





# Genome Sequences of a Green-Colored *Chlorobium phaeovibrioides* Strain Containing Two Plasmids and a Closely Related Plasmid-Free Brown-Colored Strain

 Daria I. Boldyreva,<sup>a</sup>  Vladislav V. Babenko,<sup>a</sup> Alexandra V. Kanygina,<sup>a</sup> Olga N. Lunina,<sup>b</sup> Maria A. Letarova,<sup>b</sup> Elena S. Kostryukova,<sup>a</sup> Alexander S. Savvichev,<sup>b</sup> Vladimir M. Gorlenko,<sup>b</sup> Andrey V. Letarov<sup>b</sup>

<sup>a</sup>Federal Medical Biological Agency, Federal Research and Clinical Centre of Physical-Chemical Medicine, Moscow, Russia

<sup>b</sup>Research Centre of Biotechnology, Russian Academy of Sciences, Moscow, Russia

**ABSTRACT** Here, we report the draft genome sequences of the green sulfur bacterium *Chlorobium phaeovibrioides* strains GrTcv12 and PhvTcv-s14, isolated from the chemocline zone from meromictic Lake Trekhtzvetnoe, separated from the White Sea, in Russia. This is the first report showing the presence of plasmids containing antiphage systems in the *Chlorobium* sp. genome.

In Lake Trekhtzvetnoe (Russia), brown-colored *Chlorobium phaeovibrioides* PhvTcv-s14 is located in the upper layers of the chemocline zone above green-colored *Chlorobium phaeovibrioides* GrTcv12, which has particular phenotype distinctions (1). Both strains were isolated from the chemocline of Lake Trekhtzvetnoe, which is the smallest meromictic lake with constant stratification. This lake has constant hydrological and biological stratification over a short water column and a high density of *C. phaeovibrioides* in the chemocline layer (2).

The strains were grown as described previously (1). The DNA was extracted by using the Wizard DNA extraction kit (Promega Corporation, USA) and size selected with optimized solid-phase reversible immobilization (SPRI) beads (3). The long reads were generated with MinION sequencing (Oxford Nanopore Technologies, UK). The sequencing libraries were prepared using the ligation sequencing kit (catalog number SQK-LSK109) and native barcoding expansion kit (catalog number EXP-NBD114) and run in a FLO-MIN106 flow cell. Reads were base called using Albacore v1.2.5 and trimmed and demultiplexed with Porechop v0.2.1 using default parameters (4). The short-read whole-genome sequencing (WGS) for each strain was generated using the Ion Torrent PGM (Life Technologies, USA) sequencing platform with the Ion Xpress plus fragment library kit (Life Technologies) using 400-bp chemistry. Prinseq lite v0.20.4 was used for read-quality ( $Q > 20$ ) trimming. *De novo* assembly was performed by hybrid assembler Unicycler (v0.4.8) (5) using default parameters. Identification of the protein-coding sequences and primary annotation were performed using PGAP v4.7 (6). All relevant sequencing and assembly statistics are summarized in Table 1.

PhvTcv-s14 was assembled as a single circular chromosome. GrTcv12 was assembled as two chromosome contigs (circular and linear) and two plasmids (p1 and p2). The coverages of the chromosome, p1, and linear contig were almost identical, while the coverage of p2 exceeded the chromosome coverage by a factor of 4.5. The p2 plasmid was previously detected in a shotgun metagenome of the Lake Trekhtzvetnoe chemocline water due to its elevated coverage (2).

Among known bacterial genomes of the phylum *Chlorobi*, only *Prosthecochloris aestuarii* has been identified to carry a plasmid (NCBI reference sequence accession number NC\_011061). Thus, we can suppose that the presence of plasmids in the genome is not typical for *Chlorobium* spp. Plasmid p1 contains genes related to the

**Citation** Boldyreva DI, Babenko VV, Kanygina AV, Lunina ON, Letarova MA, Kostryukova ES, Savvichev AS, Gorlenko VM, Letarov AV. 2020. Genome sequences of a green-colored *Chlorobium phaeovibrioides* strain containing two plasmids and a closely related plasmid-free brown-colored strain. *Microbiol Resour Announc* 9:e01172-19. <https://doi.org/10.1128/MRA.01172-19>.

**Editor** Kenneth M. Stedman, Portland State University

**Copyright** © 2020 Boldyreva et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Daria I. Boldyreva, [dasha.boldyreva@yandex.ru](mailto:dasha.boldyreva@yandex.ru).

**Received** 19 September 2019

**Accepted** 9 December 2019

**Published** 9 January 2020

**TABLE 1** Sequencing and assembly statistics for *C. phaeovibrioides* PhvTcv-s14 and GrTcv12

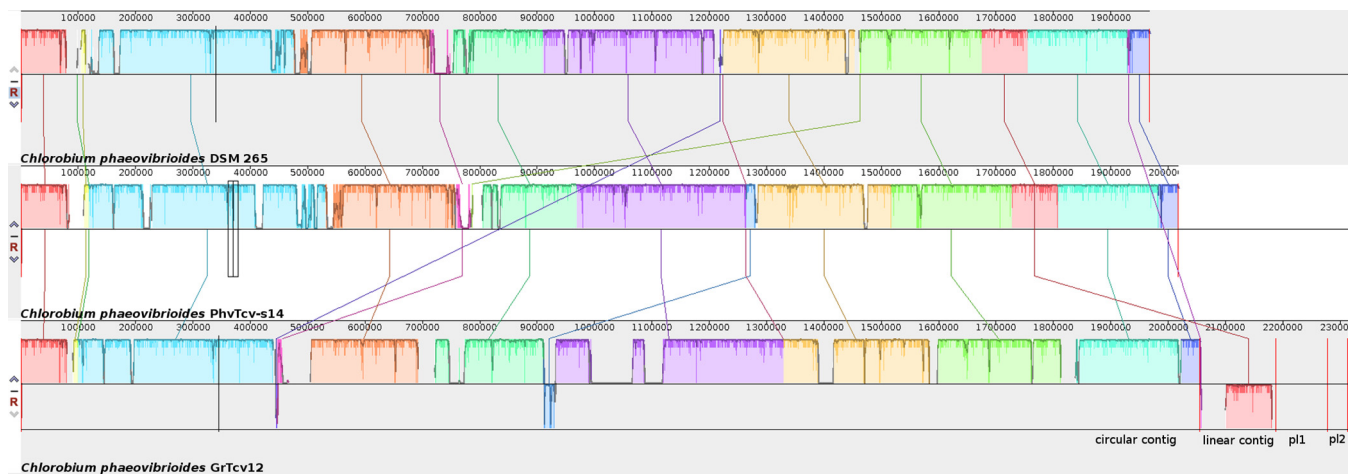
Parameter	Data for strain:					
	PhvTcv-s14	GrTcv12				Whole genome
		pl1	pl2	Linear contig	Circular chromosome	
Ion Torrent PGM						
No. of generated reads	404,071					1,001,673
Mean read length (bp)	218					216
MinION						
No. of generated reads	199,320					348,897
Mean read length (bp)	7,351					5,090
Sequence length (bp)	2,017,051	90,349	35,536	132,296	2,055,276	
G+C content (%)	53	49.4	50.2	52.6	52.7	
No. of CDSs <sup>a</sup>	1,842	77	29	139	1,704	
Coverage (×)	750					800
No. of genes	1,903					2,207
No. of pseudogenes	254					202
No. of RNA genes	61					56

<sup>a</sup> CDSs, coding DNA sequences.

type I-F CRISPR-Cas system and serine/threonine proteins (7, 8). Plasmid pl2 contains genes related to AbiEii/AbiGii toxin family proteins and to the restriction-modification system. The linear contig also encodes a bacteriophage exclusion system, BREX, a prophage-related region, and some housekeeping genes. The presence of multiple (at least 5) antiphage systems in the *C. phaeovibrioides* GrTcv12 genome suggests a significant load of the phage infection in the recent evolutionary history of this strain, which is in good agreement with the previously obtained data (9).

Comparative analysis of GrTcv12 and PhvTcv-s14 with *C. phaeovibrioides* DSM 265 (10) using Mauve (11) shows that these genomes are very close to each other (Fig. 1).

The average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) values were calculated using the ANI calculator (12) and the Genome-to-Genome Distance Calculator (GGDC) v 2.1 (13), respectively. In comparison with *C. phaeovibrioides* DSM 265, the ANI values of strains GrTcv12 and PhvTcv-s14 were 98.99%, and 98.93% and the dDDH values were 91.10% and 89.50%, respectively. The calculated



**FIG 1** Mauve genome comparison between the reference genome *C. phaeovibrioides* DSM 265 (upper) and the studied genomes *C. phaeovibrioides* PhvTcv-s14 (middle) and *C. phaeovibrioides* GrTcv12 (lower). Each contiguously colored region is a locally colinear block (LCB), a region without rearrangement of homologous backbone sequence. LCBs identified by Mauve are color coded; links between LCBs are indicated by the thin colored lines. LCBs below a genome's center line are in the reverse complement orientation relative to the reference DNA sequence. Unmatched regions within an LCB indicate the presence of a strain-specific sequence. The contigs are separated by red lines. The scale is in nucleotides.

values exceeded the proposed boundary values for species delineation (ANI, 95% to 96%; dDDH, 70%) (14), which suggests that strains GrTcv12 and PhvTcv-s14 are novel strains of the known species *Chlorobium phaeovibrioides*. Also, they have similarities with *C. phaeovibrioides* GrKhr17 and BrKhr17 from the neighboring Lake Bolshye Khruslomeny (15). Further study of green sulfur bacteria (GSB) in this area can give insight into how phages influence bacterial genome changes during evolution.

The genomes of GrTcv12 and PhvTcv-s14 contain the *gvp* genes that are required for gas vesicle formation. Also, both genomes contain genes required for biosynthesis of Bchl *a*, *b*, *c*, and *d* and genes which are responsible for the biosynthesis of isorenieratene and chlorobactin (16, 17). There are *bciD* genes required for Bchl *e* synthesis in the PhvTcv-s14 genome. Sequencing and analysis of these bacteria revealed genomic determinants of the particular phenotypes of new strains of *Chlorobium phaeovibrioides* from Arctic meromictic lakes.

**Data availability.** Genome sequence data of *Chlorobium phaeovibrioides* PhvTcv-s14 and *Chlorobium phaeovibrioides* GrTcv12 were deposited into NCBI GenBank under BioSample accession numbers [SAMN09466660](https://www.ncbi.nlm.nih.gov/biosample/SAMN09466660) and [SAMN09466659](https://www.ncbi.nlm.nih.gov/biosample/SAMN09466659), respectively, and under BioProject accession number [PRJNA438928](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA438928). Raw sequence reads are available under the SRA accession numbers [SRR10277008](https://www.ncbi.nlm.nih.gov/sra/SRR10277008) (MinION) and [SRR10277009](https://www.ncbi.nlm.nih.gov/sra/SRR10277009) (IonTorrent PGM) for *Chlorobium phaeovibrioides* PhvTcv-s14 and [SRR10277006](https://www.ncbi.nlm.nih.gov/sra/SRR10277006) (MinION) and [SRR10277007](https://www.ncbi.nlm.nih.gov/sra/SRR10277007) (IonTorrent PGM) for *Chlorobium phaeovibrioides* GrTcv12.

## ACKNOWLEDGMENT

This study was partially supported by the Russian Foundation for Basic Research (RFBR) (research project number 17-04-01263).

## REFERENCES

- Lunina ON, Savvichev AS, Babenko VV, Boldyreva DI, Kuznetsov BB, Kolganova TV, Krasnova ED, Kokryatskaya NM, Veslopolova EF, Voronov DA, Demidenko NA, Letarova MA, Letarov AV, Gorlenko VM. 2019. Seasonal variations in the structure of an anoxygenic phototrophic bacterial community from the meromictic Lake Trekhtsvetnoe (Kandalaksha Bay, White Sea). *Microbiology* 88:100–114. <https://doi.org/10.1134/S0026261719010041>.
- Savvichev AS, Babenko VV, Lunina ON, Letarova MA, Boldyreva DI, Veslopolova EF, Demidenko NA, Kokryatskaya NM, Krasnova ED, Gaisin VA, Kostyukova ES, Gorlenko VM, Letarov AV. 2018. Sharp water column stratification with an extremely dense microbial population in a small meromictic lake, Trekhtsvetnoe. *Environ Microbiol* 20:3784–3797. <https://doi.org/10.1111/1462-2920.14384>.
- Schalamun M, Nagar R, Kainer D, Beavan E, Eccles D, Rathjen JP, Lanfer R, Schwesinger B. 2019. Harnessing the MinION: an example of how to establish long-read sequencing in a laboratory using challenging plant tissue from *Eucalyptus pauciflora*. *Mol Ecol Resour* 19:77–89. <https://doi.org/10.1111/1755-0998.12938>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom* 3:e000132. <https://doi.org/10.1099/mgen.0.000132>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Cohen PT, Cohen P. 1989. Discovery of a protein phosphatase activity encoded in the genome of bacteriophage lambda. Probable identity with open reading frame 221. *Biochem J* 260:931–934. <https://doi.org/10.1042/bj2600931>.
- Depardieu F, Didier J-P, Bernheim A, Sherlock A, Molina H, Duclos B, Bikard D. 2016. A eukaryotic-like serine/threonine kinase protects *Staphylococcus* against phages. *Cell Host Microbe* 20:471–481. <https://doi.org/10.1016/j.chom.2016.08.010>.
- Llorens-Marès T, Liu Z, Allen LZ, Rusch DB, Craig MT, Dupont CL, Bryant DA, Casamayor EO. 2017. Speciation and ecological success in dimly lit waters: horizontal gene transfer in a green sulfur bacteria bloom unveiled by metagenomic assembly. *ISME J* 11:201–211. <https://doi.org/10.1038/ismej.2016.93>.
- Imhoff JF. 2003. Phylogenetic taxonomy of the family Chlorobiaceae on the basis of 16S rRNA and *fmo* (Fenna-Matthews-Olson protein) gene sequences. *Int J Syst Evol Microbiol* 53:941–951. <https://doi.org/10.1099/ijs.0.02403-0>.
- Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics* 25:2071–2073. <https://doi.org/10.1093/bioinformatics/btp356>.
- Rodriguez-R LM, Konstantinidis KT. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ PrePr* 4:e1900v1. <https://doi.org/10.7287/peerj.preprints.1900v1>.
- Auch AF, von Jan M, Klenk H-P, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2:117–134. <https://doi.org/10.4056/signs.531120>.
- Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu X-W, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68:461–466. <https://doi.org/10.1099/ijsem.0.002516>.
- Grouzdev DS, Lunina ON, Gaisin VA, Krutkina MS, Baslerov RV, Savvichev AS, Gorlenko VM, Grouzdev DS, Lunina ON, Gaisin VA, Krutkina MS, Baslerov RV, Savvichev AS, Gorlenko VM. 2019. Genome sequences of green- and brown-colored strains of *Chlorobium phaeovibrioides* with gas vesicles. *Microbiol Resour Announc* 8:e00711-19. <https://doi.org/10.1128/MRA.00711-19>.
- Maresca JA, Romberger SP, Bryant DA. 2008. Isorenieratene biosynthesis in green sulfur bacteria requires the cooperative actions of two carotenoid cyclases. *J Bacteriol* 190:6384–6391. <https://doi.org/10.1128/JB.00758-08>.
- Maresca JA, Bryant DA. 2006. Two genes encoding new carotenoid-modifying enzymes in the green sulfur bacterium *Chlorobium tepidum*. *J Bacteriol* 188:6217–6223. <https://doi.org/10.1128/JB.00766-06>.