher A + T content (77.71%) than G + C content (22.29%). These characteristics of the mitogenome are similar to those of other Phryganeidae mitogenomes (Wang et al. 2013; Zhou et al. 2016). The mitochondrial gene composition and order in *S. atrata* are identical to those in most other Trichoptera mitogenomes (Wang et al. 2013; Living Prairie Mitogenomics Consortium 2018; Marcus 2018). Four of the protein-coding genes (*COX1*, *COX2*, *ND1*, and *ND5*) have single-nucleotide stop codons (T) completed by posttranscriptional addition of 3' A residues. Similar to that in many other insects, *COX1* in *S. atrata* begins with a start codon (CGA) (Liao et al. 2010). The gene arrangement and length of tRNA and rRNA regions are typical for Trichoptera (Wang et al. 2013; Lalonde and Marcus 2017; Living Prairie Mitogenomics Consortium 2018).

Phylogenetic analysis revealed the monophyly of Trichoptera, and all suborders and infraorders were monophyletic (Figure 1). The phylogenetic tree showed that *Semblis* is a sister group to *Eubasilissa* with maximal nodal support. The results thus validated the close relationship between the two genera, and the same results have been reported by Thomas et al. (2020). The complete mitogenome of *S. atrata* presented here will be useful for investigating the phylogeny and population genetics of Phryganeidae and even Trichoptera.

### Ethics statement

The study protocol was approved by the Committee on the Ethics of Animal Experiments of the College of Life Sciences, Shenyang Normal University.

### Author contributions

Y. Z. designed and conceived this work; E. L., S. O., B. D. and B. Y. collected the samples; E. L. and J. L. wrote the first version of the manuscript; Y. Z., E. L. and J. L. carried out the

investigation. All authors read, revised and approved the final manuscript.

### Disclosure statement

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

### Funding

This study was supported financially by the Research Fund for the Educational Commission of Liaoning Province of China [Grant No. LQN201904] and the Natural Science Foundation of Liaoning Province of China [Grant No. 2020-BS-148].

### Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of the NCBI at <https://www.ncbi.nlm.nih.gov/> under accession no. MZ514855. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA752916, SRR15371432, and SAMN20667750, respectively.

### References

- Berglund S-A, Engblom E, Lingdell P. 1999. Naturally rare, threatened, or overlooked? *Semblis phalaenoides* and *S. atrata* (Trichoptera: Phryganeidae) in Sweden. *Entomologisk Tidskrift*. 7:956–16.
- 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(6): 313–319.
- Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Res*. 27(8): 1767–1780.
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 17(4): 540–552.
- Dierckxens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 45(4):e18.
- Lalonde MML, Marcus JM. 2017. The complete mitochondrial genome of the long-horned caddisfly *Triaenodes tardus* (Insecta: Trichoptera: Leptoceridae). *Mitochondrial DNA B Resour*. 2(2): 765–767.
- Lanfear R, Calcott B, Ho SY, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol*. 29(6):1695–1701.
- Liao F, Wang L, Wu S, Li YP, Zhao L, Huang GM, Niu CJ, Liu YQ, Li MG. 2010. The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *Int J Biol Sci*. 6(2):172–186.
- Living Prairie Mitogenomics Consortium. 2018. The complete mitochondrial genome of the giant casemaker caddisfly *Phryganea cinerea* (Insecta: Trichoptera: Phryganeidae). *Mitochondrial DNA Part B*. 3: 375–377.
- Löytynoja A. 2014. Phylogeny-aware alignment with PRANK Multiple sequence alignment methods. Berlin: Springer; p. 155–170.
- Marcus JM. 2018. Our love-hate relationship with DNA barcodes, the Y2K problem, and the search for next generation barcodes. *AIMS Genet*. 5(1):1–23.
- McCullagh BS, Marcus JM. 2015. The complete mitochondrial genome of Lemon Pansy, *Junonia lemonias* (Lepidoptera: Nymphalidae: Nymphalinae). *J Asia-Pac Entomol*. 18(4):749–755.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.
- Thomas JA, Frandsen PB, Prendini E, Zhou X, Holzenthal RW. 2020. A multigene phylogeny and timeline for Trichoptera (Insecta). *Syst Entomol*. 45(3):670–686.
- Wang Y, Liu X, Yang D. 2013. The first mitochondrial genome for caddisfly (insecta: Trichoptera) with phylogenetic implications. *Int J Biol Sci*. 10(1):53–63.
- Zhou X, Frandsen PB, Holzenthal RW, Beet CR, Bennett KR, Blahnik RJ, Bonada N, Cartwright D, Chuluunbat S, Cocks GV, et al. 2016. The Trichoptera barcode initiative: a strategy for generating a species-level Tree of Life. *Phil Trans R Soc B*. 371(1702):20160025.