Data

The Complete Mitochondrial Genome of the Fourhorn Sculpin *Triglopsis quadricornis* (Perciformes, Cottidae) from Sirius Passet, North Greenland

Bo-Mi Kim[1*], Ji-Hoon Kihm[2,3], and Tae-Yoon S. Park[2,3]

[1]*Research Unit of Cryogenic Novel Material, Korea Polar Research Institute, Korea Institute of Ocean Science & Technology, Incheon 21990, Korea*
[2]*Division of Earth Sciences, Korea Polar Research Institute, Korea Institute of Ocean Science & Technology, Incheon 21990, Korea*
[3]*Polar Science, University of Science & Technology, Daejeon 34113, Korea*

Abstract : *Triglopsis quadricornis* Linnaeus, 1758 (Cottidae) is distributed in the Atlantic and Arctic and has four unique bony protuberances on its head. Here, we report the complete, circular, and annotated mitochondrial genome of *T. quadricornis*. The complete *T. quadricornis* mitochondrion was sequenced by high-throughput Illumina HiSeq platform. The sequences are 16,736 bp in size and contains 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, a control region, and large and small ribosomal subunits. The overall genomic structure of *T. quadricornis* mitochondrion was conserved with the gene arrangement of *Megalocottus* and *Myoxocephalus* species, and phylogenetic analysis supports their sister relationships. Most PCGs consist of TAA or TAG as a termination codon, whereas *COII*, *ND4*, and *CYTB* have T-- as a stop codon. This complete mitochondrial DNA information of *T. quadricornis* will provide an essential genomic resource to elucidate the phylogenetic relationship and evolutionary history of the family Cottidae.

Key words : *Triglopsis quadricornis*, Greenland fish, Cottidae, mitogenome, phylogeny

1. Introduction

Sculpins as common benthic fish are distributed in most oceans of the world (Nelson 2006; Johannesen et al. 2012). The Cottidae is the largest family of sculpins, which comprise more than 70 genera and 275 species worldwide (Nelson 2006). They mainly occur in boreal and cold-temperate regions across the Arctic (Nelson 2006). Although sculpins are not normally deemed a suitable food source for higher tropic levels (e.g., fish, seabirds, and mammals) in the Arctic, they probably form an important resource contributing to food web stability, considering their abundance and availability in dynamically-changing environments in the Arctic (e.g., seasonal ice cover and open water) (Rooney et al. 2006; Walsh 2008). The genus *Triglopsis* has a wide distribution in the northern hemisphere with noticeable variations in morphology. Fourhorn sculpin *Triglopsis quadricornis* Linnaeus, 1758 is observed in nearshore open water in the Arctic (COSEWIC 2003) However, little is known about their general biology (e.g., reproduction, feeding strategies), habitats, population sizes or trophic roles.

In this study, we present the complete mitogenome of *T. quadricornis* (Accession no. MT789711; registered with homotypic synonym *Myoxocephalus quadricornis*) by *de novo* assembling from Illumina shotgun sequence data. An individual of *T. quadricornis* was sampled at J. P. Koch Fjord near Sirius Passet, North Greenland (82°46' 1.7"N, 42°32'52.4"W; Lee 2018) on 16th July 2017. The voucher specimen was registered in the Korea Polar Research Institute (KOPRI; Species ID: SP-fish; Specimen ID: 170716_ SP-fish01). Genomic DNA was isolated from muscle tissue by using a QIAamp DNA Blood Mini kit (Qiagen, Hilden, Germany). Quantity and quality of the extracted genomic DNA were

*Corresponding author. E-mail: bomikim@kopri.re.kr

measured using a Qubit 4 Fluorometer (Thermo Fisher Scientific, Inc., Waltham, MA, USA). A genomic library was constructed using TruSeq Nano DNA Kit by a commercial company (Macrogen, Seoul, South Korea) according to the manufacturer's instructions (Illumina, San Diego, CA, USA). The sequencing library was constructed by random fragmentation of the genomic DNA with 5' and 3' adapter ligation. The ligation product was amplified by PCR before loading into a flow cell. Raw reads were obtained from the sample that passed quality control by Illumina HiSeq platform. Trimming and filtering were performed using FastP and Trimmomatics, and 24,643,780 reads were finally obtained. After the trimming process, *de novo* assembly was performed with SPAdes v3.11.1 (Bankevich et al. 2012). Genomic features and annotations were performed using MITOS2 (Bernt et al. 2013) and tRNAscan-SE 2.0 (Lowe and Eddy 1997). The annotated gene structure was further confirmed using NCBI- BLAST (http://blast.ncbi.nlm.nih.gov).

The mitochondrial genome content of *T. quadricornis* was found to generally include a typical set of 13 PCGs, 22 tRNAs, and the large and small rRNA of the ribosome as observed in Perciformes species (Mascolo et al. 2019; Ceruso et al. 2020). The full circular contig was 16,736 bp in length with a GC content of 46.7% and an AT bias (A: 26.8%; T: 26.4%; G: 17.4%; C: 29.3%). Recently, the complete mitochondrial genomes of two individuals of *M. quadricornis* (Accession no. MT303953 and MT303954) collected from the White-Sea were released (Balakirev et al. 2020b). Although the two *M. quadricornis* mitochondrions have a length of 16,682 bp, identity between the individuals was 99.8% (37/16,682 bp). When we compared *T. quadricornis* mitochondrion to the two mitogenomes of *M. quadricornis*, slight differences (0.4~0.9%) in similarity were observed in *ND1*, *ND2*, *ND5*, and *CYTB* genes. Overall, gene composition, arrangement, and other features (e.g., size, GC ratio) are highly similar to the mitogenomes registered in the genus
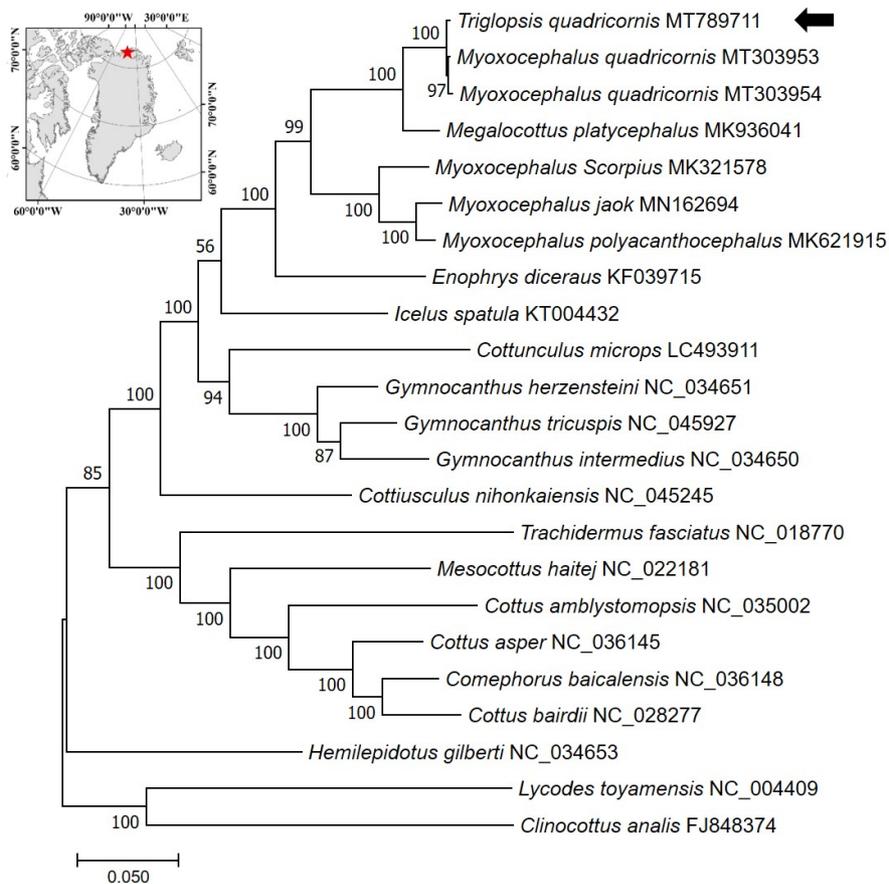

Fig. 1. Maximum-likelihood (ML) phylogeny of 20 species of the family Cottidae based on the concatenated nucleotide sequences of entire protein-coding genes (PCGs). Numbers on the branches indicate ML bootstrap percentages (1000 replicates). DDBJ/EMBL/Genbank accession numbers for published sequences are incorporated. The black arrow indicates the *T. quadricornis* analyzed in this study

Myoxocephalus and *Megalocottus* (Balakirev et al. 2020a; Li et al. 2018). In the *T. quadricornis* mitogenome, *COI* gene has GTG as a start codon and three genes, *COII*, *ND4*, and *CYTB* have an incomplete stop codon (T--), as shown in the genus *Myoxocephalus* (Li et al. 2018).

A maximum likelihood phylogenetic hypothesis was established using sequence data from the concatenated set of 13 PCGs of *T. quadricornis* mitogenome with including 21 published mitogenomes from the Cottidae. JModelTest ver. 2.1.10 (Darriba et al. 2012) was used to select the best substitution model and a substitution model (HKY + G + I) was applied to perform a maximum-likelihood (ML) method in the PhyML 2.4.5 (Guindon and Gascuel 2003) with 1000 bootstrap replicates. The mitogenome of *T. quadricornis* in this study is clustered together with *Megalocottus* and *Myoxocephalus* species (Fig. 1). In addition, our *COI* sequence was identical to the partial *COI* sequences of *T. quadricornis* (652 bp; GenBank accession no. EU524913) and *M. quadricornis* (648 bp; GenBank accession no. MG421040), whereas similarity between *T. quadricornis* in this study and *Megalocottus platycephalus* (1,009 bp; GenBank accession no. MH016197) was 96%. Complete mitochondrial genomes have been recently released from two local specimens of *Myoxocephalus quadricornis* (probably synonym of *Triglopsis quadricornis*) (MQI1, GenBank accession no. MT303953.1 and MQI3, MT303954.1), which were collected from Dvina Bay, White Sea, Russia. Difference of the whole mitochondrial DNA sequence between MQI1 and MQI3 turns out to be only 0.2%, while those of the both White Sea specimens differ from the Greenland specimen in this study appears by 0.6%. The differences in mitochondrial DNA sequence, therefore, would have been derived by different geographical distributions, but additional mitogenome data are required to understand the differences in mitogenomes in this genus.

2. Data Availability Statement

Mitogenome data supporting this study are openly available in GenBank at: https://www.ncbi.nlm.nih.gov/nuccore/MT789711. Associated BioProject, SRA, and Bio Sample accession numbers are https://www.ncbi.nlm.nih.gov/bioproject/PRJNA673747, https://www.ncbi.nlm.nih.gov/sra/SRR12964012, and https://www.ncbi.nlm.nih.gov/biosample/SAMN16622464, respectively.

References

Balakirev ES, Kravchenko AY, Cherepkova EV, Saveliev PA, Semenchenko AA, Ayala FJ (2020a) Complete mitochondrial genome of the Belligerent sculpin *Megalocottus platycephalus* (Cottoidei: Cottidae). Mitochondrial DNA B Resour **4**: 2980‒2981

Balakirev ES, Kravchenko AY, Semenchenko AA (2020b) Genetic evidence for a mixed composition of the genus *Myoxocephalus* (Cottoidei: Cottidae) necessitates generic realignment. Genes **11**:1071

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455‒477

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**:313‒319

Ceruso C, Venuti I, Osca D, Caputi L, Anastasio A, Crocetta F, Sordino P, Pepe T (2020) The complete mitochondrial genome of the sharpsnout seabream *Diplodus puntazzo* (Perciformes: Sparidae). Mitochondrial DNA B Resour **5**: 2379‒2381

COSEWIC (2003) COSEWIC assessment and update status report on the fourhorn sculpin *Myoxocephalus quadricornis* (freshwater form) in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa, 24 p

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods **9**:772

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52**:696‒704

Johannesen E, Høines ÅS, Dolgov AV, Fossheim M (2012)

Demersal fish assemblages and spatial diversity patterns in the Arctic-Atlantic transition zone in the Barents Sea. PLoS ONE **7**:34924. doi:10.1371/journal.pone.0034924

Lee WY (2018) Bird observations in Sirius Passet, North Greenland, 2016 and 2017. Dansk Orn Foren Tidsskr **112**:53‒58

Li Y, Song P, Feng J, Zhang N, Zhang R, Lin L (2018) Complete mitochondrial genome sequence and phylogenetic analysis of *Myoxocephalus scorpius* (Linnaeus, 1758). Mitochondrial DNA B Resour **4**:862‒863

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25**:955‒964

Mascolo C, Ceruso M, Chirollo C, Palma G, Anastasio A, Sordino P, Pepe T (2019) The complete mitochondrial genome of the Angolan dentex *Dentex angolensis* (Perciformes: Sparidae). Mitochondrial DNA B Resour **4**:1245‒1246

Nelson JS (2006) Fishes of the world. John Wiley & Sons, Inc., New Jersey, 752 p

Rooney N, McCann K, Gellner G, Moore JC (2006) Structural asymmetry and the stability of diverse food webs. Nature **442**:265‒269

Walsh JE (2008) Climate of the arctic marine environment. Ecol Appl **18**:3‒22

Author's Information

Bo-Mi Kim
Senior Research Scientist, Korea Polar Research Institute

Ji-Hoon Kihm
Postdoctoral Researcher, Korea Polar Research Institute

Tae-Yoon S. Park
Principal Research Scientist, Korea Polar Research Institute