Limnology and Freshwater Biology 2024 (5): 1267-1275 DOI:10.31951/2658-3518-2024-A-5-1267

Short communication

The Complete Mitochondrial Genome of *Chlorella* sp. isolated from Lake Baikal

Panova Yu.A.[1,2], Chernogor L.I.[2], Belikov S.I.[2]*

[1] Siberian Federal University, Akademgorodok, 13a, Krasnoyarsk, 660036, Russia
[2] Limnological Institute, Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya Str., 3, Irkutsk, 664033, Russia

ABSTRACT. We report here the mitochondrial genome sequence of Lake Baikal green alga extracted from complete genome of *Chlorella* sp. strain BAC9706. Sequenced mitochondrial genome of strain BAC9706 has 90.8-kbp containing including 34 protein-coding genes. Phylogenetic analysis using the mitochondrial genomes of algae of Chlorellales order indicated that the strain BAC9706 belongs *C. vulgaris* clade and is Baikal ecotype of *C. vulgaris*.

Keywords: *Chlorella vulgaris*, mitochondrial genome, Lake Baikal

For citation: Panova Yu.A., Chernogor L.I., Belikov S.I. The Complete Mitochondrial Genome of *Chlorella* sp. isolated from Lake Baikal // Limnology and Freshwater Biology. 2024. - № 5. - P. 1267-1275. DOI: 10.31951/2658-3518-2024-A-5-1267

1. Introduction

Baikal is a unique ancient oligotrophic freshwater lake - the deepest, oldest and largest by volume in the world. The lake has an unusual climatic environment and amazing of endemic flora and fauna diversity. Autotrophic picoplankton plays a significant role in the production of primary organic matter. In summer, picoplankton accounts for 60–90% of all primary production in Lake Baikal (Votintsev et al., 1975; Bondarenko and Guselnikova, 1989; Nagata et al., 1994). The greatest contribution to the phytoplankton biomass is made by diatoms and golden algae (Bondarenko et al., 2017), while green algae Chlorophyta are rare and represented by several species, with Chlorella-like algae noted only sporadically. Thus, authors reported the under-ice development of Chlorella sp. Baikal (Kozhova, 1959; 1987, Nagata et al., 1994). It was recently shown that Chlorella-like green algae dominated under the ice in 2020, and *C. vulgaris* (Beijerinck, 1890) was found at all the studied stations of the Irkutsk reservoir, indirectly adjacent to Lake Baikal. (Bashenkhaeva et al., 2020; Firsova et al., 2023).

The mass death of endemic sponges *Lubomirskia baikalensis* began in Baikal in 2011. (Belikov et al., 2019). Previously, we showed that significant changes occurred in the microbial communities of the sponge *L. baikalensis* and cell culture of the primmorph. A significant shift in microbial communities and mass death of the main symbiotic green algae *Choricystis* and their partial replacement with other algae in sponge communities were detected. (Belikov et al., 2019; Chernogor et al., 2020).

The draft genome sequence of *Chlorella* sp. strain BAC9706 (Petrushin et al., 2020) was prepared to expand our molecular biological knowledge of these microalgae and for compare them in the future with an endosymbiotic alga strain. In this study, we assembled the complete mitochondrial genome of a microalgal strain BAC9706 morphologically similar to Chlorella-like algae.

2. Materials and methods

2.1. DNA Sequencing and assembly

The genomic DNA was isolated as previously described (Petrushin et al., 2020) and sequenced using Illumina MiSeq platform. A draft assembly was built using SPAdes v. 3.15.4 (Bankevich et al., 2012). This draft assembly contained 5,837 contigs with an N50 value of 44,654 bp, and the largest contig was 317,606 bp long.

The obtained whole-genome assembly contigs were aligned on the reference genomes derived from NCBI using BLASTn to identify mitochondrial contig.

---

*Corresponding author.
E-mail address: sergeibelikov47@gmail.com (S.I. Belikov)

Received: September 30, 2024; ***Accepted:*** October 23, 2024;
Available online: October 31, 2024

2.2. Mitogenome annotation

For genome annotation we used GeSeq tool (Tillich et al., 2017) with default parameters for protein-coding genes, tRNAs, and rRNAs with NC_045362 of *Chlorella vulgaris* strain NJ-7 sequence as reference for mitochondrial genome. The protein-coding genes were verified manually through alignments of homologous mitogenomes from other *Chlorella vulgaris* strains using BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Proteins encoded in introns (orfs) were identified by translating the nucleotide sequence into protein sequence using the translate tool with standard genetic codes (https://web.expasy.org/translate/). Complete mitochondrial DNA sequence was uploaded to the GenBank database (accession number PQ043348). OGDRAW was used for genome maps visualization OGDRAW (Greiner et al., 2019), to prepare files for Genbank submission we used GB2sequin (https://chlorobox.mpimp-golm.mpg.de/GenBank2Sequin.html) (Lehwark and Greiner, 2019).

2.3. Phylogenetic Analysis

To determine the phylogenetic relationship of *Chlorella* genus a Maximum-likelihood tree was constructed using MEGA11 (Tamura et al., 2021). All currently available sequences of complete well-annotated mitochondrial genomes belonging to the *Chlorellales* order were selected from NCBI. From these, we extracted all protein-coding sequences, performed alignments, and concatenated them. To improve alignments, sequences from outlying taxa were removed. Final phylogenetic tree is based on mitochondrial protein sequences representing 31 concatenated protein-coding mitochondrial genes (3 of 34 annotated genes for intron-encoded orfs were excluded from the analysis).

3. Results and Discussion

3.1. Nucleotide Composition and Genome Structure

The complete mitochondrial genome of *Chlorella* sp. BAC9706 was identified as a circular double-stranded molecule with a length of 90770 bp and is located between related strains UTEX259 (98062 bp) and NJ-7 (87477 bp). The nucleotide composition of *Chlorella* sp. strain BAC9706 is 35.2 % A, 35.0 % T, 15.1 % G, and 14.6 % C, with a higher AT bias (70.2 %). The AT content of the strain BAC9706 mitogenome was slightly higher than that of the two related strains, equal to 70.0 % (Table 1).

The mitochondrial genome of strain BAC9706 contains 34 protein-coding genes (PCGs), 27 tRNA, and 3 rRNA genes. Among these, a total of 19 PCGs are encoded on the heavy (H) strand, while 15 genes are located on the light (L) strand (Fig. 1). Total length of protein-coding genes is 26969 bp, accounting for 29.7% of entire mitochondrial genome. The gene order within the mitochondrial genome is identical to that of two related strains.

The protein-coding genes include 13 for ribosomal proteins (rpl5-6, 16, rps2-4, 7, 10-14, 19), 9 for NAD(P)H-quinone oxidoreductases (nad1–7, 9 and nad4L), 5 for ATP synthases (atp1, atp4, atp6, atp8 and atp9), 3 for coxs (cox1–3), 1 for cob, and 3 orf for putative proteins. Transfer RNA genes for all 20 amino acids were identified, in which tRNA-Met, tRNA-Leu are triplicated, and tRNA-Ser, tRNA-Arg, and tRNA-Gly are duplicated.

All PCGs have a typical initiation codon ATG and stop with TAA, except for rps13, rps14, atp8, atp9, cox2, nad3 that uses a stop codon TAG. Among the protein-coding genes, nad5 is the longest gene with a length of 2019 bp, while the shortest is the atp9 gene with a length of 225 bp. The number of bases in the 13 PCGs follows the pattern A (35.2%) > T (35.00%) > G (15.1%) > C (14.6%), suggesting that AT is more preferred in PCGs.

3.2. Phylogenetic analysis

Phylogenetic analysis using the mitochondrial genomes of *Chlorellales* species indicated that strain BAC9706 is closely related to other *C. vulgaris* strains - UTEX259, NJ-7, ITBBA3-12, KNUA007 and they clustered in the *Chlorella* lineage with 100% bootstrap support (Fig. 2).

Table 1. Nucleotide distribution strains.

Nucleotide	BAC9706		UTEX259		NJ-7	
	Count, bp	%	Count, bp	%	Count, bp	%
Adenine (A)	31978	35.2	34356	35.0	30602	35.0
Cytosine (C)	13285	14.6	14462	14.7	12883	14.7
Guanine (G)	13728	15.1	14965	15.3	13348	15.3
Thymine (T)	31779	35.0	34279	35.0	30644	35.0
Purines (A+G)	45706	50.4	49321	50.3	43950	50.2
Pyrimidines (C+T)	45064	49.6	48741	49.7	43527	49.8
C + G	27013	29.8	29427	30.0	26231	30.0
A + T	63757	70.2	68635	70.0	61246	70.0
Total length	90770		98062		87477	

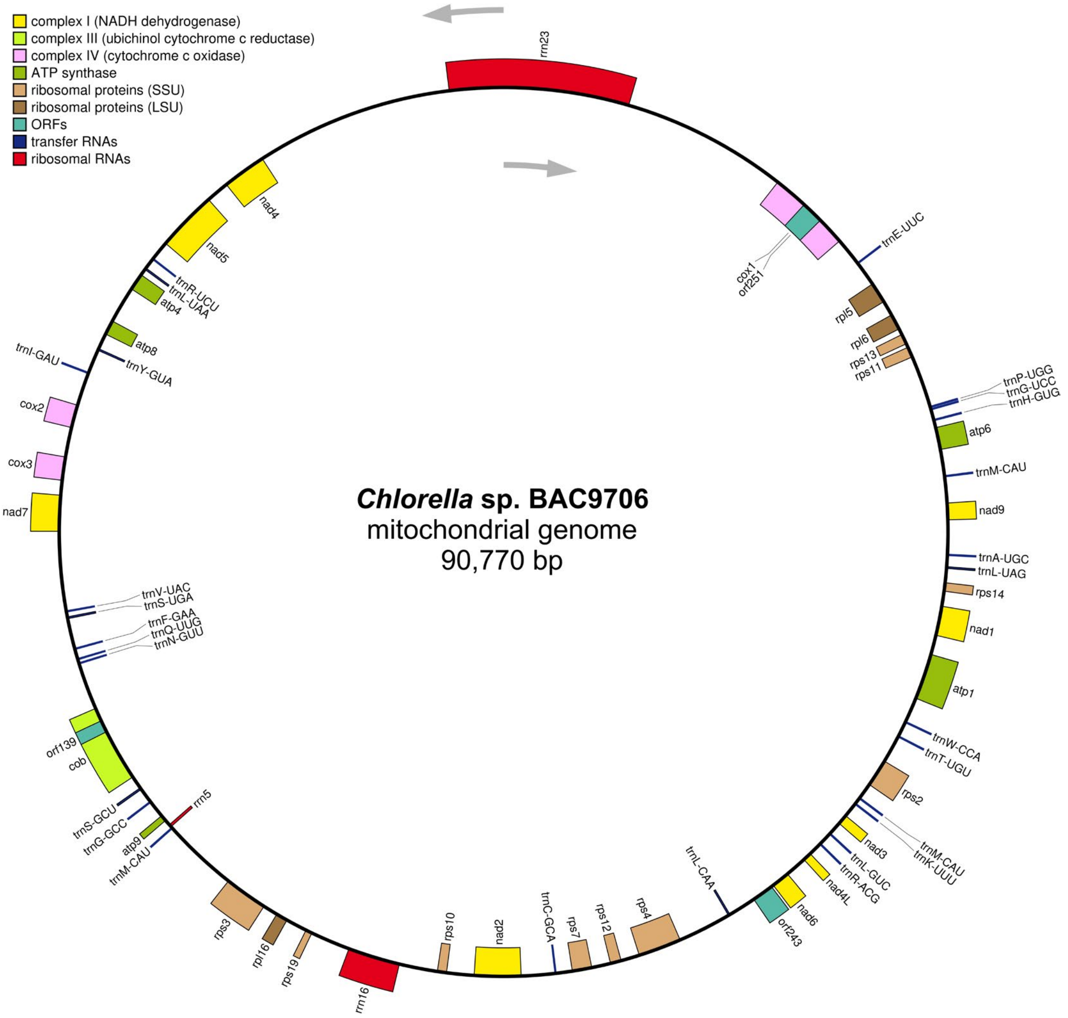

Fig.1. Mitochondrial genome map of *Chlorella* sp. strain BAC9706.

3.3. Whole mitogenome alignment

Whole genome alignment analysis for three strains *C. vulgaris* - UTEX259, NJ-7 and *Chlorella* sp. BAC9706 (Fig. 3) shows the following major differences in the mitochondrial genomes: in the UTEX259 strain, the intron of cox1 gene contains orf441, which is absent in the other two genomes; significant differences are in the gene encoding 23S rRNA and in the adjacent region, similarly, low conservation is observed next to the gene encoding 16S rRNA.

Thus, the main differences between the three closely related strains are related to the intergenic regions closed to rRNA genes.

4. Conclusions

In this study, we successfully assembled and analyzed the complete mitochondrial genome Chlorophyta green algae of the Lake Baikal. Next-generation whole genome sequencing of Chlorella-like green algae and bioinformatics analysis revealed a circular double-stranded molecule with a typical set of mitochondrial genes.

The constructed phylogenetic tree clearly indicates that the Chlorella-like strain BAC9706 is one of the representatives of the algae species *C. vulgaris*, Baikal ecotype of *C. vulgaris*.

Acknowledgements

We acknowledge D. Kuzmin, V. Sharov, K. Krutovsky and S. Feranchuk for their assistance in computational cluster administration and computation. The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation within the framework of State Assignment No. 0279-2021-0011.

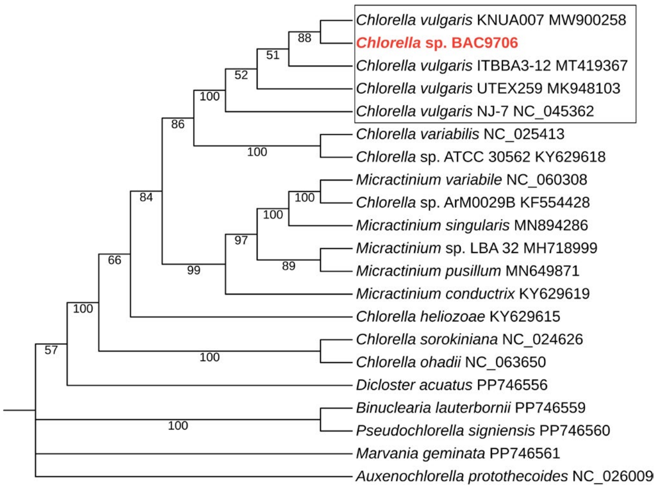

Fig.2. The Maximum Likelihood phylogenetic tree of *Chlorella* sp. strain BAC9706 based on mitochondrial concatenated protein-coding genes from Chlorellales taxa.

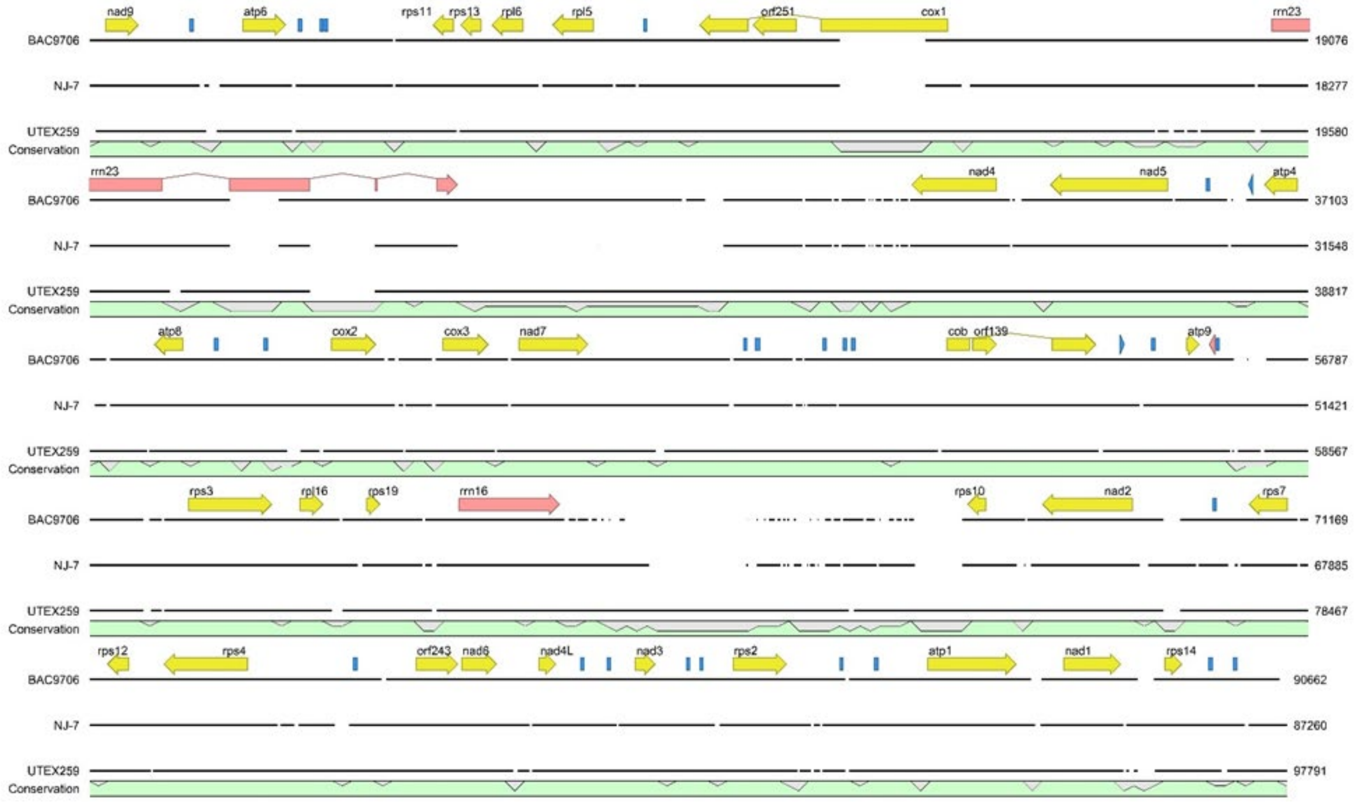

Fig.3. Alignment of whole genomes for three *Chlorella* strains - UTEX259 (MK948103), NJ7 (NC_045362) and *Chlorella* sp. BAC9706 (PQ043348). CDSs are shown in yellow, rRNAs in pink and tRNAs in blue. The light green line plot shows the conservation of the corresponding alignment region.

Conflict of interests

The authors hereby state that this research work and manuscript production complied with ethical standards, and none of the authors have any potential conflict of interests. We further declare that this research was not funded by any agency except RSF.

References

Bankevich A., Nurk S., Antipov D. et al. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455-477. DOI: 10.1089/cmb.2012.0021

Bashenkhaeva M.V., Galachyants Y.P., Khanaev I.V. et al. 2020. Comparative analysis of free-living and particle-associated bacterial communities of Lake Baikal during the ice-covered period. Journal of Great Lakes Research 46(3): 508-518. DOI: 10.1016/j.jglr.2020.03.015

Belikov S., Belkova N., Butina T. et al. 2019. Diversity and shifts of the bacterial community associated with Baikal sponge mass mortalities. PLoS ONE 14(3): e0213926. DOI: 10.1371/journal.pone.0213926

Bondarenko N., Sheveleva N., Rozhkova N. et al. 2017. Remote mountain lakes of Eastern Siberia: a pattern of ecologically pure non-industrialized water-bodies. Environmental Earth Sciences 76(10): 1-10. DOI: 10.1007/s12665-017-6708-4

Bondarenko N.A., Guselnikova N.E. 1989. Production of phytoplankton of Southern Baikal. Bulletin of the Siberian Branch of the USSR Academy of Sciences. Issue 1. p. 77-80.

Chernogor L., Klimenko E., Khanaev I. et al. 2020. Microbiome analysis of healthy and diseased sponges Lubomirskia baicalensis by using cell cultures of primmorphs. PeerJ 8(e9080): 1-24. DOI: 10.7717/peerj.9080

Firsova A., Galachyants Yu., Bessudova A. et al. 2023. Summer phytoplankton species composition and abundance in the southern basin of Lake Baikal and Irkutsk Reservoir. Limnology and Freshwater Biology 6: 204-228. DOI: 10.31951/2658-3518-2023-A-6-204

Greiner S., Lehwark P., Bock R. 2019. Organellar genome DRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research 47(W1): W59-W64. DOI: 10.1093/nar/gkz238

Kozhova O.M. 1959. About the subglacial “bloom” of phytoplankton in Lake Baikal. Bot. Zh. 44 (7). p. 1001.

Kozhova O.M. 1987. Phytoplankton of Lake Baikal: structural and functional characteristics. Arch. Hydrobiol. Beih. Ergebn. Limnol. 25. p. 19.

Lehwark P., Greiner S. 2019. GB2sequin - A file converter preparing custom GenBank files for database submission. Genomics 111(4): 759-761. DOI: 10.1016/j.ygeno.2018.05.003

Nagata T., Takai K., Kawanobe K. et al. 1994. Autotrophic picoplankton in southern Lake Baikal: abundance, growth and grazing mortality during summer. J. Plankton Research 16(8): 945-959.

Petrushin I.S., Belikov S.I., Belykh O.I. et al. 2020. Draft genome sequence of the green microalga Chlorella sp. strain BAC9706, isolated from Lake Baikal, Russia. Microbiology Resource Announcements 9(43). DOI: 10.1128/mra.00966-20

Tamura K., Stecher G., Kumar S. 2021. MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution 38(7): 3022-3027. DOI: 10.1093/molbev/msab120

Tillich M., Lehwark P., Pellizzer T. et al. 2017. GeSeq - versatile and accurate annotation of organelle genomes. Nucleic Acids Research 45(W1): W6-W11. DOI: 10.1093/nar/gkx391

Votintsev K.K., Meshcheryakova A.I., Popovskaya G.I. 1975. Organic matter cycle in Lake Baikal. USSR Academy of Sciences. SO Limnol.in-t. Novosibirsk: Nauka. 189 p.