

The complete chloroplast genome of *Chlorella vulgaris* strain BAC9706 isolated from Lake Baikal

Panova Yu.A.^{1,2}, Chernogor L.I.¹, Belikov S.I.^{1*}

¹Limnological Institute, Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya Str., 3, Irkutsk, 664033, Russia

²Siberian Federal University, Svobodny prosp., 79, Krasnoyarsk, 660041, Russia

ABSTRACT. We report here the complete chloroplast genome sequence of the strain BAC9706 of green microalga *Chlorella vulgaris* isolated from Lake Baikal. The chloroplast genome is a circular DNA molecule of 161,721 bp and contains 78 protein-coding genes, 33 tRNA genes, and 3 rRNA genes. Comparative analysis revealed a high level of conservation of gene content and gene order among other *C. vulgaris* strains with minor differences mainly associated with introns and non-coding regions. Phylogenetic analysis based on concatenated chloroplast protein-coding genes showed that strain BAC9706 clusters within the *C. vulgaris* clade with high bootstrap support. These results indicate that strain BAC9706 represents a Baikal ecotype of *C. vulgaris* rather than a new species.

Keywords: *Chlorella vulgaris*, chloroplast genome, Lake Baikal, Chlorellales

For citation: Panova Yu.A., Chernogor L.I., Belikov S.I. The complete chloroplast genome of *Chlorella vulgaris* strain BAC9706 isolated from Lake Baikal // Limnology and Freshwater Biology. 2025. - № 6. - P. 1380-1389. DOI: 10.31951/2658-3518-2025-A-6-1380

1. Introduction

Lake Baikal is a unique ancient oligotrophic freshwater lake characterized by exceptional biodiversity and a high level of endemism. It is the deepest and one of the oldest freshwater lakes in the world, containing about 20% of the global freshwater reserves (Gao et al., 1994; Martens, 1997; Timoshkin, 2011). The long geological history of Lake Baikal, combined with its specific climatic and hydrological conditions, has resulted in the formation of highly specialized aquatic communities.

Autotrophic picoplankton plays a significant role in primary production in Lake Baikal ecosystems. During the summer period, picoplankton contributes up to 60–90% of total primary production in the lake (Votintsev et al., 1975; Bondarenko and Guselnikova, 1989; Nagata et al., 1994). The main contribution to phytoplankton biomass is typically made by diatoms and chrysophytes (Bondarenko et al., 2020), whereas green algae of the phylum Chlorophyta are generally considered rare and are represented by a limited number of taxa.

Chlorella-like green algae in Lake Baikal were first described based on microscopic observations and were reported sporadically, including under-ice phytoplankton development (Kozhova, 1959; Kozhova,

1987; Nagata et al., 1994). More recent studies have demonstrated notable changes in phytoplankton composition in Baikal and adjacent water bodies. In particular, *Chlorella vulgaris* was detected at all investigated stations of the Irkutsk Reservoir, which is indirectly connected to Lake Baikal, and was shown to dominate in under-ice and near-ice phytoplankton communities during certain periods (Bashenkhaeva et al., 2020; Firsova et al., 2023).

These observations raise questions about the origin, persistence, and ecological characteristics of *Chlorella* strains occurring in cold freshwater environments of the Baikal region. It remains unclear whether these strains represent transient populations introduced from surrounding water bodies or locally adapted ecotypes capable of surviving and reproducing under low-temperature conditions.

Previously, a draft nuclear genome sequence of *Chlorella* sp. strain BAC9706 isolated from Lake Baikal was published (Petrushin et al., 2020), followed by a detailed analysis of its mitochondrial genome (Panova et al., 2024). These studies provided important insights into the genomic features and phylogenetic position of this strain, indicating its affiliation with the *C. vulgaris* lineage and suggesting that strain BAC9706 represents a Baikal ecotype of this species.

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

Received: November 12, 2025;

Accepted after revised: December 20, 2025;

Available online: December 25, 2025

© Author(s) 2025. This work is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.



To further expand our understanding of the genomic organization of this Baikal strain and to enable comparative analyses with other representatives of the genus *Chlorella*, we assembled and analyzed the complete chloroplast genome of *C. vulgaris* strain BAC9706 isolated from Lake Baikal.

2. Materials and Methods

2.1. DNA sequencing and genome assembly

The genomic DNA was isolated as previously described (Petrushin et al., 2020), and sequenced using the 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 against reference chloroplast genomes of *C. vulgaris* available in NCBI using BLASTn to identify chloroplast contigs.

2.2. Genome annotation and visualization

For genome annotation we used the GeSeq tool (Tillich et al., 2017) with default parameters for protein-coding genes, tRNAs, and rRNAs, using the *C. vulgaris* strain NJ-7 (GenBank accession MK948100) chloroplast genome as reference. OGDRAW was used for chloroplast genome map visualization (Greiner et al., 2019), to prepare files for GB2sequin (Lehwark and Greiner, 2019).

2.3. Phylogenetic analysis

The evolutionary analysis was conducted using the Maximum Likelihood method implemented in MEGA11 (Tamura et al., 2021). The best-fit substitution model was selected automatically and branch support was assessed using 1000 bootstrap replicates. All currently available sequences of complete well-annotated chloroplast 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 alignment quality, poorly aligned regions were removed.

The final phylogenetic tree is based on 73 concatenated protein-coding chloroplast genes from 16 taxa (5 of the annotated *ycf* genes were excluded from this analysis).

3. Results

3.1. Chloroplast genome features

The complete chloroplast genome of *C. vulgaris* strain BAC9706 is a circular double-stranded DNA molecule with a length of 161,721 bp. The genome contains a total of 114 genes, including 78 protein-coding genes, 33 tRNA genes, and 3 rRNA genes (Fig. 1).

The protein-coding genes include 1 *accD* gene, 6 genes for ATP synthases (*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *atpI*), 1 for cell envelope (*cemA*), 3 protochlorophyllide reductase genes (*chlB*, *chlL*, *chlN*) and 1 Mg-protoporphyrin IX chelatase (*chlI*), 2 genes related to sulfate transport (*cysA* and *cysT*), 2 genes related to cell division and protein quality control (*ftsH*, *minD*), 2 hydrolases (*I-CvuI* and *clpP*), 1 translation initiation factor (*infA*), 6 cytochrome-related genes (*ccsA*, *petA*, *petB*, *petD*, *petG*, *petL*), 22 involved in photosynthesis (*rbcl*, *psaA*, *psaB*, *psaC*, *psaI*, *psaJ*, *psaM*, *psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*), 21 for ribosomal proteins (*rpl12*, *rpl14*, *rpl16*, *rpl19*, *rpl2*, *rpl20*, *rpl23*, *rpl32*, *rpl36*, *rpl5*, *rps11*, *rps12*, *rps14*, *rps18*, *rps19*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps9*), 4 for RNA synthesis (*rpoA*, *rpoB*, *rpoC1*, and *rpoC2*), 1 for elongation factor *Tu* (*tufA*), 5 *ycf* genes (*ycf3*, *ycf4*, *ycf62*, *ycf12*, *ycf78*). The 33 tRNA genes cover the transfer of all 20 amino acids, in which 3 are tRNA-Gly, 2 tRNA-Ile, 4 tRNA-Leu, 3 tRNA-Arg, 3 tRNA-Ser, 2 tRNA-Thr, 2 tRNA-Val genes.

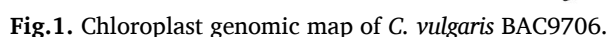
The complete chloroplast genome sequence of *C. vulgaris* BAC9706 has been deposited in the GenBank database under the accession number PQ037191.1.

3.2. Comparative analysis with other *C. vulgaris* strains

Currently published complete chloroplast genomes of *C. vulgaris* vary in genome size and gene content, reflecting differences in annotation approaches and assembly completeness (Table 1). Differences in the reported numbers of protein-coding genes among

Table 1. General features of currently published chloroplast genomes of *C. vulgaris*

Accession	Strain	Genome size, bp	Protein- coding genes	tRNA/rRNA genes	Reference
PQ037191.1	<i>C. vulgaris</i> BAC9706	161721	78	33/3	Present study
MW900257.1	<i>C. vulgaris</i>	156202	73	32/3	Unpublished
NC_001865.1	<i>C. vulgaris</i> C-27	150613	174	33/3	Wakasugi et al., 1997
CM041649.1	<i>C. vulgaris</i> 211/11P	165504	70	NA	Cecchin et al., 2019
MT577052.1	<i>C. vulgaris</i>	165412	78	33/3	Wen and Wan, 2020
MT920676.1	<i>C. vulgaris</i> ITBBA3-12	168369	111	29/3	Han et al., 2021
MK948102.1	<i>C. vulgaris</i> UTEX259	176851	79	33/3	Unpublished
MK948100.1	<i>C. vulgaris</i> NJ-7	154201	79	33/3	Unpublished



Comparative analysis revealed a high degree of sequence similarity between the chloroplast genomes. Gene content and gene order were highly conserved among the analyzed genomes, and no large-scale rearrangements affecting genome organization were detected.

Overall, most nucleotide substitutions between BAC9706 and other *C. vulgaris* strains were located in non-coding intergenic regions, while protein-coding genes were highly conserved. In particular, no amino acid substitutions were detected in the *rbcL* gene, indi-

3.3. Phylogenetic analysis

4. Discussion

The complete chloroplast genome of *C. vulgaris* strain BAC9706 isolated from Lake Baikal complements previously published mitochondrial genome data for this strain. Comparative analysis revealed a high level of conservation in gene content, gene order, and overall genome organization relative to other *C. vulgaris* strains, with no evidence of large-scale rearrangements.

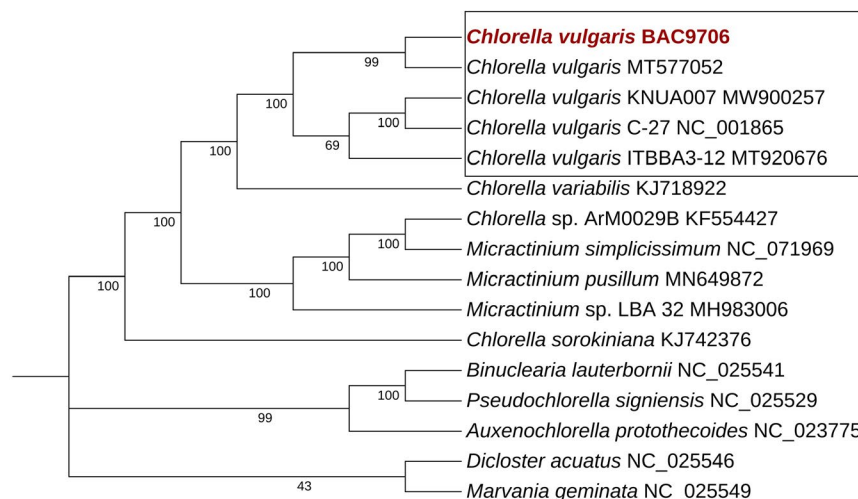


Fig.2. Phylogenetic tree of *C. vulgaris* BAC9706 based on chloroplast concatenated protein-coding genes from Chlorellales taxa.

Most sequence variation was confined to intronic and intergenic regions, whereas protein-coding genes remained highly conserved. In particular, differences observed in the *psbA* intron, including variation in the length of the intron-encoded endonuclease, likely reflect intraspecific variability of mobile introns rather than functional divergence of photosynthetic genes.

Phylogenetic analysis based on complete chloroplast genome sequences placed strain BAC9706 firmly within the *C. vulgaris* clade, consistent with previous mitochondrial genome-based results (Panova et al., 2024). Together, these data indicate that BAC9706 represents a Baikal ecotype of *C. vulgaris* rather than a distinct species and provide an additional organellar reference for future comparative studies.

5. Conclusions

The complete chloroplast genome of *C. vulgaris* strain BAC9706 isolated from Lake Baikal was assembled and analyzed. The genome exhibits a high level of conservation in gene content and organization compared to other *C. vulgaris* strains, with differences mainly confined to intronic and intergenic regions. Phylogenetic analysis confirms that BAC9706 represents a Baikal ecotype of *C. vulgaris* rather than a distinct species. These data provide a genomic basis for future studies of *Chlorella* strains in cold freshwater environments.

Acknowledgements

The State Assignments of Limnological Institute Siberian Branch of Russian Academy of Sciences funded this research No. 0279-2021-0011. This material is based upon work supported by the Russian Science Foundation Project No. 19-14-00088 and the Russian Foundation for Basic Research No. 16-04-00065; 16-54-150007; 18-04-00224.

Conflict of interest

The authors declare no conflicts of interest.

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):455-77. DOI: [10.1089/cmb.2012.0021](https://doi.org/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](https://doi.org/10.1016/j.jglr.2020.03.015)
- Bondarenko N.A., Vorobyova S.S., Zhuchenko N.A. et al. 2020. Current state of phytoplankton in the littoral area of lake Baikal, spring 2017. *Journal of Great Lakes Research* 46 (1): 17-28. DOI: [10.1016/j.jglr.2019.10.001](https://doi.org/10.1016/j.jglr.2019.10.001)
- Bondarenko N.A., Guselnikova N.E. 1989. Production of phytoplankton of Southern Baikal. *Bulletin of the Siberian Branch of the USSR Academy of Sciences* 1: 77-80.
- Cecchin M., Marcolungo L., Rossato M. et al. 2019. *Chlorella vulgaris* genome assembly and annotation reveal the molecular basis for metabolic acclimation to high light conditions. *Plant Journal* 100(6): 1289-1305. DOI: [10.1111/tpl.14508](https://doi.org/10.1111/tpl.14508)
- 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](https://doi.org/10.31951/2658-3518-2023-A-6-204)
- Gao S., Davis P., Liu K. et al. 1994. Asymmetric upwarp of the asthenosphere beneath the Baikal rift zone, Siberia. *Journal of Geophysical Research* 99: 15319-15330. DOI: [10.1029/94JB00808](https://doi.org/10.1029/94JB00808)
- Greiner S., Lehwark P., Bock R. 2019. Organellar-GenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* 47(W1). DOI: [10.1093/nar/gkz238](https://doi.org/10.1093/nar/gkz238)
- Han B., Mu Y., Tan D. 2021. The chloroplast genome of a unicellular green alga strain isolated from the rubber processing wastewater. *Mitochondrial DNA Part B: Resources* 6(1): 15-16. DOI: [10.1080/23802359.2020.1844090](https://doi.org/10.1080/23802359.2020.1844090)
- Kozhova O.M. 1959. About the subglacial "bloom" of phytoplankton in Lake Baikal. *Botanical Journal* 44(7): 1001.
- Kozhova O.M. 1987. Phytoplankton of Lake Baikal: structural and functional characteristics. *Arch. Hydrobiol. Beih. Ergebn. Limnol* 25: 19-37.
- 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](https://doi.org/10.1016/j.ygeno.2018.05.003)

Martens K. 1997. Speciation in ancient lakes. Trends In Ecology & Evolution 12(5): 177-82. DOI: [10.1016/S0169-5347\(97\)01039-2](https://doi.org/10.1016/S0169-5347(97)01039-2)

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

Panova Y.A., Chernogor L.I., Belikov S.I. 2024. The complete mitochondrial genome of *Chlorella* sp. isolated from Lake Baikal. Limnology and Freshwater Biology 2024 (5): 1267-1275. DOI: [10.31951/2658-3518-2024-A-5-1267](https://doi.org/10.31951/2658-3518-2024-A-5-1267)

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): e00966-20. DOI: [10.1128/MRA.00966-20](https://doi.org/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](https://doi.org/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](https://doi.org/10.1093/nar/gkx391)

Timoshkin O.A. 2011. Main tendencies in research of ancient Lake biodiversity; most interesting recent discoveries in biodiversity of Lake Baikal. In: Timoshkin O.A.(Ed.) Index of animal species inhabiting Lake Baikal and its catchment area. Novosibirsk: Nauka. pp. 1423-1428.

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.

Wakasugi T., Nagai T., Kapoor M. et al. 1997. Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. Proceedings of the National Academy of Sciences 94(11): 5967-5972. DOI: [10.1073/pnas.94.11.5967](https://doi.org/10.1073/pnas.94.11.5967)

Wen Y., Wan D. 2020. The complete chloroplast genome sequence of *Chlorella vulgaris* and phylogenetic analysis. Mitochondrial DNA Part B: Resources 5(3):2713-2715. DOI: [10.1080/23802359.2020.1787896](https://doi.org/10.1080/23802359.2020.1787896)

Хлоропластный геном штамма ВАС9706 *Chlorella vulgaris*, выделенного из озера Байкал

Краткое сообщение

LIMNOLOGY
FRESHWATER
BIOLOGY

Панова Ю.А.^{1,2}, Черногор Л.И.¹, Беликов С.И.^{1*}

¹ Лимнологический институт СО РАН, ул. Улан-Баторская, 3, Иркутск, 664033, Россия

² Сибирский федеральный университет, Академгородок, 13а, Красноярск, 660036, Россия

АННОТАЦИЯ. В настоящей работе представлена полная последовательность хлоропластного генома зелёной микроводоросли *Chlorella vulgaris* штамма ВАС9706, выделенного из озера Байкал. Хлоропластный геном представляет собой кольцевую молекулу ДНК длиной 161 721 п.н. и содержит 78 белок-кодирующих генов, 33 гена тРНК и 3 гена рРНК. Сравнительный анализ выявил высокий уровень консервативности генного состава и порядка генов по сравнению с другими штаммами *C. vulgaris*, при этом незначительные различия в основном связаны с интронами и некодирующими областями. Филогенетический анализ, основанный на конкатенированных белок-кодирующих генах хлоропластного генома, показал, что штамм ВАС9706 группируется внутри клады *C. vulgaris* с высокой бутстреп-поддержкой. Полученные результаты указывают на то, что ВАС9706 представляет собой байкальский экотип *C. vulgaris*, а не новый вид.

Ключевые слова: *Chlorella vulgaris*, хлоропластный геном, озеро Байкал, Chlorellales

Для цитирования: Панова Ю.А., Черногор Л.И., Беликов С.И. Хлоропластный геном штамма ВАС9706 *Chlorella vulgaris*, выделенного из озера Байкал // Limnology and Freshwater Biology. 2025. - № 6. - С. 1380-1389. DOI: 10.31951/2658-3518-2025-A-6-1380

1. Введение

Озеро Байкал — уникальное древнее олиготрофное пресноводное озеро, характеризующееся исключительным биоразнообразием и высоким уровнем эндемизма. Это самое глубокое и одно из древнейших пресноводных озёр в мире, содержащее около 20 % мировых запасов пресной воды (Gao et al., 1994; Martens, 1997; Timoshkin, 2011). Длительная геологическая история озера Байкал в сочетании со специфическими климатическими и гидрологическими условиями привела к формированию высокоспециализированных водных сообществ.

Автотрофный пикопланктон играет значимую роль в первичной продукции экосистем озера Байкал. В летний период вклад пикопланктона может достигать 60–90 % от общей первичной продукции озера (Votintsev et al., 1975; Bondarenko and Guselnikova, 1989; Nagata et al., 1994). Основной вклад в биомассу фитопланктона, как правило, вносят диатомовые и золотистые водоросли (Bondarenko et al., 2020), тогда как зелёные водоросли отдела

Chlorophyta обычно считаются редкими и представлены ограниченным числом таксонов.

Хлореллоподобные зелёные водоросли в озере Байкал были впервые описаны на основе микроскопических наблюдений и отмечались эпизодически, в том числе в составе подлёдного фитопланктона (Kozhova, 1959; Kozhova, 1987; Nagata et al., 1994). Более поздние исследования показали заметные изменения в составе фитопланктона Байкала и прилегающих водоёмов. В частности, *Chlorella vulgaris* была обнаружена на всех исследованных станциях Иркутского водохранилища, которое опосредованно связано с озером Байкал, и в отдельные периоды доминировала в подлёдных и припайных фитопланктонных сообществах (Bashenkhaeva et al., 2020; Firsova et al., 2023).

Эти наблюдения поднимают вопросы о происхождении, устойчивости и экологических характеристиках штаммов *Chlorella*, обитающих в холодных пресноводных экосистемах Байкальского региона. Остаётся неясным, представляют ли такие штаммы транзиторные популяции, занесённые из окружающих водоёмов, или же локально адаптированные

*Автор для переписки. Адрес e-mail: sergeibelikov47@gmail.com (С.И. Беликов)

Поступила: 12 ноября 2025;

Принята после доработки: 20 декабря 2025;

Опубликована online: 25 декабря 2025

© Автор(ы) 2025. Эта работа распространяется под международной лицензией Creative Commons Attribution-NonCommercial 4.0.



экотипы, способные к выживанию и размножению в условиях низких температур.

Ранее была опубликована черновая последовательность ядерного генома *Chlorella* sp. штамма ВАС9706, выделенного из озера Байкал (Petrushin et al., 2020), а затем выполнен детальный анализ его митохондриального генома (Panova et al., 2024). Эти исследования предоставили важную информацию о геномных особенностях и филогенетическом положении данного штамма, указав на его принадлежность к линии *C. vulgaris* и позволив предположить, что ВАС9706 представляет собой байкальский экотип этого вида.

Для дальнейшего расширения представлений о геномной организации данного байкальского штамма и проведения сравнительного анализа с другими представителями рода *Chlorella* в настоящей работе была выполнена сборка и анализ полного хлоропластного генома *C. vulgaris* штамма ВАС9706, выделенного из озера Байкал.

2. Материалы и методы

2.1. Секвенирование ДНК и сборка генома

Геномная ДНК была выделена ранее описанным методом (Petrushin et al., 2020) и секвенирована с использованием платформы Illumina MiSeq. Черновая сборка была выполнена с помощью программы SPAdes v. 3.15.4 (Bankevich et al., 2012). Полученная сборка включала 5837 контигов со значением N50 44654 п.н., при этом длина наибольшего контига составляла 317606 п.н.

Полученные контиги полногеномной сборки были выровнены по референсным хлоропластным геномам *C. vulgaris*, доступным в базе данных NCBI, с использованием BLASTn для идентификации контигов хлоропластного происхождения.

2.2. Аннотация генома и визуализация

Для аннотации генома использовали инструмент GeSeq (Tillich et al., 2017) с параметрами по умолчанию для белок-кодирующих генов, тРНК и рРНК, используя хлоропластный геном *C. vulgaris* штамма NJ-7 (номер доступа GenBank MK948100) в качестве референса. Визуализацию геномных карт выполняли с помощью OGDRAW (Greiner et al., 2019), а также подготавливали файлы для GB2sequin (Lehwark and Greiner, 2019).

2.3. Филогенетический анализ

Эволюционный анализ был проведён методом максимального правдоподобия с использованием программы MEGA11 (Tamura et al., 2021). Оптимальная модель нуклеотидных замен подбиралась автоматически, надёжность ветвей оценивали на основе 1000 бутстреп-реплик. Из базы данных NCBI были отобраны все доступные на момент исследования полностью аннотированные хлоропластные геномы представителей порядка Chlorellales. Из них были извлечены все белок-ко-

дирующие последовательности, выполнены выравнивания и проведена их конкатенация. Для повышения качества выравнивания плохо выровненные участки были удалены.

Итоговое филогенетическое дерево построено на основе 73 конкатенированных белок-кодирующих генов хлоропластного генома, полученных для 16 таксонов (пять аннотированных генов *ycf* были исключены из анализа).

3. Результаты и обсуждения

3.1. Характеристика хлоропластного генома

Полный хлоропластный геном *C. vulgaris* штамма ВАС9706 представляет собой кольцевую двуцепочечную молекулу ДНК длиной 161721 п.н. Геном содержит в общей сложности 114 генов, включая 78 белок-кодирующих генов, 33 гена тРНК и 3 гена рРНК (Рис. 1).

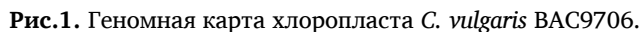
К числу белок-кодирующих генов относятся: один ген *accD*; шесть генов субъединиц АТФ-синтазы (*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *atpI*); один ген, связанный с формированием клеточной оболочки (*cemA*); три гена протохлорофиллидредуктазы (*chlB*, *chlL*, *chlN*) и один ген Mg-протопорфирин IX-хелатазы (*chlI*); два гена, связанные с транспортом сульфатов (*cysA* и *cysT*); два гена, участвующие в клеточном делении и контроле качества белков (*ftsH*, *minD*); две гидролазы (*I-CvuI* и *clpP*); один фактор инициации трансляции (*infA*); шесть генов, связанных с цитохромами (*ccsA*, *petA*, *petB*, *petD*, *petG*, *petL*); 22 гена, участвующие в фотосинтезе (*rbcl*, *psaA*, *psaB*, *psaC*, *psaI*, *psaJ*, *psaM*, *psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*); 21 ген рибосомных белков (*rpl12*, *rpl14*, *rpl16*, *rpl19*, *rpl2*, *rpl20*, *rpl23*, *rpl32*, *rpl36*, *rpl5*, *rps11*, *rps12*, *rps14*, *rps18*, *rps19*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps9*); четыре гена, участвующие в синтезе РНК (*rpoA*, *rpoB*, *rpoC1*, *rpoC2*); один ген фактора элонгации Tu (*tufA*); а также пять генов семейства *ycf* (*ycf3*, *ycf4*, *ycf62*, *ycf12*, *ycf78*).

Тридцать три гена тРНК обеспечивают перенос всех 20 аминокислот; при этом представлены три гена тРНК-Gly, два тРНК-Ile, четыре тРНК-Leu, три тРНК-Arg, три тРНК-Ser, два тРНК-Thr и два тРНК-Val.

Полная последовательность хлоропластного генома *C. vulgaris* ВАС9706 депонирована в базе данных GenBank под номером доступа PQ037191.1.

3.2. Сравнительный анализ с другими штаммами *C. vulgaris*

Опубликованные на настоящий момент полные хлоропластные геномы *C. vulgaris* различаются по размеру генома и генной насыщенности, что отражает различия в подходах к аннотации и полноте сборок (Таблица 1). Различия в числе аннотированных белок-кодирующих генов в значительной степени обусловлены используемыми стратегиями аннотации и критериями включения генов.



Заметное различие было обнаружено в интроне гена *psbA*. У штамма ВАС9706 данный интрон кодирует предполагаемую сайт-специфичную ДНК-эндонуклеазу длиной 234 аминокислоты, тогда как у штамма NJ-7 присутствует укороченный вариант. Это различие обусловлено несколькими нуклеотидными заменами и одной однонуклеотид-

Номер доступа	Штамм	Размер генома, п.н.	Белок-кодирующие гены	Гены тРНК/ рРНК	Источник
PQ037191.1	<i>C. vulgaris</i> BAC9706	161721	78	33/3	Данное исследование
MW900257.1	<i>C. vulgaris</i>	156202	73	32/3	Не опубликовано
NC_001865.1	<i>C. vulgaris</i> C-27	150613	174	33/3	(Wakasugi et al., 1997)
CM041649.1	<i>C. vulgaris</i> 211/11P	165504	70	NA	(Cecchin et al., 2019)
MT577052.1	<i>C. vulgaris</i>	165412	78	33/3	(Wen and Wan, 2020)
MT920676.1	<i>C. vulgaris</i> ITBBA3-12	168369	111	29/3	(Han et al., 2021)
MK948102.1	<i>C. vulgaris</i> UTEX259	176851	79	33/3	Не опубликовано
MK948100.1	<i>C. vulgaris</i> NJ-7	154201	79	33/3	Не опубликовано

ной вставкой, что приводит к формированию более длинной открытой рамки считывания у BAC9706.

В целом большинство нуклеотидных замен между BAC9706 и другими штаммами *C. vulgaris* локализованы в некодирующих межгенных областях, тогда как белок-кодирующие гены отличаются высокой степенью консервативности. В частности, в гене *rbcL* не было выявлено ни одной аминокислотной замены, что указывает на жёсткие функциональные ограничения, характерные для ключевых белков фотосинтетического аппарата.

3.3. Филогенетический анализ

Филогенетический анализ показал, что штамм BAC9706 группируется с другими штаммами *C. vulgaris*, образуя хорошо поддерживаемую кладу со 100% бутстреп-поддержкой (Рис. 2). Полученный результат согласуется с ранее выполненными анализами на основе митохондриального генома и подтверждает таксономическую принадлежность BAC9706 к виду *C. vulgaris*.

4. Обсуждение

Полный хлоропластный геном штамма BAC9706, выделенного из озера Байкал, дополняет ранее опубликованные данные о митохондриальном геноме данного штамма. Сравнительный анализ выявил у *C. vulgaris* высокий уровень консервативности генового состава, порядка генов и общей организации генома по сравнению с другими штаммами, при этом признаков крупномасштабных перестроек обнаружено не было.

Большинство различий в последовательностях было ограничено интронными и межгенными областями, тогда как белок-кодирующие гены оставались высоко консервативными. В частности, различия, выявленные в интроне гена *psbA*, включая вариации длины интрон-кодируемой эндонуклеазы, вероятно, отражают внутривидовую изменчивость мобильных интронов, а не функциональную дивергенцию фотосинтетических генов.

Филогенетический анализ, основанный на полных последовательностях хлоропластных геномов, уверенно помещает штамм BAC9706 в кладу *C. vulgaris*, что согласуется с ранее полученными результатами на основе митохондриального генома (Panova et al., 2024). В совокупности эти данные указывают на то, что BAC9706 представляет собой байкальский экотип *C. vulgaris*, а не новый вид, и обеспечивают дополнительный органеллярный референс для будущих сравнительных исследований.

5. Заключение

В настоящей работе был собран и проанализирован полный хлоропластный геном *C. vulgaris* штамма BAC9706, выделенного из озера Байкал. Геном характеризуется высоким уровнем консервативности генового состава и организации по сравнению с другими штаммами *C. vulgaris*, при этом выявленные различия в основном ограничены интронными и межгенными областями. Филогенетический анализ подтверждает, что BAC9706 представляет собой байкальский экотип *C. vulgaris*, а не отдельный вид. Полученные данные формируют геномную основу для дальнейших исследований штаммов *Chlorella*, обитающих в холодных пресноводных экосистемах.

Благодарности

Работа выполнена в рамках государственного задания Лимнологического института Сибирского отделения Российской академии наук № 0279-2021-0011. Исследование также поддержано Российским научным фондом (проект № 19-14-00088) и Российским фондом фундаментальных исследований (проекты № 16-04-00065, 16-54-150007, 18-04-00224).

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

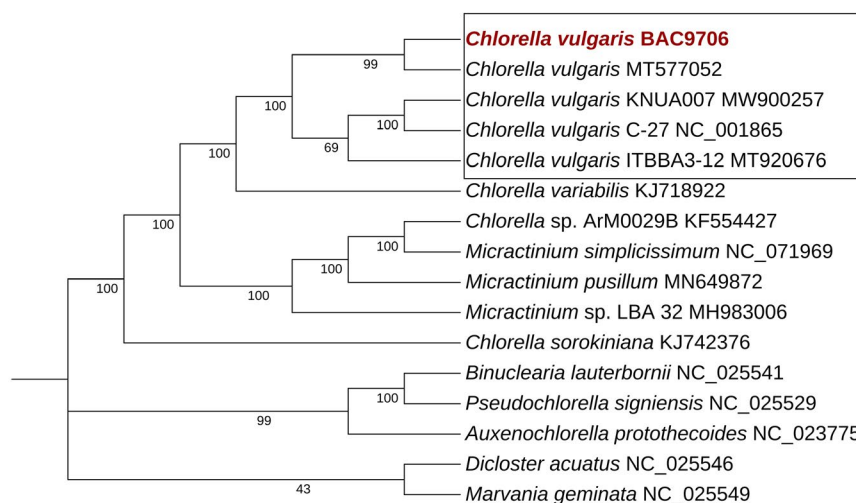


Рис.2. Филогенетическое дерево *C. vulgaris* BAC9706, построенное на основе конкатенированных белок-кодирующих генов хлоропластного генома представителей порядка Chlorellales.

Список литературы

- 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):455-77. DOI: [10.1089/cmb.2012.0021](https://doi.org/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](https://doi.org/10.1016/j.jglr.2020.03.015)
- Bondarenko N.A., Vorobyova S.S., Zhuchenko N.A. et al. 2020. Current state of phytoplankton in the littoral area of lake Baikal, spring 2017. *Journal of Great Lakes Research* 46(1): 17-28. DOI: [10.1016/j.jglr.2019.10.001](https://doi.org/10.1016/j.jglr.2019.10.001)
- Bondarenko N.A., Guseynikova N.E. 1989. Production of phytoplankton of Southern Baikal. *Bulletin of the Siberian Branch of the USSR Academy of Sciences* 1: 77-80.
- Cecchin M., Marcolungo L., Rossato M. et al. 2019. *Chlorella vulgaris* genome assembly and annotation reveal the molecular basis for metabolic acclimation to high light conditions. *Plant Journal* 100(6): 1289-1305. DOI: [10.1111/tpj.14508](https://doi.org/10.1111/tpj.14508)
- 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](https://doi.org/10.31951/2658-3518-2023-A-6-204)
- Gao S., Davis P., Liu K. et al. 1994. Asymmetric upwarp of the asthenosphere beneath the Baikal rift zone, Siberia. *Journal of Geophysical Research* 99: 15319-15330. DOI: [10.1029/94JB00808](https://doi.org/10.1029/94JB00808)
- Greiner S., Lehwark P., Bock R. 2019. Organellar-GenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* 47(W1). DOI: [10.1093/nar/gkz238](https://doi.org/10.1093/nar/gkz238)
- Han B., Mu Y., Tan D. 2021. The chloroplast genome of a unicellular green alga strain isolated from the rubber processing wastewater. *Mitochondrial DNA Part B: Resources* 6(1): 15-16. DOI: [10.1080/23802359.2020.1844090](https://doi.org/10.1080/23802359.2020.1844090)
- Kozhova O.M. 1959. About the subglacial "bloom" of phytoplankton in Lake Baikal. *Botanical Journal* 44(7): 1001.
- Kozhova O.M. 1987. Phytoplankton of Lake Baikal: structural and functional characteristics. *Arch. Hydrobiol. Beih. Ergebn. Limnol* 25: 19-37.
- 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](https://doi.org/10.1016/j.ygeno.2018.05.003)
- Martens K. 1997. Speciation in ancient lakes. *Trends In Ecology & Evolution* 12(5): 177-82. DOI: [10.1016/S0169-5347\(97\)01039-2](https://doi.org/10.1016/S0169-5347(97)01039-2)
- Nagata T., Takai K., Kawanobe K. et al. 1994. Autotrophic picoplankton in southern Lake Baikal: abundance, growth and grazing mortality during summer. *Journal of Plankton Research* 16(8): 945-959.
- Panova Y.A., Chernogor L.I., Belikov S.I. 2024. The complete mitochondrial genome of *Chlorella* sp. isolated from Lake Baikal. *Limnology and Freshwater Biology* 2024 (5): 1267-1275. DOI: [10.31951/2658-3518-2024-A-5-1267](https://doi.org/10.31951/2658-3518-2024-A-5-1267)
- 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): e00966-20. DOI: [10.1128/MRA.00966-20](https://doi.org/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](https://doi.org/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](https://doi.org/10.1093/nar/gkx391)
- Timoshkin O.A. 2011. Main tendencies in research of ancient Lake biodiversity; most interesting recent discoveries in biodiversity of Lake Baikal. In: Timoshkin O.A.(Ed.) *Index of animal species inhabiting Lake Baikal and its catchment area*. Novosibirsk: Nauka. pp. 1423-1428.
- 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.
- Wakasugi T., Nagai T., Kapoor M. et al. 1997. Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. *Proceedings of the National Academy of Sciences* 94(11): 5967-5972. DOI: [10.1073/pnas.94.11.5967](https://doi.org/10.1073/pnas.94.11.5967)
- Wen Y., Wan D. 2020. The complete chloroplast genome sequence of *Chlorella vulgaris* and phylogenetic analysis. *Mitochondrial DNA Part B: Resources* 5(3):2713-2715. DOI: [10.1080/23802359.2020.1787896](https://doi.org/10.1080/23802359.2020.1787896)