

Complete Genome Sequence of an Uncultured Bacterium of the Candidate Phylum *Bipolaricaulota*

V. V. Kadnikov^{a,b,*}, A. V. Mardanov^a, A. V. Beletsky^a, Y. A. Frank^c, O. V. Karnachuk^c, and N. V. Ravin^{a,b}

^aInstitute of Bioengineering, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, 119071 Russia

^bFaculty of Biology, Moscow State University, Moscow, 119192 Russia

^cTomsk State University, Tomsk, 634050 Russia

*e-mail: vkadnikov@bk.ru

Received December 25, 2018; revised January 6, 2019; accepted January 7, 2019

Abstract—Bacteria of the candidate phylum *Bipolaricaulota*, previously known as OP1 and *Acetothermia*, have been identified in various terrestrial and marine ecosystems as a result of molecular analysis of microbial communities. To date, none of the members of *Bipolaricaulota* has been isolated in a pure culture. We assembled the complete closed genome sequence of the bacterium of this phylum from the metagenome of a subsurface thermal aquifer and used the genomic data to analyze the phylogenetic position and metabolic potential of this bacterium. Analysis of the genome of this bacterium, designated Ch78, revealed various transporters of sugars and peptides and limited capacities for the extracellular hydrolysis of polysaccharides. Reconstruction of the central metabolic pathways suggested that the bacterium Ch78 is an anaerobic organotroph capable of fermenting sugars and proteinaceous substrates. Known pathways of aerobic and anaerobic respiration, as well as of autotrophic carbon fixation, were not found in the Ch78 genome. In the subsurface aquifer *Bipolaricaulota* probably consume low-molecular weight organic compounds and produce hydrogen and acetate. Based on the results of phylogenetic and genomic analysis, classification of the novel bacterium as “*Candidatus Bipolaricaulis sibiricus*” is proposed.

Keywords: candidate phylum OP1, *Bipolaricaulota*, *Acetothermia*, metagenome, complete genome

DOI: 10.1134/S0026261719040064

The 16S rRNA gene sequences of the candidate phylum OP1 were originally revealed in the sediments of the Obsidian Pool hot spring (75–95°C), Yellowstone National Park, United States (Hugenholtz et al., 1998). This is one of the first 12 candidate phyla described as a result of molecular analysis of microbial diversity. The 16S rRNA gene sequences assigned to this candidate phylum were subsequently detected in various terrestrial and marine ecosystems, including deep-sea hydrotherms (Teske et al., 2002; Stott et al., 2008; Kato et al., 2009), hot springs (Costa et al., 2009; Tobler and Benning, 2011), anaerobic bioreactors (Nobu et al., 2015), and groundwater (Hirayama et al., 2005). In spite of its wide occurrence in nature, the share of OP1 in most habitats did not exceed 5%. The SILVA database (Quast et al., 2013) contains presently over 1000 16S rRNA gene sequences assigned to the candidate phylum OP1.

Although no OP1 member has been isolated in pure culture, some information on the properties of bacteria of this candidate phylum was obtained by sequencing of metagenomes and single-cell genomes. The Genome Taxonomy Database (GTDB, Parks et al., 2018) contains presently 24 OP1 genomes, although only one of them is complete. The first

incomplete genome (MAG, metagenome-assembled genome) of an organism of this group was obtained from the metagenome of a hot spring microbial mat (Takami et al., 2012). Its analysis revealed that this bacterium is acetogenic and possess the Wood–Ljungdahl folate-dependent pathway of CO₂ fixation. The name “*Candidatus Acetothermum autotrophicum*” was therefore proposed for this organism (Takami et al., 2012), and the name *Acetothermia* was proposed for the candidate phylum OP1 (Rinke et al., 2013). Another MAG (*Acetothermia* bacterium 64_32) was obtained from the metagenome of the sediments of a sea shelf oil deposit (Hu et al., 2016). These genome lacked the key genes encoding autotrophic CO₂ fixation pathways, which indicated heterotrophic metabolism. The only complete genome (one circular chromosome) of an OP1 bacterium was obtained from the metagenome of an anaerobic bioreactor for organic waste treatment (Hao et al., 2018). Genome annotation and reconstruction of the metabolic pathways indicated anaerobic chemoheterotrophic metabolism, with carbon and energy obtained via fermentation of peptides, amino acids, and simple sugars to acetate, formate, and hydrogen. The name “*Candidatus Bipolaricaulis anaerobius*” was proposed for this

bacterium due to its unusual cell morphology, and renaming of the OP1 phylum to “*Candidatus* Bipolaricaulota” was proposed due to nomenclature problems with the previously used name (Hao et al., 2018).

Deep subsurface ecosystems are extreme habitats with high temperature and pressure. Microbial communities of these ecosystems are usually characterized by the presence of various uncultured bacterial and archaeal groups. Partial or complete genomes have been determined for many of them using metagenomics or single-cell genome sequencing (Anantharaman et al., 2016; Magnabosco et al., 2016; HERNSDORF et al., 2017; Probst et al., 2017). We have previously investigated microbial communities of groundwater in Western Siberia, Russia. These subsurface thermal waters located in Cretaceous sedimentary rocks at the depth of 1 to 3 km are available due to numerous oil exploration wells, through which groundwater flows to the surface under natural pressure. We analyzed the composition of microbial communities (Kadnikov et al., 2017a, 2017b) and sequenced the metagenome of groundwater flowing out through the 5P borehole in Tomsk region, Russia (Kadnikov et al., 2017c). Approximately 50% of this community consisted of methanogenic archaea, while its other half was represented by bacteria of various uncultured lineages (Kadnikov et al., 2017a). Analysis of the community composition based on the 16S rRNA gene sequences revealed bacteria of the phylum *Bipolaricaulota*, and their share in the community was about 1.5%.

The goal of the present work was to use the results of metagenomic analysis in order to determine the complete genome (at the single chromosome level) of a new member of *Bipolaricaulota* and to characterize its metabolic pathways and global distribution.

MATERIALS AND METHODS

Well characterization, sampling, and metagenomic DNA isolation. Water was collected from the oil-exploration well 5P, drilled in the 1950s to the depth of 2.8 km in the vicinity of the Chazhemto settlement, Tomsk oblast, Russia. The water sample (20 L) was collected in April 2016 (Kadnikov et al., 2017c). The water temperature was ~20°C, it had near-neutral pH (7.43–7.6), and a negative redox potential (–304 to –338 mV). Sulfate concentration was low (90.4 mg/L; Kadnikov et al., 2017b).

To collect microbial biomass, the sample was filtered through 0.22- μ m cellulose nitrate membranes. The filters were homogenized by grinding with liquid nitrogen, and the total community DNA was isolated using the Power Soil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, United States). A total of ~1 μ g DNA was obtained.

Sequencing of metagenomic DNA using Illumina platform and assembly of microbial genomes. Sequencing of metagenomic DNA using Illumina HiSeq2500

(Illumina, United States) was described previously (Kadnikov et al., 2017c). Sequencing (in the paired reads format, 2 \times 250 nt) with subsequent filtration ($Q > 33$) resulted in obtaining DNA sequences with total length of ~16.9 Gbp. The sequences were assembled into contigs using metaSPAdes v. 3.7.1. The contigs were then binned into clusters corresponding to individual microbial genomes (MAG) using CONCOCT (Alneberg et al., 2014). Completeness of assembled MAGs and the degree of their contamination (presence of the contigs belonging to other microorganisms) were determined using CheckM (Parks et al., 2015).

Taxonomic position of assembled genomes was determined according to the GTDB database using GTDB-Tk v. 0.1.3 (Parks et al., 2018). One of the genomes, designated Ch78, was identified as belonging to a *Bipolaricaulota* member.

Sequencing of metagenomic DNA using the MinION system and assembly of a complete genome of a member of *Bipolaricaulota*. Metagenomic DNA was additionally sequenced on MinION (Oxford Nanopore, United Kingdom) using Ligation Sequencing kit 1D protocol according to the manufacturer’s recommendations. Sequencing resulted in 1418419 reads with the total length of ~1.54 Gbp. These long reads were used to join the contigs belonging to the Ch78 genome into longer sequences. For this purpose, the MinION reads exhibiting high homology to the sequences of the Ch78 contigs were selected using BWA v. 0.7.15 (Li and Durbin, 2009). The contigs were combined using npScarf (Cao, 2017). The possible errors in the consensus sequence were corrected with Pilon 1.22 (Walker et al., 2014) basing on mapping of Illumina reads performed using Bowtie 2 (Langmead and Salzberg, 2012). The genome sequence of the Ch78 bacterium was deposited to NCBI GenBank under accession no. CP034928.

Similarity of genomic sequences. The average amino acid identity (AAI) between the genomes was determined using the aai.rb script from the Enveomics Collection (Rodriguez-R and Konstantinidis, 2016). The values of DNA–DNA hybridization *in silico* were calculated using the GGDC2 tool (Meier-Kolthoff et al., 2013) available at <http://ggdc.dsmz.de/>, using the recommended formula 2.

Genome annotation and analysis. Search for the genes and their annotation were carried out using the RAST server; the annotation was then checked and manually corrected by comparing the predicted protein sequences with the NCBI databases. The N-terminal signal peptides were predicted using Signal P v. 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>) and PRED-TAT (<http://www.compugen.org/tools/PRED-TAT/>); transmembrane domains were identified using TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). The index of genomic DNA repli-

Table 1. Main characteristics of the *Bipolaricaulota* genomes

Organism/genome	GenBank	Total scaffold length, nt	Genome completeness, %	Scaffold number
<i>Candidatus</i> <i>Bipolaricaulis</i> anaerobius Ran1	LS483254	1324338	100	1
<i>Candidatus</i> <i>Bipolaricaulis</i> sibiricus Ch78	CP034928	1701655	100	1
<i>Acetothermia</i> bacterium 64_32	GCA_001508155	1810203	89.83	33
<i>Candidatus</i> <i>Acetothermum</i> autotrophicum	AP011800–AP011803	1970005	77.97	4
<i>Candidatus</i> <i>Fraserbacteria</i> bacterium RBG_16_55_9	GCA_001778455	1955059	76.27	133
<i>Acetothermia</i> bacterium UBA7950	GCA_003535305	1040075	75.42	165
<i>Acetothermia</i> bacterium UBA9294	GCA_003545155	1055817	77.12	144
<i>Candidatus</i> <i>Acetothermia</i> bacterium JdFR-49	GCA_002010355	1698648	68.41	312
<i>Candidatus</i> <i>Acetothermia</i> bacterium JdFR-50	GCA_002010365	1292279	58.63	255
<i>Candidatus</i> <i>Acetothermia</i> bacterium JdFR-52	GCA_002010385	1710598	86.44	155
<i>Candidatus</i> <i>Acetothermia</i> bacterium JdFR-46	GCA_002011365	1088071	54.01	254
<i>Candidatus</i> <i>Acetothermia</i> bacterium JdFR-47	GCA_002011415	1622290	82.33	230
<i>Candidatus</i> <i>Acetothermia</i> bacterium JdFR-48	GCA_002011425	1893129	91.53	62
<i>Candidatus</i> <i>Acetothermia</i> bacterium JdFR-51	GCA_002011435	1763844	88.14	185
<i>Acetothermia</i> bacterium UBA3574	GCA_002375935	1576779	91.53	83
<i>Acetothermia</i> bacterium UBA3571	GCA_002375995	1842568	91.53	46
<i>Acetothermia</i> bacterium UBA3565	GCA_002376075	1833941	88.14	96
<i>Acetothermia</i> bacterium UBA3564	GCA_002376085	1497836	83.9	72
<i>Acetothermia</i> bacterium UBA3563	GCA_002376125	1519271	80.35	162
<i>Acetothermia</i> bacterium UBA3560	GCA_002376485	1872529	88.14	51
<i>Acetothermia</i> bacterium UBA4801	GCA_002402715	1727801	83.05	125
<i>Acetothermia</i> bacterium UBA4777	GCA_002403185	1150656	82.2	70
<i>Acetothermia</i> bacterium UBA7521	GCA_002479455	1790708	86.44	68

cation was calculated using the iRep software (Brown et al., 2016). Hydrogenases were classified using the HydDB online service (<https://services.birc.au.dk/hyddb/>) (Søndergaard et al., 2016).

Phylogenetic analysis. The dataset used for genome-based phylogenetic analysis included the Ch78 genome, genomes of 21 other members of *Bipolaricaulota* (Table 1), genomes of the members of *Synergistetes*, *Deinococcus-Thermus*, and *Thermotogae*, members of the candidate phyla *Aerophobetes* and *Calescamantes*, and the genome of *Aquifex aeolicus*. In each of these genomes, 43 conservative single-copy marker genes were identified using CheckM, and multiple alignment of concatenated sequences of these marker genes was carried out. This multiple alignment was used to construct the phylogenetic tree by the Maximum Likelihood method using PhyML v. 3.3 (Guindon et al., 2010) with the default parameters.

RESULTS AND DISCUSSION

Genome assembly of a member of the candidate phylum *Bipolaricaulota*. Sequencing of the metagenome of the microbial community from well 5P groundwater

using the Illumina technology resulted in generation of ~16.9 Gbp assembled into contigs (Kadnikov et al., 2017c). The contigs were clustered using CONCOCT. One of the MAGs obtained, designated Ch78, was represented by 6 contigs with a total length of 1.7 Mbp, with an average 55-fold sequence coverage. The relative abundance of this genotype in the community, defined as a fraction of this MAG in the metagenome, was about 0.6%. Analysis of the taxonomic position of this genome using Genome Taxonomy database supported its affiliation with the candidate phylum *Bipolaricaulota* (OP1).

Complete genome of the Ch78 bacterium was assembled using 1.4×10^6 long reads (~1.5 Gbp in total) obtained by monomolecular nanopore sequencing. As a result, six contigs were assembled into a single circular 1701655 nt-long sequence. According to CheckM assessment, this genome was 100% complete in the absence of possible contamination. This is the second known complete genome of a member of *Bipolaricaulota*.

Analysis of the Ch78 genome revealed the 5S–23S rRNA gene operon, a separately located 16S rRNA gene, and 47 transfer RNA genes. Based on genome

annotation, 1668 potential protein-encoding genes were identified, of which the functions of 808 could be predicted by comparison with NCBI databases. The Ch78 genome contained a single CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) locus, containing 50 repeat–spacer units, and a set of genes encoding the type I-C CRISPR system. Genome size and gene number of Ch78 were close to those of other *Bipolaricaulota* (Table 1).

The rate of DNA replication *in situ* (in groundwater) for Ch78 was evaluated using the iRep software (Brown et al., 2016). The iRep replication index of 1.15 was calculated for the Ch78 genome, indicating slow growth of these bacteria (~15% of cells were actively replicating at the time of sampling), but confirming them to be a metabolically active part of the microbial community.

Phylogenetic position of Ch78. Search for phylogenetically related microorganisms based on genome similarity revealed “*Candidatus Bipolaricaulis anaerobius*” Ran1 (Hao et al., 2018) to be the closest relative of Ch78 with an average AAI of 76.08%. The AAI values between Ch78 and other *Bipolaricaulota* genomes did not exceed 60.1%. According to the AAI thresholds proposed by Konstantinidis et al. (2017) for determination of the phylogenetic position of uncultured microorganisms, Ch78 and “*Ca. Bipolaricaulis anaerobius*” belong to different species of the same genus. The estimated value of their DNA–DNA hybridization *in silico* (18.4–23.0%) also supports assigning these microorganisms to different species. Identity of the 16S rRNA gene sequences of Ch78 and “*Ca. Bipolaricaulis anaerobius*” Ran1 was 97%.

A phylogenetic tree based on concatenated sequences of 43 conservative marker genes from Ch78 and other available *Bipolaricaulota* genomes was constructed in order to elucidate the phylogeny of this phylum. The results obtained (Fig. 1) confirm that “*Ca. Bipolaricaulis anaerobius*” is the closest relative of Ch78. The previously characterized “*Ca. Acetothermum autotrophicum*” falls into a cluster together with “*Ca. Fraserbacteria bacterium*” RBG_16_55_9 and forms a separate remote branch within the candidate phylum *Bipolaricaulota*.

Analysis of metabolic pathways of Ch78. Analysis of the Ch78 genome revealed the genes of ABC type transporters responsible for uptake of sugars, amino acids, and peptides into the cell. Search for glycosyl hydrolases containing the N-terminal secretion signal peptides revealed only two α -amylases, probably involved in extracellular hydrolysis of starch and its derivatives. Ch78 also possessed ABC type transporters for maltose/maltodextrin import and intracellular enzymes for utilization of these carbohydrates (maltodextrin glucosidase, α -amylase, amyloamylase, and α -glucosidase). No genes of other hydrolytic enzymes involved in polysaccharide hydrolysis were detected. Imported sugars may be metabolized via the

Embden–Meyerhof pathway (Fig. 2). Genome analysis revealed the genes of all enzymes of glycolysis, gluconeogenesis, and the nonoxidative branch of the pentose phosphate pathway. Glycogen may be used by Ch78 as a storage polysaccharide; the pathways for its synthesis and degradation are encoded in its genome. Pyruvate produced by glycolysis may be converted to acetyl-CoA by pyruvate-ferredoxin oxidoreductase. Oxidation of acetyl-CoA with formation of acetate and generation of ATP may be carried out by ADP-producing acetyl-CoA synthetase. The tricarboxylic acids cycle is not closed (succinate dehydrogenase is absent) and is probably used only for biosynthetic purposes.

Amino acids and peptides may also be used as carbon source; their import may be carried out by the ABC type transporters specific for amino acids and oligo- or dipeptides. The genes encoding potentially secreted proteolytic enzymes were not found. Amino acids (imported or formed by intracellular peptide hydrolysis) may be deaminated and converted to 2-keto acids (pyruvate, 2-oxoglutarate, etc.). The 2-keto acids may be oxidized to acyl-CoA derivatives by ferredoxin-dependent oxidoreductases with various specificities; the latter may be then cleaved with production of ATP and relative acids (Fig. 2).

Since the genome of “*Ca. Acetothermum autotrophicum*”, the first characterized OP1 member, encoded the complete Wood–Ljungdahl pathway of CO₂ fixation, the possibility of its autotrophic metabolism was proposed (Takami et al., 2012). The key enzymes of this pathway, including the CO dehydrogenase and acetyl-CoA synthetase complex, were not found in the genome of Ch78. Genome analysis of other OP1 members revealed the Wood–Ljungdahl pathway in the closest relative of “*Ca. Acetothermum autotrophicum*,” “*Ca. Fraserbacteria bacterium*” RBG_16_55_9, but not in other OP1 members. The metabolic capabilities of Ch78 are probably limited to fermentation of organic compounds, while the known pathways of aerobic or anaerobic respiration are absent.

The presence of several NiFe hydrogenases is an evidence of important role of hydrogen in Ch78 metabolism (Fig. 2). The first one, belonging to group 3c, is a component of a cytoplasmic complex with heterodisulfide reductase (Hdr-Mvh complex), which may carry out H₂ oxidation coupled to heterodisulfide and ferredoxin reduction due to electron bifurcation (Greening et al., 2016). The second hydrogenase (Mbh), of group 4d, is a multisubunit complex associated with the cytoplasmic membrane. Hydrogenases of this group oxidize reduced ferredoxin with production of molecular hydrogen and transport the H⁺ or Na⁺ ions through the membrane, generating the transmembrane ion gradient (Greening et al., 2016). The third enzyme (Hox) belongs to group 3d.

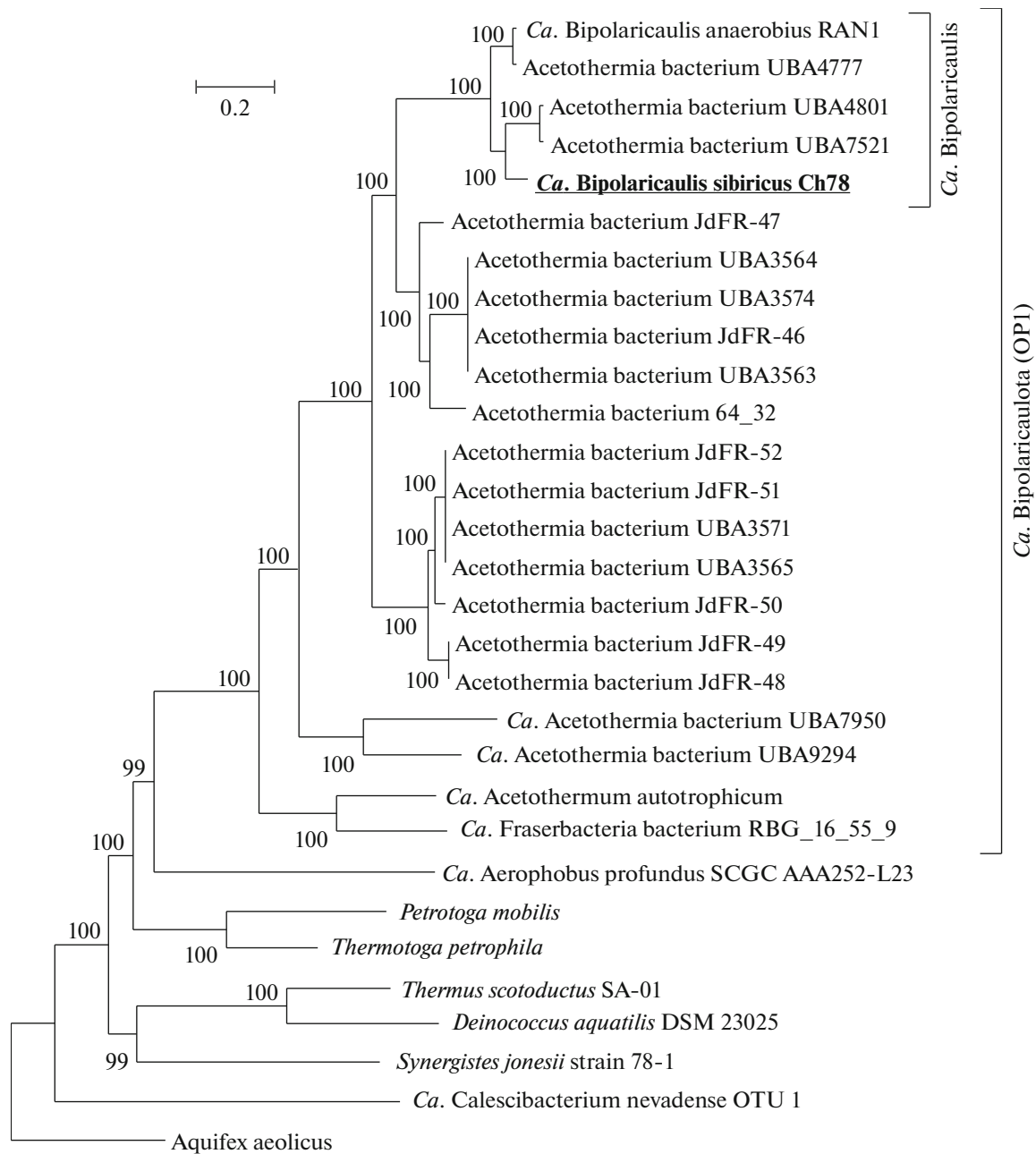


Fig. 1. Position of “*Candidatus Bipolaricaulis sibiricus*” Ch78 on the phylogenetic tree constructed using concatenated sequences of 43 conservative single-copy marker genes.

Depending on the cell redox status, these cytoplasmic hydrogenases may either oxidize NADH and reduce protons to H₂ or to carry out the reverse reaction (Greening et al., 2016). Activities of these hydrogenases can not only drive the reoxidation of NADH and reduced ferredoxin produced during the course of fermentation, but also generate the transmembrane ion gradient, that can be used by the membrane F₀F₁-type ATP synthase for ATP synthesis. Apart from the group 4d NiFe hydrogenase, Na⁺-transporting membrane pyrophosphatase may also be involved in formation of

the transmembrane ion gradient. Overall, Ch78 has limited abilities of generation of the transmembrane ion gradient because it has no NADH dehydrogenase or other ion pumps.

Interestingly, the genome of “*Ca. Bipolaricaulis anaerobius*” encodes a group 4d (NiFe)-hydrogenase capable of generating the transmembrane ion gradient but does not encode ATP synthase. Homologs of the Ch78 ATP synthase were found in many OPI members, which may indicate its loss in “*Ca. Bipolaricaulis anaerobius*.” Comparison of the genome region flank-

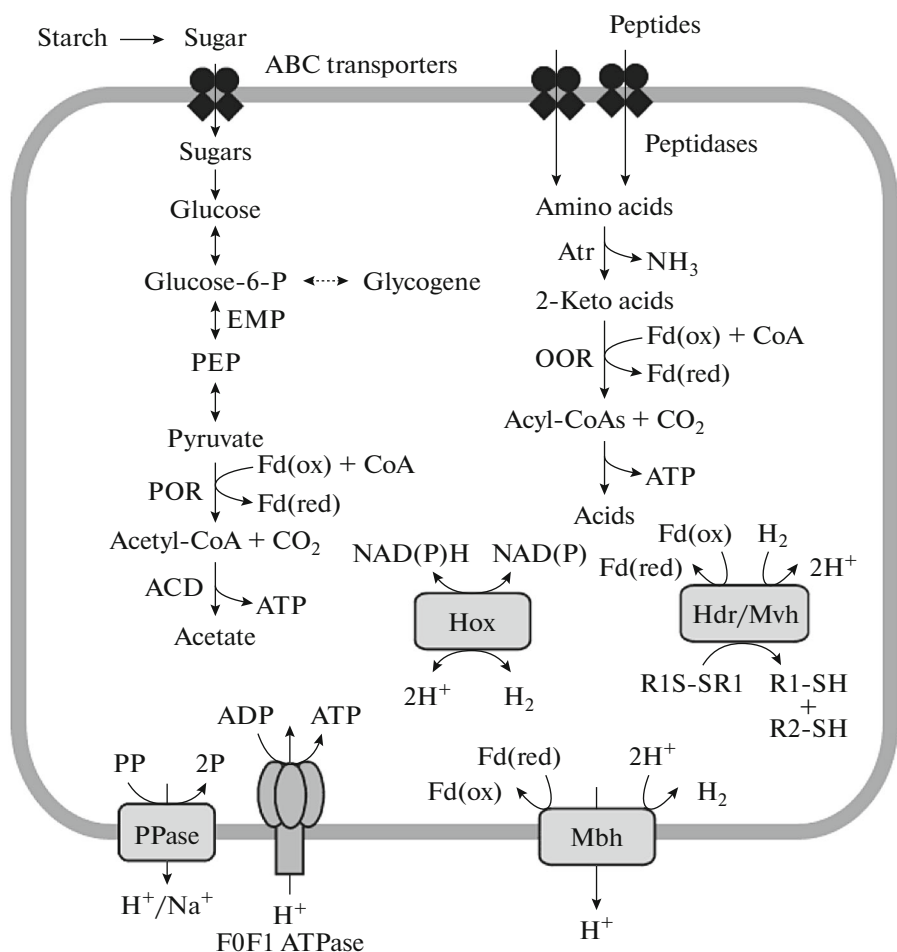


Fig. 2. The main metabolic pathways of “*Candidatus Bipolaricaulis sibiricus*” Ch78. Enzyme designations: Atr, aminotransferases; POR, pyruvate-ferredoxin oxidoreductase; OOR, ferredoxin-dependent oxidoreductases of 2-keto acids; ACD, acetyl-CoA synthetase (ADP-forming); Hdr/Mvh, complex of heterodisulfide reductase and group 3c hydrogenase; Mbh, group 4d hydrogenase; Hox, group 3d hydrogenase; PPase, pyrophosphatase. Other designations: EMP, Embden–Meyerhof pathway; PEP, phosphoenolpyruvate; Fd(ox), oxidized ferredoxin; Fd(red), reduced ferredoxin; CoA, coenzyme A; PP, pyrophosphate; P, phosphate.

ing the ATP synthase gene cluster in Ch78 and the corresponding region in “*Ca. Bipolaricaulis anaerobius*” genome revealed their similarity and deletion of ATP synthase locus in “*Ca. Bipolaricaulis anaerobius*.” Thus “*Ca. Bipolaricaulis anaerobius*” probably uses the transmembrane gradient only for active transport, while ATP could be formed only in the reactions of substrate phosphorylation.

Global distribution of *Bipolaricaulota* and their ecological role. Although the SILVA database (Quast et al., 2013) contains 1054 16S rRNA gene sequences representing the phylum *Bipolaricaulota*, their length usually does not exceed 300–700 bp. Out of 366 near complete sequences (over 1200 bp long), most were found in deep sea and lake sediments, hot springs, oil reservoirs, and salt lake sediments. Organic-rich environment and anoxic conditions are common to all these habitats, which may favor development of fermenting organotrophs of the phylum *Bipolaricaulota*.

The microbial community of the subsurface aquifer in which Ch78 was found consist of methanogenic archaea (~50%) and various uncultured bacterial lineages phylogenetically distant from the known species and belonging mainly to the phyla *Firmicutes*, *Chloroflexi*, *Proteobacteria*, *Bacteroidetes*, and *Ignavibacteriae* (Kadnikov et al., 2017a). Members of these phyla are known to be able to hydrolyze complex polymers. Particularly, members of the phyla *Ignavibacteriae* and *Chloroflexi* (family *Anaerolineaceae*) found in subsurface waters in the Western Siberian region may be involved in the degradation of polysaccharides and proteins (Kadnikov et al., 2018). In such ecosystems, members of the *Bipolaricaulota* probably act as scavengers, consuming low-molecular organic compounds produced by hydrolytic microorganisms and producing hydrogen and acetate, the substrates for methanogenic archaea.

Description of “*Candidatus Bipolaricaulis sibiricus*”

Ch78. Based on analysis of its complete genome, a new species, “*Candidatus Bipolaricaulis sibiricus*,” is proposed. *Bipolaricaulis sibiricus* sp. nov. (si.bi’ri.cus N.L. masc. adj. sibiricus, originating from Siberia). Uncultured. The organism was detected in a deep subsurface thermal aquifer in Western Siberia. Probably anaerobic, capable of fermenting sugars and proteinaceous substrates. The DNA G + C content is 67.23 mol %. The organism is represented by complete genome (GenBank CP034928) obtained from the metagenome of subsurface thermal water from the 5P borehole (Chazhemto, Tomsk region, Russia).

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research, project no. 16-34-60124 and by the Ministry of Science and Higher Education of the Russian Federation.

The work was carried out using the scientific equipment of the Core Research Facility “Bioengineering” (Research Center of Biotechnology Russian Academy of Sciences).

FUNDING

The work was supported by the Russian Foundation for Basic Research, project no. 16-34-60124 and by the Ministry of Science and Higher Education of the Russian Federation.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. This article does not contain any studies involving animals or human participants performed by any of the authors.

REFERENCES

- Alneberg, J., Bjarnason, B.S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F., and Quince, C., Binning metagenomic contigs by coverage and composition, *Nat. Methods*, 2014, vol. 11, pp. 1144–1146.
- Anantharaman, K., Brown, C.T., Hug, L.A., Sharon, I., Castelle, C.J., Probst, A.J., Thomas, B.C., Singh, A., Wilkins, M.J., Karaoz, U., Brodie, E.L., Williams, K.H., Hubbard, S.S., and Banfield, J.F., Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system, *Nat. Commun.*, 2016, vol. 7, p. 13219.
- Brown, C.T., Olm, M.R., Thomas, B.C., and Banfield, J.F., Measurement of bacterial replication rates in microbial communities, *Nat. Biotechnol.*, 2016, vol. 34, pp. 1256–1263.
- Cao, M.D., Nguyen, S.H., Ganesamoorthy, D., Elliott, A.G., Cooper, M.A., and Coin, L.J., Scaffolding and completing genome assemblies in real-time with nanopore sequencing, *Nat. Commun.*, 2017, vol. 8, p. 14515.
- Costa, K.C., Navarro, J.B., Shock, E.L., Zhang, C.L., Soukup, D., and Hedlund, B.P., Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin, *Extremophiles*, 2009, vol. 13, pp. 447–459.
- Greening, C., Biswas, A., Carere, C.R., Jackson, C.J., Taylor, M.C., Stott, M.B., Cook, G.M., and Morales, S.E., Genomic and metagenomic surveys of hydrogenase distribution indicate H₂ is a widely utilised energy source for microbial growth and survival, *ISME J.*, 2016, vol. 10, pp. 761–777.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O., New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0, *Syst. Biol.*, 2010, vol. 59, pp. 307–321.
- Hao, L., McIlroy, S.J., Kirkegaard, R.H., Karst, S.M., Fernando, W.E.Y., Aslan, H., Meyer, R.L., Albertsen, M., Nielsen, P.H., and Dueholm, M.S., Novel prosthecate bacteria from the candidate phylum Acetothermia, *ISME J.*, 2018, vol. 12, pp. 2225–2237.
- Hernsdorf, A.W., Amano, Y., Miyakawa, K., Ise, K., Suzuki, Y., Anantharaman, K., Probst, A., Burstein, D., Thomas, B.C., and Banfield, J.F., Potential for microbial H₂ and metal transformations associated with novel bacteria and archaea in deep terrestrial subsurface sediments, *ISME J.*, 2017, vol. 11, pp. 1915–1929.
- Hirayama, H., Takai, K., Inagaki, F., Yamato, Y., Suzuki, M., Nealson, K.H., and Horikoshi, K., Bacterial community shift along a subsurface geothermal water stream in a Japanese gold mine, *Extremophiles*, 2005, vol. 9, pp. 169–184.
- Hu, P., Tom, L., Singh, A., Thomas, B.C., Baker, B.J., Piceno, Y.M., Andersen, G.L., and Banfield, J.F., Genome-resolved metagenomic analysis reveals roles for candidate phyla and other microbial community members in biogeochemical transformations in oil reservoirs, *MBio*, 2016, vol. 7, p. e01669-15.
- Hugenholtz, P., Pitulle, C., Hershberger, K.L., and Pace, N.R., Novel division level bacterial diversity in a Yellowstone hot spring, *J. Bacteriol.*, 1998, vol. 180, pp. 366–376.
- Kadnikov, V.V., Frank, Y.A., Mardanov, A.V., Beletskii, A.V., Ivasenko, D.A., Pimenov, N.V., Karnachuk, O.V., and Ravin, N.V., Uncultured bacteria and methanogenic archaea predominate in the microbial community of Western Siberian deep subsurface aquifer, *Microbiology (Moscow)*, 2017a, vol. 86, pp. 412–415.
- Kadnikov, V.V., Frank, Y.A., Mardanov, A.V., Beletsky, A.V., Ivasenko, D.A., Pimenov, N.V., Karnachuk, O.V., and Ravin, N.V., Variability of the composition of the microbial community of the deep subsurface thermal aquifer in Western Siberia, *Microbiology (Moscow)*, 2017b, vol. 86, pp. 765–772.
- Kadnikov, V.V., Frank, Y.A., Mardanov, A.V., Beletsky, A.V., Karnachuk, O.V., and Ravin, N.V., Metagenome of the Siberian underground water reservoir, *Genome Announc.*, 2017c, vol. 5, p. e01317-17.
- Kadnikov, V.V., Mardanov, A.V., Beletsky, A.V., Banks, D., Pimenov, N.V., Frank, Y.A., Karnachuk, O.V., and Ravin, N.V., A metagenomic window into the 2-km-deep terrestrial subsurface aquifer revealed multiple path-

- ways of organic matter decomposition, *FEMS Microbiol. Ecol.*, 2018, vol. 94, p. fyy152.
- Kato, S., Kobayashi, C., Kakegawa, T., and Yamagishi, A., Microbial communities in iron silica rich microbial mats at deep sea hydrothermal fields of the Southern Mariana Trough, *Environ. Microbiol.*, 2009, vol. 11, pp. 2094–2111.
- Konstantinidis, K.T., Rosselló-Móra, R., and Amann, R., Uncultivated microbes in need of their own taxonomy, *ISME J.*, 2017, vol. 11, pp. 2399–2406.
- Langmead, B. and Salzberg, S.L., Fast gapped-read alignment with Bowtie 2, *Nat. Methods*, 2012, vol. 9, pp. 357–359.
- Li, H. and Durbin, R., Fast and accurate short read alignment with Burrows-Wheeler transform, *Bioinformatics*, 2009, vol. 25, pp. 1754–1760.
- Magnabosco, C., Ryan, K., Lau, M.C., Kuloyo, O., Lollar, B.S., Kieft, T.L., van Heerden, E., and Onstott, T.C., A metagenomic window into carbon metabolism at 3 km depth in Precambrian continental crust, *ISME J.*, 2016, vol. 10, pp. 730–741.
- Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.-P., and Göker, M., Genome sequence-based species delimitation with confidence intervals and improved distance functions, *BMC Bioinform.*, 2013, vol. 14, p. 60.
- Nobu, M.K., Narihiro, T., Rinke, C., Kamagata, Y., Tringe, S.G., Woyke, T., and Liu, W.T., Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor, *ISME J.*, 2015, vol. 9, pp. 1710–1722.
- Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., and Hugenholtz, P., A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life, *Nat. Biotechnol.*, 2018, vol. 36, pp. 996–1004.
- Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W., CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes, *Genome Res.*, 2015, vol. 25, pp. 1043–1055.
- Probst, A.J., Castelle, C.J., Singh, A., Brown, C.T., Anantharaman, K., Sharon, I., Hug, L.A., Burstein, D., Emerson, J.B., Thomas, B.C., and Banfield, J.F., Genomic resolution of a cold subsurface aquifer community provides metabolic insights for novel microbes adapted to high CO₂ concentrations, *Environ. Microbiol.*, 2017, vol. 19, pp. 459–474.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O., The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, *Nucleic Acids Res.*, 2013, vol. 41, pp. D590–D596.
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A., Dodsworth, J.A., Hedlund, B.P., Tsiamis, G., Sievert, S.M., Liu, W.T., Eisen, J.A., et al., Insights into the phylogeny and coding potential of microbial dark matter, *Nature*, 2013, vol. 499, pp. 431–437.
- Rodriguez-R, L.M. and Konstantinidis, K.T., The envomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes, *PeerJ Preprints*, 2016, vol. 4, p. e1900v1.
- Søndergaard, D., Pedersen, C.N., and Greening, C., HydDB: A web tool for hydrogenase classification and analysis, *Sci. Rep.*, 2016, vol. 6, p. 34212.
- Stott, M.B., Saito, J.A., Crowe, M.A., Dunfield, P.F., Hou, S., Nakasone, E., Daughney, C.J., Smirnova, A.V., Mountain, B.W., Takai, K., and Alam, M., Culture-independent characterization of a novel microbial community at a hydrothermal vent at Brothers volcano, Kermadec arc, New Zealand, *J. Geophys. Res.*, 2008, vol. 113, p. B08S06.
- Takami, H., Noguchi, H., Takaki, Y., Uchiyama, I., Toyoda, A., Nishi, S., Chee, G.J., Arai, W., Nunoura, T., Itoh, T., Hattori, M., and Takai, K., A deeply branching thermophilic bacterium with an ancient acetyl-CoA pathway dominates a subsurface ecosystem, *PLoS One*, 2012, vol. 7, e30559.
- Teske, A., Hinrichs, K.U., Edgcomb, V., de Vera Gomez, A., Kysela, D., Sylva, S.P., Sogin, M.L., and Jannasch, H.W., Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities, *Appl. Environ. Microbiol.*, 2002, vol. 68, pp. 1994–2007.
- Tobler, D.J. and Benning, L.G., Bacterial diversity in five Icelandic geothermal waters: temperature and sinter growth rate effects, *Extremophiles*, 2011, vol. 15, pp. 473–485.
- Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M., Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement, *PLoS One*, 2014, vol. 9, p. e112963.

Translated by P. Sigalevich