

Sequencing of Two Mitochondrial Genomes of Endangered Form of the Sevan Trout

Salmo ischchan aestivalis

Artem V. Nedoluzhko^{a*}, Evgeniy Simonov^{b,c}, Sergey M. Rastorguev^a, Eugenia S. Boulygina^a,

Fedor S. Sharko^d, Svetlana V. Tsygankova^a, Van Quan Nguyen^e, Bardukh K. Gabrielyan^f,

Haikaz R. Roubenyan^f, and Boris A. Levin^{b,g}

^a*National Research Centre “Kurchatov Institute”, 123182 Moscow, Russia*

^b*Papanin Institute for Biology of Inland Waters, Russian Academy of Sciences, Laboratory of*

Fish Ecology, 152742, Borok, Yaroslavl Region, Russia

^c*Siberian Federal University, 660041, Krasnoyarsk, Russia*

^d*Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of*

Sciences, 119071, Moscow, Russia

^e*Institute of Marine Environment and Resources, 246 Danang Street, Hai Phong, Vietnam.*

^f*Scientific Center of Zoology and Hydroecology, National Academy of Sciences of Republic of*

Armenia, Yerevan, Armenia

^g*Cherepovets State University, 162600, Cherepovets Vologda Region, Russia*

* Corresponding author: Tel. +74991967937 E-mail: nedoluzhko@gmail.com

Acknowledgments

The authors are grateful to Mikhail V. Kovalchuk (National Research Centre “Kurchatov Institute”, Moscow, Russia) for his ongoing support. This work is supported by the National Research Centre “Kurchatov Institute” (DNA extraction, libraries preparation, NGS-sequencing, and bioinformatics) and partially by the Russian Science Foundation, grant no. 15-14-10020 (sampling and bioinformatics).

Abstract

The two complete mitochondrial genomes of endangered form of the Sevan trout *Salmo ischchan aestivalis* are published in this paper. The mitochondrial DNA (mtDNA) is 16,677 base pairs (bp) in length and contained 13 protein-coding genes, 2 rRNA genes, 22 tRNA genes. The overall base composition of the genome in descending order was 29.4% - C, 27.9% - A, 26.0% - T, 16.7% - G, without a significant AT bias of 53.9%.

Keywords

Mitochondrial genome, endangered species, *Salmo ischchan aestivalis*, Sevan trout, Illumina sequencing

Main text

DNA samples of the Sevan trout, endangered subspecies *Salmo ischchan aestivalis*, were obtained from two individuals: Li4 individual was collected from local aquaculture farm near Djermuk town on 2012 (39.7305 N and 45.5950 E) and Li30 was collected from the natural population, the Makenis River, tributary of Lake Sevan, on 06 May 1974 (40.1771 N and 45.5841 E). For the latter specimen we used historical scales from scale book, a DNA source that is appropriate for mitogenome sequencing by NGS technology .

DNA was isolated from caudal fin (Li4) and dried scales (Li30) in the DNA facilities of the National Research Center “Kurchatov institute” (Moscow, Russia) using a standard method of DNA extraction from animal tissue (phenol-chloroform). The concentration of DNA in the samples was measured using a Qubit fluorimeter (Thermo Fisher Scientific, USA).

Two DNA libraries were prepared using an Ovation Ultralow Systems V2 kit (NuGEN, San Carlos, CA). Mitochondrial genome was sequenced used Illumina Hiseq 2500 (Illumina, USA) with 150 bp paired-end reads.

33,981,829 and 144,017,525 Illumina paired-end reads were generated for DNA library of Li4 sample and Li30 sample, respectively. Illumina reads from Li30 sample were used for building mtDNA sequence de-novo by Norgal software package , but for Li4 library, there were not enough reads for building whole mitochondrial genome continuous contig. The Li4 reads were mapped to the Li30 mitochondrial genome using the bowtie2 software version 2.2.3 with --very-sensitive-local preset options with --very-sensitive-local preset options .

As result, the mitogenome of *S. ischchan aestivalis* consists of 16,677 bp (GenBank accession numbers Li4: MG599465 and Li30: MG599466) and includes 13 protein coding genes (PCGs), 2 rRNA genes and 22 tRNA genes.

Eleven of the 13 PCGs (*NAD4*, *NAD5*, *NAD4L*, *NAD3*, *COB*, *NAD1*, *NAD2*, *COX2*, *ATP8*, *ATP6*, *COX3*) used ATG as start codon, another one (*COX1*) used GTG and *NAD6* used ATA. Twelve genes (*NAD1*, *NAD2*, *COX1*, *COX2*, *ATP8*, *ATP6*, *COX3*, *NAD3*, *NAD4L*, *NAD4*, *NAD5* and *COB*) ended with a TAA stop codon, but for three ones of them (*COX2*, *NAD4* and *COB*) TAA stop codon is completed by the addition of 3' A residues to the mRNA, *NAD6* gene ended with a TAG stop codon.

The phylogenetic analysis for whole mithogenome sequences was performed for the *S. ischchan aestivalis* and other Salmonidae species: *S. ischchan danilewskii* (MG599465 and MG599466), *S. trutta fario* (LC137015); *S. trutta* (MF621760); *S. trutta* (MF621762); *S. salar* (JQ390055); *S. salar* (JQ390056) and *Oncorhynchus kisutch* (MF621749) (Figure 1). Sequences were aligned using multiple sequence alignment program Muscle 3.8.31 . All gaps and poorly aligned positions were removed using Gblocks 0.91b , resulting in 16,511 bp length alignment. The phylogenetic relationships were reconstructed using the maximum likelihood (ML) method in the PhyML 2.4.5 . The best substitution model (averaged for whole mitogenome) was chosen in the jModelTest 2.1.10 on the basis of the Bayesian information criterion (BIC). According to jModelTest, the best model describing the evolution

of the mitogenomes was TVM+G ($-\ln L = 34659.93$, BIC = 69611.21), and therefore, it was used for ML analysis (Figure 1).

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

References

- Al-Nakeeb K, Petersen TN, Sicheritz-Ponten T. 2017. Norgal: extraction and de novo assembly of mitochondrial DNA from whole-genome sequencing data. *BMC Bioinformatics*. Nov 21;18.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*. Aug;9:772-772.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. Mar;32:1792-1797.
- Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biol*. Oct;52:696-704.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. Apr;9:357-U354.
- Nedoluzhko AV, Rastorguev SM, Simonov E, Boulygina ES, Sharko FS, Tsygankova SV, Gabrielyan BK, Roubenyan HR, Levin BA. 2018. Two complete mitochondrial genomes of extinct form of the Sevan trout *Salmo ischchan danilewskii*. *Mitochondrial DNA Part B*. 2018/01/02;3:40-41.
- Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biol*.56:564-577.