JOINT INSTITUTE FOR NUCLEAR RESEARCH
Dzhelepov Laboratory of Nuclear Problems

FINAL REPORT ON THE START PROGRAM

Identification of sponge species in Lake Sevan based on ITS sequence analysis

Supervisor:
Dr. Elena Kravchenko

Student:
Ksenia Gavrilenko, Russia,
Irkutsk State University

Participation period:
July 06 – August 16,
Summer Session 2025

Dubna, 2025

Abstract

The paper presents a genetic analysis of sponges collected in Lake Sevan (Armenia). The ITS region of the ribosomal DNA was amplified using PCR, cloned, and sequenced. Comparison of the obtained sequences with the NCBI database (BLAST) allowed the species to be identified as Ephydatia fluviatilis. Based on these data, a phylogenetic tree was constructed to reflect the position of the species among other representatives of the Porifera. The results confirm the effectiveness of ITS sequencing for sponge taxonomy and indicate the presence of E. fluviatilis in Lake Sevan.

Keywords: *Ephydatia fluviatilis*, Lake Sevan, ITS region, Porifera, PCR, cloning, DNA sequencing, phylogeny, NCBI BLAST.

Introduction

Sponges are found all over the world. There are about 8,000 species that live in fresh and salt water bodies, of which only 238 species are freshwater [6].

A sponge is a fairly simple organism that does not have organs or tissues, but at the same time has different types of cells that can transform into each other, which allows sponges to quickly regenerate.

The sponge filters and purifies water from organic substances. This is the reason for its very important role in aquatic ecosystems. It plays an important role in the food chain, digesting a huge number of both single-celled organisms and organic compounds, but the sponge is also a permanent or temporary habitat for many organisms that have entered into symbiotic relationships with it. The sponge community includes various microorganisms - fungi, algae, archaea, bacteria and viruses [7].

Purpose of the work: Conducting a comprehensive molecular genetic analysis of freshwater sponges (Porifera) of Lake Sevan is of significant scientific interest due to the uniqueness of this high-mountain ecosystem. The main objective of this work was to accurately determine the species affiliation of sponges inhabiting Lake Sevan using modern DNA analysis methods. Particular attention was paid to the characterization of the ITS region of ribosomal DNA as the most informative molecular marker for solving taxonomic problems in sponges.

This goal defined the following tasks:

1) Select optimal conditions for PCR amplification of the ITS region, including the choice of primers and amplification temperature conditions.

2) Clone the PCR products into the pAL2-T plasmid vector followed by sequencing to obtain high-quality sequences for further analysis.

3) Process the obtained data in two stages using bioinformatic

methods:

- Alignment and editing of sequences using specialized software.
- Construction of phylogenetic trees.

Relevance:

Features of sponge biology, in particular their exceptional morphological plasticity and low level of organism integrity, significantly complicate their classification based on traditional morphological features. Wide intraspecific variability and interspecific similarity make morphological criteria unreliable. In this regard, the use of molecular genetic methods becomes a necessary tool for refining the taxonomy and phylogeny of these organisms.

Successful experience of using molecular markers in studying marine sponges demonstrates their advantages over morphological analysis, especially in cases where external signs do not allow for unambiguous conclusions [8].

Thus, a comprehensive study of the sponges of Lake Sevan using molecular genetic methods will help clarify their taxonomic status.

Materials and methods of research

PCR amplification

The PCR method is based on the principle of artificially doubling DNA fragments, which allows the target section of genetic material to be isolated and studied. This process occurs with the help of the enzyme DNA polymerase, which synthesizes new copies of DNA based on the existing sample [10].

To perform PCR, it is necessary to prepare a MasterMix mixture (Table 1).

Table 1: Chemical composition of the MasterMix mixture.

Reagents	Per 1 sample (µl)	For 5 samples (µl)	For 11 samples (µl)
Primer mix: forward and reverse	5.5	1	12.1
Dilluent	55	10	121
H2O	44	8	96.8

Then, 19 µl of MasterMix and 1 µl of DNA from the samples are added to the test tubes with dehydrated DNA polymerase. After mixing all the reagents, the test tubes are shaken and placed in the amplifier, where PCR is carried out according to the established program in several stages. The first stage is denaturation, during which the DNA double helix is broken into two strands under the influence of high temperature (95 °C). Next comes primer annealing: the temperature is reduced to 55 °C, and special short DNA sections (primers) are attached to the target fragments. During the elongation stage, DNA polymerase synthesizes new DNA fragments, doubling the target section. These cycles are repeated many times, which leads to an exponential increase in the number of DNA copies. In laboratory conditions, PCR usually includes dozens of cycles, providing a sufficient

volume of material for subsequent analysis.

Electrophoresis

When working with DNA, one often encounters a mixture of molecules of different lengths – for example, after treatment with restriction enzymes. Since DNA in solution carries a negative charge, such fragments can be separated by size using gel electrophoresis [1]. In an electric field, the molecules move toward the anode, with shorter fragments moving faster and longer ones moving slower due to the resistance of the gel matrix. After electrophoresis, distinct bands are formed in the gel, corresponding to fragments of a certain length, and comparison with a marker allows us to determine their size.

For electrophoresis, 1% agarose gel was prepared: 0.3 g of agarose was dissolved in 30 ml of TAE buffer (1X), heated in a microwave oven until completely melted, and ethidium bromide was added (final concentration 0.5 μg/ml). Warm agarose was poured into a cuvette, a comb was inserted, left for 30 minutes to solidify, after which it was carefully removed. The gel was placed in an electrophoresis chamber filled with TAE buffer (1X). DNA samples were mixed with dye (6x TriTrack DNA Loading Dye) and added to the wells, simultaneously loading 2 μl of GeneRuler DNA Ladder Mix marker. Electrophoresis was carried out at a voltage of 1–10 V/cm of the gel. After completion of the process, the gel was analyzed under UV light on a transilluminator, and the target DNA fragments were excised with a scalpel.

Ligation

Ligation was performed using the Quick-TA kit according to the standard protocol. The lyophilized pAL2-T vector was resuspended in 25 μl of deionized water to a final concentration of 50 ng/μl. PCR products were pre-purified from impurities using a column kit.

The ligation reaction was carried out for 14-16 h at 14°C in a mixture

containing 1 µl vector (50 ng), purified PCR product at a vector: insert molar ratio of 1:3 for single amplicons or 1:8-10 for amplicon mixtures, 1 µl 10X Overnight ligation buffer, 1 µl Quick-TA T4 DNA ligase (200 U/µl), and deionized water to a final volume of 10 µl. After incubation, the ligate was stored at -20°C until transformation. Transformation was performed into competent E. coli XL1-Blue cells using 5-10 µl ligate per 100 µl *E. coli* cell suspension.

Preparation of Petri dishes and seeding of transformants

Petri dishes were sterilized in a dry-heat incubator at 115°C for 20 minutes. The nutrient medium was prepared on the basis of agar with the addition of ampicillin (100 µg/ml). After autoclaving, the medium was cooled to 50°C and poured into sterile dishes in a layer of 3–4 mm. Solidified agar was used for seeding: the ligated sample was applied to the surface with a sterile spatula and evenly distributed to obtain isolated colonies. The dishes were incubated in an inverted position at 37°C for 12–16 hours.

Cultivation and selection of recombinant clones of *E. coli*

Only white colonies corresponding to clones with successful integration of the target insert into the pAL2-T plasmid were selected for cultivation. Blue colonies (non-recombinant) were excluded from the analysis, since they contained a vector without an insert, which was due to the restoration of the functional lacZ gene in the presence of X-gal. The selected white colonies were transferred with sterile plastic loops into tubes with 5–7 ml of liquid medium with ampicillin. The tubes were incubated at 37°C for 12–16 hours on a shaker (200 rpm) at an angle of 45° for optimal aeration. This regime ensured active culture growth and accumulation of target recombinant plasmids for subsequent DNA extraction.

Isolation of plasmid DNA

Bacterial culture (1–2 ml) grown in liquid medium with ampicillin was centrifuged at 1700 g for 1 min to collect the cells. The supernatant was discarded, the pellet was resuspended in 250 µl of buffer P1 (with RNase), after which 250 µl of buffer P2 were added and gently mixed until a viscous lysate was formed. For neutralization, 350 µl of cold buffer N3 was added, mixed until a homogeneous pellet appeared and centrifuged at 13,000 g for 10 min. The supernatant was transferred to a spin column, centrifuged at 7000 g (1 min), then washed sequentially with 700 µl of buffer PB and 750 µl of buffer PE. DNA was eluted with 30 µl of deionized water.

Sanger sequencing

The reaction was carried out in a mixture containing:

1) Specific primer - a short single-stranded DNA complementary to the beginning of the target region (serves as a "primer" for DNA polymerase);

2) Standard deoxynucleotides (dATP, dCTP, dGTP, dTTP);

3) Fluorescently labeled dideoxynucleotides (ddNTPs), the random inclusion of which in the growing chain stops synthesis.

The reaction products were separated by capillary electrophoresis. DNA fragments were detected in real time by fluorescence of ddNTP labels. The data were analyzed by software, obtaining a chromatogram with peaks of four colors (A, C, G, T) (Fig. 1). The method provides reading of sequences 500–1000 nucleotides long in one run [9].

Fig. 1. Example of a chromatogram

Bioinformatics analysis

Sequencing data were processed in BioEdit [3]: chromatograms were visually assessed, low-quality regions were corrected, and artifacts were removed. Complete ITS sequences were reconstructed by assembling contigs from forward/reverse reads.

For taxonomic identification, BLASTn (NCBI) [2] was used, selecting reference sequences with similar length and variable homology (85–100%). The selected sequences were combined into a multifile for analysis.

Multiple alignment was performed on the MAFFT [5] server. Phylogenetic trees were constructed in IQ-TREE [4] with automatic selection of the evolutionary model (BIC or AIC criterion) and node support assessment (using the ultrafast bootstrap support method). Visualization was performed in FigTree, where the display of branches was configured and node support values were noted.

Results

Based on the alignment of ITS sequences, a phylogenetic tree was reconstructed (Fig. 2), determining the taxonomic position of the studied sponge (Sponge_sample).

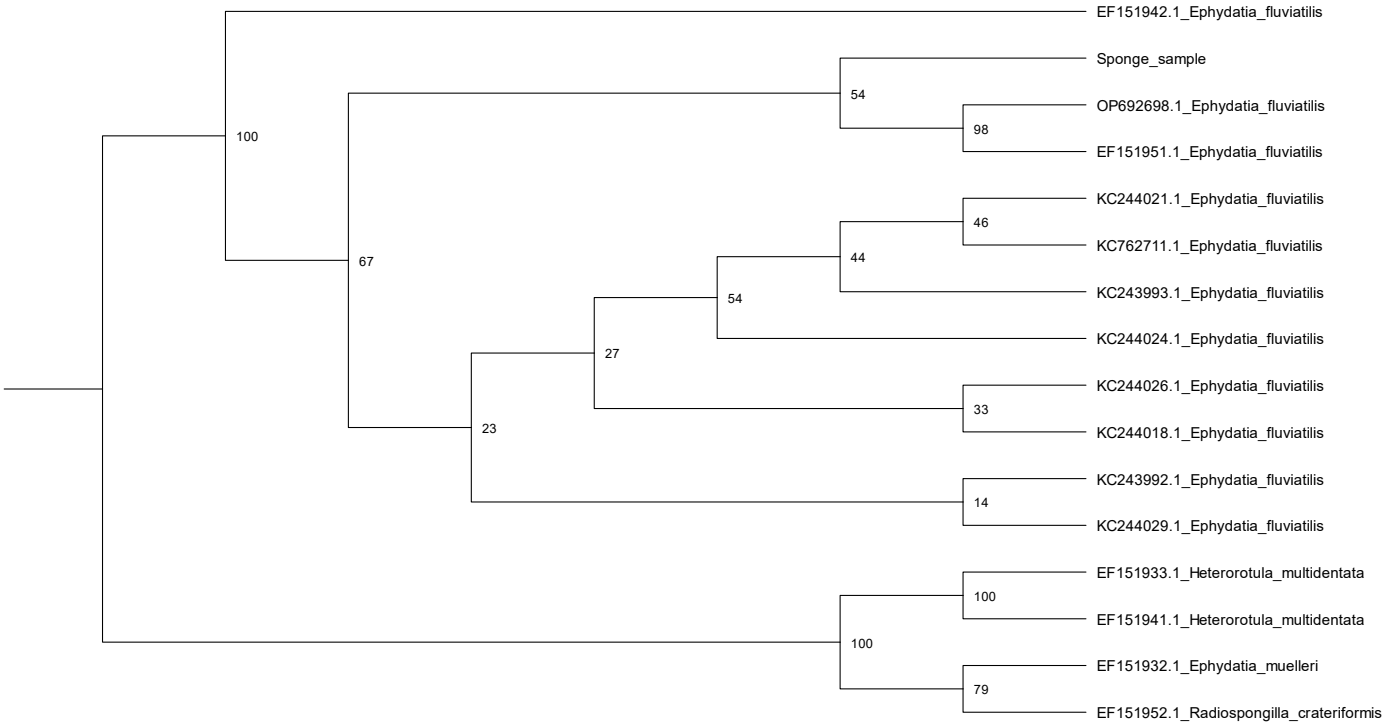

Fig. 2. Phylogenetic tree, support in UFBoot nodes (%)

Ultrafast Bootstrap Support (UFBoot) estimates the reliability of branches in a phylogenetic tree, where high values (95-100%) indicate reliable evolutionary relationships, such as those of EF151941.1 and EF151933.1_Heterorotula_multidentata.

Moderate support (70-95%) for EF151932.1_Ephydatia_muelleri and EF151952.1_Radiospongilla_crateriformis indicates possible uncertainties in the data. Of particular interest is the sponge under study, which forms a cluster with EF151942.1_Ephydatia_fluviatilis with 54% support, which although suggests its affiliation to this species, but requires additional confirmation due to its moderate statistical reliability.

Low values (<70%) as in KC243993.1 and KC244029.1_Ephydatia_fluviatilis indicate the weak reliability of such branches, which may change with the acquisition of new data.

Conclusions

Molecular genetic analysis of the ITS region of ribosomal DNA using Sanger sequencing followed by identification through the NCBI database (BLAST) identified the studied freshwater sponges from Lake Sevan as the species *Ephydatia fluviatilis*.

The applied methods demonstrated high efficiency: PCR amplification with optimized primers successfully isolated the target ITS fragment (~900 bp), which was confirmed by electrophoresis in agarose gel.

It is worth noting that the moderate support of the node (54%) in the phylogenetic tree indicates the need for further research to confirm the species assignment of the specimen. This may include the analysis of additional genetic markers or morphological studies that will help clarify the taxonomic status of the sponge under study.

Also, the presence of *E. fluviatilis* in the ecosystem of the high-mountain Lake Sevan (Armenia) was molecularly confirmed, and the effectiveness of the ITS marker for the taxonomy of freshwater sponges was proven. This is especially relevant for a group of organisms with pronounced morphological plasticity, where traditional identification methods are often insufficient.

References

1) Alberts Bruce. Essential cell biology / Bruce. Alberts, Garland Science, 2010. 845 p.

2) Basic Local Alignment Search Tool // National Library of Medicine URL: https://blast.ncbi.nlm.nih.gov/Blast.cgi (дата обращения: 05.08.2025).

3) BioEdit 7.7 // Software Informer URL: https://bioedit.software.informer.com/ (дата обращения: 05.08.2025).

4) FigTree v1.4.5 pre-release // GitHub URL: https://github.com/rambaut/figtree/releases (дата обращения: 05.08.2025).

5) MAFFT version 7 // Research Institute of Microbial Diseases URL: https://mafft.cbrc.jp/alignment/server/index.html (дата обращения: 05.08.2025).

6) Manconi R., Pronzato R. Phylum Porifera / R. Manconi, R. Pronzato // Elsevier Inc., 2015. – С. 133–157.

7) Metagenomic analysis of viral communities in diseased Baikal sponge *Lubomirskia baikalensis* / T. V. Butina, Yu. S. Bukin, I. V. Khanaev [et al.] // Limnology and Freshwater Biology. – 2019. – № 1. – С. 155–162.

8) Ицкович, Валерия Борисовна. Молекулярная филогения и систематика пресноводных губок: автореферат дис. кандидата биологических наук: 03.00.15 / Ин-т биологии моря ДВО РАН. — Владивосток, 2005. — 22 с.

9) Метод Сэнгера // РУВИКИ URL: https://ru.ruwiki.ru/wiki/Метод_Сэнгера (дата обращения: 08.08.2025).

10) ПЦР в диагностике: принципы, применение и подготовка к исследованию // ЕВРОДОН URL: https://eurodon61.ru/about/articles/ptsr-v-diagnostike-printsipy-primenenie-i-podgotovka-k-issledovaniyu/ (дата обращения: 06.08.2025).

Acknowledgments

I would like to express my sincere gratitude to all those who contributed to the implementation of this work. First of all, I would like to thank my scientific supervisor, **Dr. Elena Kravchenko**, for her professional mentoring and expert advice at all stages of the work. Her critical comments and support were invaluable for the successful completion of the project.

Special thanks to **Kirill Antonovich Tarasov** for help in mastering molecular biology methods (PCR, electrophoresis, cloning), patient teaching of the intricacies of experiments and prompt resolution of technical issues.

I would like to sincerely thank **Alena Sergeevna Yakhnenko** for her key contribution to bioinformatics analysis: training in working with programs for processing sequences, constructing phylogenetic trees, and interpreting molecular results.

Special words of gratitude go to the **START** program administration and the **Joint Institute for Nuclear Research (JINR)** for the unique opportunity to do an internship at a world-class scientific center. The professional environment and access to modern equipment became the foundation for acquiring invaluable experience.