


The complete mitochondrial genome of *Hemiclepsis yangtzensis* (Clitellata: Glossiphoniidae)

Ti-Lin Yi^{a,b} , Meng-Ting Pei^c, Zhi-Wei Xu^a and Dai-Qing Yang^{a,b}

^aSchool of Animal Science, Yangtze University, Jingzhou, China; ^bHubei Provincial Engineering and Technology Research Center for *Monopterus albus*, Jingzhou, China; ^cJingzhou Institute of Technology, Jingzhou, China

ABSTRACT

We present the complete mitochondrial genome sequence of a recently described new leech species named *Hemiclepsis yangtzensis* Yang & Bolotov 2021 collected in central China. The mitochondrial genome is 14,984 bp in length and consists of 13 protein-coding genes, 2 ribosomal RNA genes, and 22 transfer RNA genes, all of which are encoded on a single strand. It exhibited a strong A + T bias of 72.87%. There is a large non-coding region (614 bp) located between the *tRNA-Arg* and *tRNA-His* genes, wherein we identified 40 short dispersed repeats, 13–22 bp long, 8 of which were direct, 20 inverted, and 12 palindromic. Phylogenetic analysis of 20 Hirudinea mitogenome sequences resolved monophyletic Glossiphoniidae, and *H. yangtzensis* formed a sister lineage with *Glossiphonia concolor*.

ARTICLE HISTORY

Received 18 October 2021
Accepted 2 April 2022

KEYWORDS

Hemiclepsis yangtzensis;
leech; mitochondrial
genome; phylogenetics



Leeches (Annelida: Clitellata: Hirudinea) are a specialized group (>680 species) of ectoparasites or predators that mostly inhabit freshwater habitats (Sket and Trontelj 2008). Leeches play an important role in the ecosystem, so they are used as an indicator of environmental stress (Koperski 2017). They are also used in neurobiological and developmental studies (Le Marrec-Croq et al. 2013). Due to a lack of molecular data, phylogeny and taxonomy of Hirudinea (and Annelida) remain only partially resolved (Weigert and Bleidorn 2016). Importantly, in contrast to early reports, which suggested that mitochondrial architecture is highly conserved in Annelida, with all genes encoded on a single strand; a more complex picture emerged during the last ten years, with a number of lineages exhibiting highly rearranged architecture, and some even have genes encoded on both strands (Oceguera-Figueroa et al. 2016; Weigert et al. 2016; Jiménez-Armenta et al. 2020; Daffe et al. 2021; Ye et al. 2021).


To contribute to the understanding of mitogenomic evolution in annelids, and generate data for future taxonomic and phylogenetic studies, in this study, we sequenced and characterized the complete mitochondrial genome of a very recently discovered and described new leech species *Hemiclepsis yangtzensis* Yang & Bolotov 2021 (Rhynchobdellida: Glossiphoniidae) (Xu et al. 2021). The specimen was collected from the skin of a *Monopterus albus* host in a fish breeding pond at the Aquatic Economic Animal Research Center of Yangtze University, 30.3590°N, 112.1376°E, Jingzhou, Hubei Province, China. The animal use protocol for this study was approved by the Animal Care and

Use Committee of Hubei Province (China). The host being farmed fish, and the target species an unprotected invertebrate, no additional permits were required for sampling. The type specimen is deposited in the Museum of Hydrobiological Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei Province, China (http://english.ihb.cas.cn/rh/rd/center12/201304/t20130418_101034.html; contact person: Zhiwei-Xu; e-mail: 2692420332@qq.COM) under the voucher number 18972378353.

The mitogenome was sequenced, assembled, and annotated following the methodology outlined before (Zou et al. 2017). Briefly, primers designed to match generally conserved regions of target genes were used to amplify short fragments of *16S*, *12S*, *cox1*, *atp6*, *cytb*, and *nad5*. Specific primers were designed based on these conserved regions sequences and used to amplify the remaining mitogenome sequence in several PCR reactions. The PCR reactions were carried out with LA Taq polymerase (Supplementary Data); 35 cycles comprised 94 °C for 30 s, 50 °C 30 s, and 72 °C for 1 min per 1 kb. PCR products were sequenced using Sanger sequencing. Sequences were assembled using DNASTar (Burland 2000), whereas annotation was conducted using Geneious (Kearse et al. 2012) and adjusted manually.

The complete mitochondrial genome (GenBank No. MN106285) is 14,984 bp in length. It encodes 13 protein-coding genes, 2 ribosomal RNA genes, and 22 transfer RNA genes (Supplementary data). All genes are encoded on a single strand. It exhibited very similar architectural features to other available Glossiphoniidae species. As this is the first sequenced mitogenome for the genus *Hemiclepsis*,

CONTACT Dai-Qing Yang  yangdaiq@163.com  School of Animal Science, Yangtze University, Jingzhou 434023, China

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2022.2070039>

© 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.

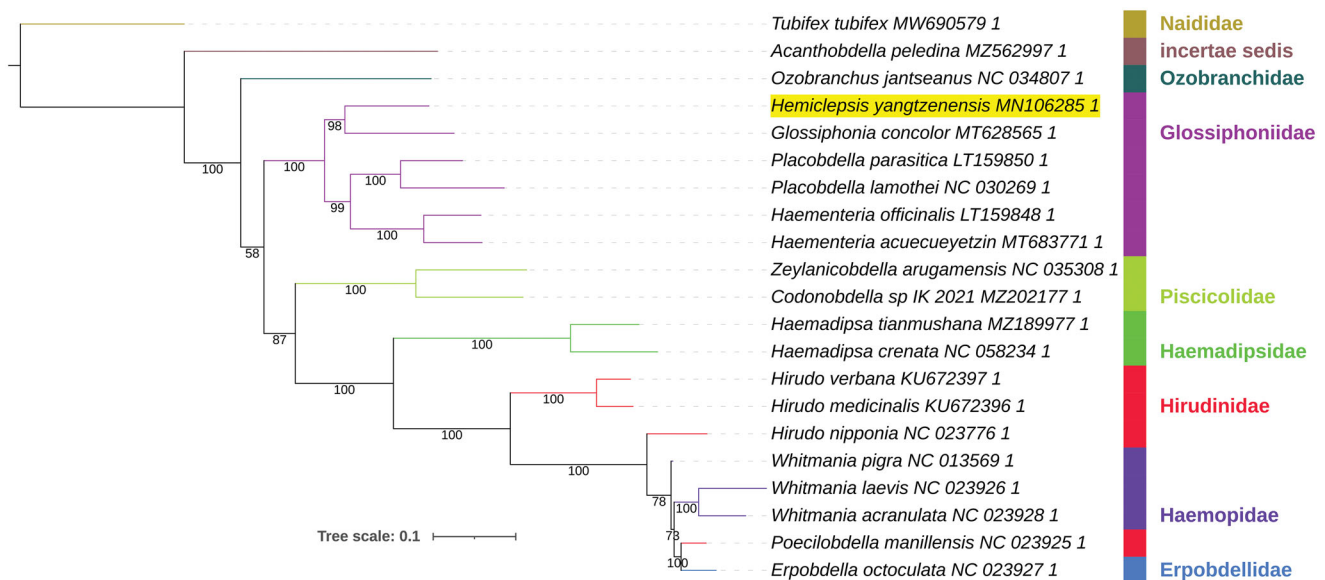


Figure 1. Phylogenetic tree inferred from the nucleotide sequences of 13 concatenated protein-coding genes using IQ-Tree. Bootstrap support values are shown at the nodes. Family identity is shown, *Hemiclepsis yangtzenensis* is highlighted by a yellow background, and *Tubifex tubifex* (Clitellata: Haplotaxida) is the outgroup.

comparisons to congeneric species currently remain impossible. It exhibited an identical gene order to another Glossiphoniidae species, *Haementeria officinalis* (Oceguera-Figueroa et al. 2016). Overall, gene order was almost identical to most other Hirudinea mitogenomes, notwithstanding rearrangements of individual tRNA genes (Supplementary Data).

The nucleotide composition of the mitogenome was 37.66% T, 15.04% C, 35.21% A, and 12.09% G, with a strong A+T bias of 72.87%, which is similar to other annelids (Ye et al. 2021). The G+C content of protein-coding genes was 26.61% (length: 11,134 bp), tRNAs 26.79% (length: 1,407 bp), and rRNAs 53.54% (length: 1,903) (Supplementary Data).

All protein-coding genes shared the same start codon, ATG, with the exception of *cox3*, which started with TTG. Seven protein-coding genes used the TAA stop codon (*cox1*, *cox3*, *nad6*, *cytb*, *atp6*, *nad4l*, and *nad2*), *nad3* used the TAG stop codon, and five genes (*cox2*, *atp8*, *nad4*, *nad1*, and *nad5*) had an incomplete stop codon T—. Incomplete stop codons are very frequent in mitogenomes of leeches, and they are most likely completed by post-transcriptional polyadenylation (Ojala et al. 1981; Oceguera-Figueroa et al. 2016). Five short non-coding regions (NCR) ranging from 1 bp to 5 bp and one large NCR (614 bp) were identified in the mitochondrial genome of *H. yangtzenensis*. This large NCR is located between *tRNA-Arg* and *tRNA-His* genes. REPuter tool (Kurtz et al. 2001) identified 40 short dispersed repeats, 13–22 bp long, in the NCR. Among these, 8 were direct, 20 inverted, and 12 were palindromic repeats. Tandem repeats have been reported in NCRs of some other annelids (Li et al. 2015).

Phylogenetic analysis was conducted on a dataset comprising almost all available Hirudinea mitogenomes (Supplementary Data). PhyloSuite (Zhang et al. 2020) was used to standardize annotation, extract data, and conduct phylogenetic analysis using nucleotide sequences of 13 concatenated protein-coding genes with the help of several

plug-in programs: genes were aligned using the codon mode, and the accurate G-INS-i strategy in MAFFT (Katoh and Standley 2013), concatenated using PhyloSuite, and phylogeny was reconstructed using IQ-tree (Trifinopoulos et al. 2016) with 10,000 ultrafast bootstraps (Minh et al. 2013) and the most parameter-rich evolutionary model GTR+G+I (Supplementary Data). iTOL was used to visualize the phylogeny (Letunic and Bork 2007). Glossiphoniidae was monophyletic, and *H. yangtzenensis* formed a sister lineage to *Glossiphonia concolor* (Figure 1).

Authors' contributions

Ti-Lin Yi and Dai-Qing Yang conceived and designed the project; Zhi-Wei Xu and Meng-Ting Pei acquired, analyzed, and interpreted the data; and Ti-Lin Yi and Dai-Qing Yang co-wrote the paper. All authors agree to be accountable for all aspects of the work.

Disclosure statement

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

Funding

This work was supported by the China Agriculture Research System of MOF and MARA under Grant No. [CARS-46]; Technical Innovation Project of Hubei Province under Grant No. [2019ABA086]; and Key Laboratory Open Project Fund of Engineering Research Center for Ecology and Agriculture of Wetlands, Ministry of Education under Grant No. [KFT201902].

ORCID

Ti-Lin Yi  <http://orcid.org/0000-0003-0033-8426>

Data availability statement

The 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 number MN106285 (<https://www.ncbi.nlm.nih.gov/nuccore/MN106285.1/>).

References

- Burland TG. 2000. DNASTAR's Lasergene sequence analysis software. In: Misener S, Krawetz SA, editors. *Methods in Molecular Biology*TM. Totowa (NJ): Humana Press; p. 71–91.
- Daffe G, Sun Y, Ah Yong ST, Kupriyanova EK. 2021. Mitochondrial genome of *Sabella spallanzanii* (Gmelin, 1791) (Sabellida: Sabellidae). *Mitochondrial DNA Part B*. 6(2):499–501.
- Jiménez-Armenta J, Kvist S, Ocegüera-Figueroa A. 2020. An exceptional case of mitochondrial tRNA duplication-deletion events in blood-feeding leeches. *Org Divers Evol*. 20(2):221–231.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- 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.
- Koperski P. 2017. Taxonomic, phylogenetic and functional diversity of leeches (Hirudinea) and their suitability in biological assessment of environmental quality. *Knowl Manag Aquat Ecosyst*. 418:49.
- Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res*. 29(22):4633–4642.
- Le Marrec-Croq F, Drago F, Vizioli J, Sautière P-E, Lefebvre C. 2013. The leech nervous system: a valuable model to study the microglia involvement in regenerative processes. *Clin Dev Immunol*. 2013:274019.
- Letunic I, Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*. 23(1):127–128.
- Li Y, Kocot KM, Schander C, Santos SR, Thornhill DJ, Halanych KM. 2015. Mitogenomics reveals phylogeny and repeated motifs in control regions of the deep-sea family Siboglinidae (Annelida). *Mol Phylogenet Evol*. 85:221–229.
- Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol*. 30(5):1188–1195.
- Ocegüera-Figueroa A, Manzano-Marín A, Kvist S, Moya A, Siddall ME, Latorre A. 2016. Comparative mitogenomics of leeches (Annelida: Clitellata): genome conservation and placobdella-specific trnD gene duplication. *PLOS One*. 11(5):e0155441.
- Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*. 290(5806):470–474.
- Sket B, Trontelj P. 2008. Global diversity of leeches (Hirudinea) in freshwater. *Hydrobiologia*. 595(1):129–137.
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res*. 44(W1):W232–W235.
- Weigert A, Bleidorn C. 2016. Current status of annelid phylogeny. *Org Divers Evol*. 16(2):345–362.
- Weigert A, Golombek A, Gerth M, Schwarz F, Struck TH, Bleidorn C. 2016. Evolution of mitochondrial gene order in Annelida. *Mol Phylogenet Evol*. 94(A):196–206.
- Xu Z, Yang C, Gofarov MY, Eliseeva TA, Kondakov AV, Yuan H, Bolotov IN, Yang D. 2021. A new freshwater leech species from Asian Swamp Eel stocks in China. *Parasitol Res*. 120(8):2769–2778.
- Ye L, Yao T, Lu J, Jiang J, Bai C. 2021. Mitochondrial genomes of two *Polydora* (Spionidae) species provide further evidence that mitochondrial architecture in the Sedentaria (Annelida) is not conserved. *Sci Rep*. 11(1):13552.
- Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour*. 20(1):348–355.
- Zou H, Jakovlić I, Chen R, Zhang D, Zhang J, Li W-X, Wang G-T. 2017. The complete mitochondrial genome of parasitic nematode *Camallanus cotti*: extreme discontinuity in the rate of mitogenomic architecture evolution within the Chromadorea class. *BMC Genomics*. 18(1):840.