





The complete mitochondrial genome of *Yuukianura szeptyckii* Deharveng & Weiner 1984 (Collembola: Neanuridae)

June Wee^{a*}, Yun-Sik Lee^{b*} , Taekjun Lee^c , Philjae Kim^d , Jino Son^e , Yongeun Kim^b and Kijong Cho^a

^aDivision of Environmental Science and Ecological Engineering, Korea University, Seoul, Republic of Korea; ^bO-Jeong Resilience Institute, Korea University, Seoul, Republic of Korea; ^cMarine Biological Resource Institute, Sahmyook University, Seoul, Korea; ^dDivision of Ecological Conservation, Bureau of Ecological Research, National Institute of Ecology, Choongnam, Korea; ^eBiological and Genetics Resources Assessment Division, National Institute of Biological Resources, Incheon, Republic of Korea

ABSTRACT

The complete mitochondrial genome of *Yuukianura szeptyckii* Deharveng & Weiner 1984 was sequenced, assembled, and annotated. The mitochondrial genome of *Y. szeptyckii* has a length of 15,771 bp and contains 13 protein-coding genes (PCGs), 22 transfer (tRNA) genes, and 2 ribosomal RNA (rRNA) genes. *Y. szeptyckii* was closely clustered with the following species of Neanuridae: *Bilobella aurantiaca* and *Friesea grisea*.

ARTICLE HISTORY

Received 5 November 2020
Accepted 5 February 2021

KEYWORDS

Springtail; gene order; phylogeny; mitogenome

Collembola are the most abundant organisms in the soil ecosystem and play an important role in nutrient cycling (Rusek 1998; Filser 2002). The family Neanuridae is a diverse family of Collembola that are characterized by the absence of the mandibular plate (Hopkin 1997). Approximately 1,600 species have been reported worldwide (Bellinger 1996–2020), but only a few complete mitochondrial genome records belong to the Neanuridea group (NCBI, accessed 2020 Oct 29; Dong et al. 2020). In the present study, sequencing, assembly, and annotation of the mitochondrial genome of *Yuukianura szeptyckii* were performed and their molecular characteristics were described.

Specimens of *Y. szeptyckii* were first collected from near a stream of Ansan City, Gyeonggi province, Korea (37.288°N, 126.822°E) on 15 June 2006. Since then, *Y. szeptyckii* has been cultured in the laboratory for approximately 14 years. The DNA of the sample used in the present study was deposited at the Korea University (<http://insect.korea.ac.kr>, ssamppong@korea.ac.kr, specimen accession number: KUEMCOL002). Mitochondrial DNA was extracted using the Qproteome® Mitochondria Isolation kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and isolated using a QIAamp DNA mini kit (Qiagen). Mitochondrial DNA was amplified using the REPLI-g Mitochondrial DNA Kit (Qiagen). Next-generation sequencing (NGS) analysis was performed using genome analysis units at the National Instrumentation Center for Environmental Management of Seoul National University in Korea. A genomic library was constructed from the genomic DNA using a Kapa Hyper Prep Kit (Kapa Biosystems, Woburn, MA, USA)

using paired-end reading, which was followed by next-generation sequencing (NGS) on the Illumina Hi-Seq 2500 platform (San Diego, CA, USA). The complete mitogenome was annotated using the Geneious ver. 11.1.5 (Biomatters Ltd, Auckland, New Zealand). Transfer RNA (tRNA) genes were identified using tRNAscan-SE online (Lowe and Chan 2016), with the following search mode: the sequence source was 'other mitochondrial', and the genetic code for tRNA isotype prediction was 'Invertebrate Mito'. Mitogenome sequences were aligned using MAFFT (Katoh and Standley 2013) and the mitogenome dataset was analyzed using maximum likelihood (ML) with RAxML 8.2 (Stamatakis 2014). The best-fit substitution was estimated using jModelTest 2.1.1 (Guindon and Gascuel 2003; Darriba et al. 2012) for the nucleotide dataset of 13 protein-coding genes (PCGs). For ML analyses, bootstrap resampling was performed using the rapid option with 1000 iterations.

The mitogenome of *Y. szeptyckii* is 15,771 bp long and contains 13 protein-coding genes (PCGs), 22 transfer (tRNA) genes, and 2 ribosomal RNA (rRNA) genes. The overall nucleotide composition was 34.0% A, 15.6% C, 10.1% G, and 40.3% T, indicating an obvious A + T bias (74.3%). For the 13 PCGs, three start codons are found: ATG (ATP6, COX3, ND4), ATT (ND2) and other nine PCGs has ATA. The 'TAG' stop codon reveal COX3, CYTB, ND5 and ND6, and other nine PCGs has 'TAA'.

The gene arrangement of *Y. szeptyckii* was the same as that of the two other species of the Neanuridea family from GenBank (*Bilobella aurantiaca* (NC_011195) and *Friesea grisea* (NC_010535)); the gene arrangement of Collembola

CONTACT Kijong Cho  kjcho@korea.ac.kr  Division of Environmental Science and Ecological Engineering, Korea University, Seoul, 02841, South Korea
*These authors contributed equally to this work

© 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.

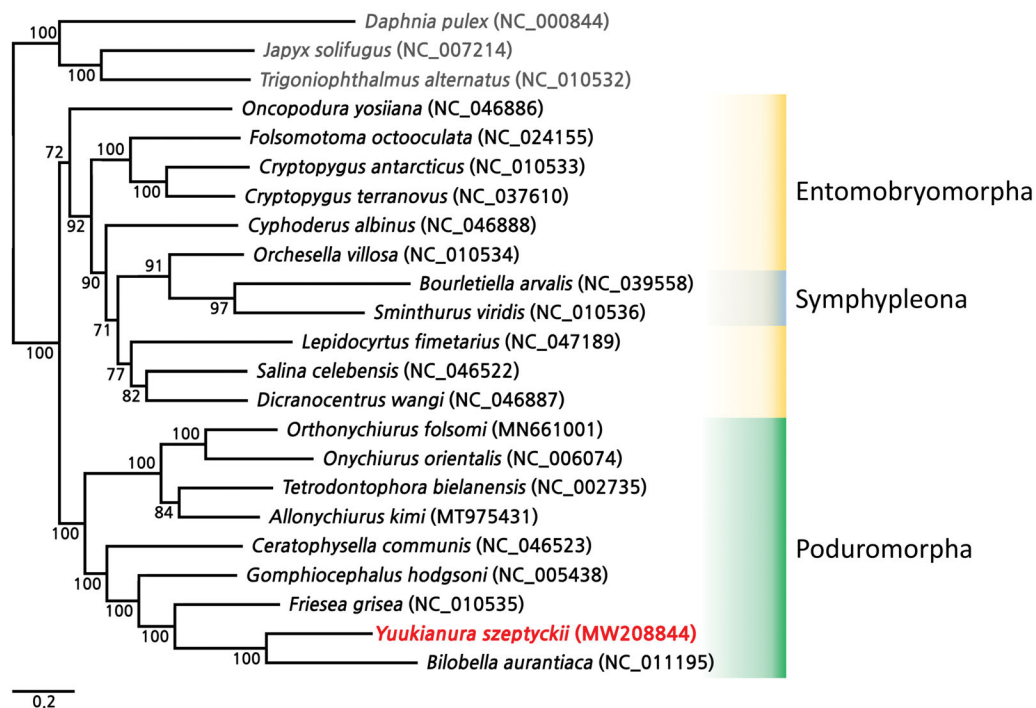


Figure 1. Phylogenetic tree of *Yuukianura szeptyckii* and 19 springtails from GenBank developed using the maximum likelihood method based on the nucleotide sequences of 13 protein-coding genes. The bootstrap support values are indicated on each node.

mitogenomes was relatively conserved compared to other taxa, including *Y. szeptyckii*. In the phylogenetic tree of ML, *Y. szeptyckii* clustered a sister group with closely related species that belonged to the family Neanuridae (*B. aurantiaca* and *F. grisea* (Figure 1)) and to a monophyletic clade of order Poduromorpha (Figure 1).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education [NRF-2018R1D1A1B07047657].

ORCID

Yun-Sik Lee  <http://orcid.org/0000-0002-9328-8490>
 Taekjun Lee  <http://orcid.org/0000-0003-4407-7862>
 Philjae Kim  <http://orcid.org/0000-0001-9938-6895>
 Jino Son  <http://orcid.org/0000-0003-1752-142X>

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 no. MW208844. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA694843, SRR13529482, and SAMN17575579, respectively.

References

- Bellinger PF, Christiansen KA, Janssens F. 1996–2020. Checklist of the Collembola of the World; [accessed 2020 Sep 1]. <http://www.collembola.org>.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 9(8):772
- Dong J, Zhang F, Wang X. 2020. Complete mitochondrial genome of *Pseudachorutes palmiensis* (Collembola: Neanuridae). *Mitochondrial DNA Part B*. 5(1):394–395.
- Filser J. 2002. The role of Collembola in carbon and nitrogen cycling in soil. *Pedobiologia*. 46(3–4):234–245.
- Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 52(5): 696–704.
- Hopkin SP. 1997. *Biology of the springtails (Insecta: Collembola)*. Oxford (UK): Oxford University Press.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res*. 44(W1): W54–W57.
- Rusek J. 1998. Biodiversity of Collembola and their functional role in the ecosystem. *Biodivers Conserv*. 7(9):1207–1219.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.