Characterization and phylogenetic analysis of the complete mitochondrial genome of *Notostomum cyclostomum*

Jun Young Chae, Jin Ho Kim, Tae-Wook Kang, Hyung-Ho Lee and Moo-Sang Kim

**ABSTRACT**

*Notostomum cyclostomum* is a parasite that lays eggs on the snow crab shells and causes various diseases by parasitizing fish. Although there have been some studies on the life history of this parasite and the associated fish diseases, little is known about the molecular biology of this parasite. Thus, here we report the mitochondrial genome of *N. cyclostomum*, which is 16,972 bp long and contains 13 mitochondrial protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, two ribosomal RNA (rRNA) genes, and a putative control region with a 92% AT-rich sequences between *tRNA-R* and *tRNA-H*. Phylogenetic analysis using 13 PCGs confirmed that *N. cyclostomum* belongs to the family Piscicolidae. This is the first study revealing the complete mitochondrial genome sequence of *N. cyclostomum*.

*N. cyclostomum* (Johansson 1898) is a parasite belonging to *Piscicolidae* and inhabits the East Sea of Russia (Epshtein 1996) and the northeastern Pacific Ocean of British Colombia Canada, and Alaska USA (Sloan et al. 1984; Khan and Paul 1995). *Notostomum cyclostomum* lays eggs on crustacean carapace like snow crab (*Chionoecetes opilio*) and hatches there before reaching fish. After hatching and reaching to the juvenile stage, it seeks fish for parasitizing on blood. *Notostomum cyclostomum* doesn’t affect snow crab, but as a vector it mediates infection of Cryptobia sp., one of hemiaflagellate fish parasites (Meyers and Burton 2009). Although having a unique life cycle, little is known about genetic and molecular biology of *N. cyclostomum*. The mitogenome of *Piscicolidae* is still uncovered and only that of Zeylanicodellara aguramensis has been reported (Wang et al. 2018). In this study, we identified the complete mitochondrial DNA (mtDNA) sequence of *N. cyclostomum*.

The *N. cyclostomum* was obtained from the carapace of snow crab (*Chionoecetes opilio*) purchased at a fish market in Yeongdeok-gun (36°97’N, 130°10’E), South Korea, and deposited at the Pukyong National University (Voucher no. PKNU_2021_001: Jun Young Chae, jychae@pukyong.ac.kr). Genomic DNA (gDNA) was extracted from whole body of *N. cyclostomum* using BeadSTM Genomic DNA Prep Kit for Animal Tissue (Biofact, Republic of Korea) according to the manufacturer recommendations. PCR was performed on *cox1* gene using the invertebrate universal primer set (Folmer et al. 1994), LCO1490 (5’-GGTCAAAAATCAATATTTG-G-3’), HCO2198 (5’-TAAACTT-CAGGGTACAAAATCA-3’). PCR reaction conditions were as follows: Pre-denaturation at 95 °C for 2 min, 30 cycles of 20 sec at 95 °C, 20 sec at 40 °C, and 45 sec at 72 °C and a final extension for 5 min at 72 °C. The *cox1* sequence was used to identify the species using the Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI). An NCBI BlastN search (Johnson et al. 2008) of *cox1* sequence showed 99.25% identity to *cox1* sequence (DQ414327.1) of the *N. cyclostomum*, confirming that the parasite used in this study is *N. cyclostomum*.

The genomic library for next-generation sequencing was prepared with the extracted gDNA (1 μg) using MGIEasy DNA library prep kit (MGi, China). Sequencing was conducted using MGISEQ-2000 (MGI, China) with 150 bp paired-end reads. Raw data were deposited in Sequence Read Archive (SRA) database (SRR19136323). The raw data were trimmed with Trimmomatic ver. 0.39 (Bolger et al. 2014), and a contig sequence was produced by default option in the de novo assembler of the CLC Genomics Workbench ver. 20.04 (QIAGEN). The circular form of mitogenome was confirmed by using ‘Map to Reference’ tool of Geneious software ver. 2021.2.2 (https://www.geneious.com) from mapping the filtered data into the contig sequence, and finally showed a 16,972 bp long complete mtDNA sequence of *N. cyclostomum*.

The complete mtDNA sequence was used to obtain basic annotation information using the MITOS WebServer (Bernt, Donath, et al. 2013), and the detailed annotation was manually corrected using SnapGene software (ver. 5.3.2, GSL Biotech LLC, snapgene.com; Wang et al. 2018). The complete mtDNA genome constituted of 13 protein-coding genes...
(PCGs), 22 tRNA genes, two rRNA genes, and putative control AT-rich region. All PCGs were encoded on the positive strand, and started with an ATG codon. The rest 12 PCGs, except for nad4l having standard TAA codon, were stopped with the truncated codons T—. The genome showed a total of 22 tRNAs including two tRNA-L and two tRNA-S genes. Predicted secondary structure of 14 tRNAs among these tRNAs showed the standard cloverleaf structure, however tRNA-E was with missing T-arm, and six tRNAs (tRNA-F, tRNA-G, tRNA-H, tRNA-P, tRNA-V, and tRNA-Y) and one tRNAs (tRNA-S1) were with missing loop structure in T-arm and D-arm, respectively. Small rRNA with a 739 bp length was located in between tRNA-M and tRNA-V and 1137 bp long large rRNA was observed in between tRNA-V and tRNA-L1. The putative control region was 2602 bp long, located in between tRNA-R and tRNA-H and possessed AT-rich region with 92% composition. This high AT composition was also observed in Zeylanicobdella arugamensis, the closest species among the reported mtDNA sequences (Wang et al. 2018).

The phylogenetic tree was performed within the scope of class Clitellata with 15 available mitogenomes including Lumbricus terrestris (U24570) and Aporrectodea rosea (NC_047733) as outgroups. GenBank accession numbers of mitogenome sequences are given adjacent to the species names. Node numbers indicate posterior probabilities from the Bayesian inference. The black arrow indicates the N. cyclostomum analyzed in this study.

![Figure 1. Phylogenetic tree of Notostomum cyclostomum and related species. Phylogenetic tree based on the nucleotide sequence of 13 mitochondrial PCGs using Bayesian inference (BI) was conducted on 15 available mitogenomic sequences from the class Clitellata including Lumbricus terrestris (U24570) and Aporrectodea rosea (NC_047733) as outgroups. GenBank accession numbers of mitogenome sequences are given adjacent to the species names. Node numbers indicate posterior probabilities from the Bayesian inference. The black arrow indicates the N. cyclostomum analyzed in this study.](image_url)

**Acknowledgments**

We greatly appreciate Dr. Khawaja Muhammad Imran Bashir for revising the English language of this manuscript.

**Ethical approval**

There is no human or animal involvement in the study. Since the sample is an insect, there is no need for ethical approval or permission to collect the sample.

**Authors contributions**

Jun Young Chae, Hyung-Ho Lee, and Moo-Sang Kim conceived original idea. Jun Young Chae carried out experiments. Jun Young Chae wrote manuscript with support by Moo-Sang Kim and Hyung-Ho Lee. Tae-Wook Kang and JinHo Kim analyzed experimental data. 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 research was supported by Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries [Project No. 20200425-04].
Data availability statement

The genome sequence data that supported the findings of this study are openly available in GenBank of NCBI at (https://www.ncbi.nlm.nih.gov/) under the accession no. ON312494. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA836134, SRR19136323 and SAMN28132021, respectively.

References

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 69(2): 313–319.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15):2114–2120.

Epshtein VM. 1996. The geographical distribution and hosts of the Notostomum leeches (Hirudinea, Piscicolidae). Vest. Zool. 30(3):26–31.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol. 3(5): 294–299.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 17(8):754–755.

Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36(Web Server issue):W5–W9. doi:10.1093/nar/gkn201.

Khan RA, Paul AJ. 1995. Life cycle studies on Arcto-boreal Leeches (Hirudinea). J. Helminthol. Soc. Wash. 62(2):105–110.

Meyers TR, Burton T. 2009. Diseases of wild and cultured shellfish in Alaska. Juneau: Alaska Department of Fish and Game, Fish Pathology Laboratories.

Sloan NA, Bower SM, Robinson SMC. 1984. Cocoon deposition on three crab species and fish parasitism by the leech Notostomum cyclostoma from deep fjords in northern British Columbia. Mar Ecol Prog Ser. 20(1):51–58.

Wang Y, Huang M, Wang R, Fu L. 2018. Complete mitochondrial genome of the fish leech Zeylanicobdella arugamensis. Mitochondrial DNA B Resour. 3(2):659–660.