
EXPERIMENTAL ARTICLES

Metagenomic Analysis of the Microbial Community in the Underground Coal Fire Area (Kemerovo Region, Russia) Revealed Predominance of Thermophilic Members of the Phyla *Deinococcus-Thermus*, *Aquificae*, and *Firmicutes*

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

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

^b Laboratory of Biochemistry and Molecular Biology, Tomsk State University, Tomsk, 634050 Russia

*e-mail: vkadnikov@bk.ru

Received May 12, 2021; revised May 13, 2021; accepted May 14, 2021

Abstract—Underground burning of coal seams accompanied by release of gases leads to development of local thermal ecosystems. We investigated the microbial community of the ground heated to 72°C in the release area of hot gases resulting from underground combustion of coal mining waste at the Bungurskiy-Severnoy coal deposit in the Kemerovo region of Russia. Analysis of the composition of the microbial community by 16S rRNA gene profiling revealed predominance of thermophilic bacteria of the phyla *Deinococcus-Thermus*, *Aquificae*, and *Firmicutes*. As a result of metagenomic analysis, 18 genomes of the main members of the microbial community were assembled, including the complete genomes of *Hydrogenobacter thermophiles*, a member of the candidate genus UBA11096 of the phylum *Aquificae* (RBS10-58), *Thermoflexus hugenholtzii*, and *Thermus antranikianii*. Analysis of the RBS10-58 genome indicates that this bacterium can autotrophically fix carbon in the reductive tricarboxylic acid cycle and obtain energy via oxidation of hydrogen and sulfur compounds with oxygen or nitrate as electron acceptors. Genome analysis of the two dominant *Firmicutes* species, *Hydrogenibacillus schlegelii* and an uncultured member of the class *Thermaerobacteria*, showed that these bacteria could grow aerobically by oxidizing hydrogen and carbon monoxide. Overall, the community was dominated by aerobic bacteria capable of growing autotrophically and obtaining energy via oxidation of the main components of coal gases, hydrogen and carbon monoxide. *Thermus antranikianii*, which makes up about a half of the microbial community, probably uses organic matter produced by autotrophic members of *Firmicutes* and *Aquificae*.

Keywords: thermophiles, coal gases, microbial community, hydrogenotrophs, *Thermus*, *Aquificae*

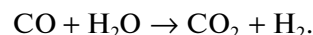
DOI: 10.1134/S0026261721050088

Studies of thermophilic microorganisms have expanded our knowledge of microbial diversity, evolution, and mechanisms of adaptation to extreme environmental conditions (Urbieta et al., 2015; Counts et al., 2017). Most research on thermophilic microorganisms has focused on thermal ecosystems associated with volcanic activity, such as terrestrial hot springs and deep-sea hydrothermal vents, or on technogenic biotopes (e.g., high-temperature bioreactors). Similarly to volcanic activity, natural combustion of fossil hydrocarbons and coal can also underlie formation of local thermal ecosystems. Characterization of microbial communities in such ecosystems contributes to our understanding of the diversity of thermophilic microorganisms and processes mediated by them.

Underground burning of coal seams is a widespread natural phenomenon and has been observed in Australia, Germany, the United States, China, Russia,

India, and other countries (Stracher and Taylor, 2004). Such underground fires can last for centuries; for example, the coal seam in Dudweiler (Saarland, Germany) has been burning since 1668. An example of long-term natural underground coal combustion is the Burning Mountain in Australia, which is estimated to have burned for about 6000 years (Rattigan, 1967).

Coal burning under conditions of oxygen deficit and in the presence of water leads to the formation of coal gases according to the following reactions (Shafirovich and Varma, 2009):



Underground coal combustion is a natural analog of the synthesis gas production by coal gasification. In addition to CO₂, coal gases contain mainly hydrogen, CO, and gaseous hydrocarbons (Stracher and Taylor,

2004; Engle et al., 2012). These gases can also contain hydrogen sulfide, sulfur oxides, and other toxic compounds, such as benzene, xylene, or aliphatic and halogenated compounds (Engle et al., 2012; Zhang et al., 2013). Sulfur and other elements present in coal seams can be transferred to the surface with the gas flow (Pone et al., 2007).

Local extreme ecosystems developing in areas where hot coal gases come to the surface are characterized by high temperatures ($>50^{\circ}\text{C}$) and the presence of toxic substances (Tammy et al., 2005). The high-energy compounds contained in coal gases, such as hydrogen and CO, can be used by microorganisms as substrates, while oxygen is used as an electron acceptor, which enables formation of specific communities of thermophilic microorganisms. However, relatively little is known about the composition of microbial communities of such ecosystems and the genetic potential of their microorganisms.

The phenomenon of underground coal combustion is also found in coal deposits in Russia. The Kuznetsk coal basin (Kuzbass) located in the south of Western Siberia is one of the largest coal mining areas in the world. Currently, coal in this region is mainly mined by open pit mining, which produces large quantities of waste, including overburden. Overburden containing residual coal is stored as dumps directly at the mining sites, where ignition and long-term burning can occur due to natural and anthropogenic causes. A well-known example of an environmental disaster is the burning of coal dumps near the mining town of Kiselevsk in Kuzbass (Russia) (Kadnikov et al., 2021).

Start new paragraph here we studied the soil microbial community associated with the zone of underground coal combustion and the release of hot coal gases at the dumps of the Bungurskiy-Severny coal deposit in the Novokuznetsk district of the Kemerovo region (Russia). The goal of the present work was to investigate the composition and the genetic potential of this microbial community. We present data on the composition of the community obtained by high-throughput sequencing of 16S rRNA gene amplicons and the soil metagenome in the area of hot coal gas release to the surface. As a result of metagenomic analysis, high-quality genomes (metagenome-assembled genomes, MAGs) of most members of the community were obtained, which made it possible to characterize the metabolic potential of the corresponding microorganisms.

MATERIALS AND METHODS

Sampling and DNA extraction. Samples were taken at the storage site of coal mining waste at the Bungurskiy-Severny coal deposit near the Apanas village of the Novokuznetsk district, Kemerovo region (53.542314 N, 86.862370 E). A ground sample col-

lected at a depth of 5–10 cm and designated RBS10 represented wet ground on the slope of a mining waste dump near the site of hot steam and gas release. The sample collected was a finely dispersed rock containing coal. Metagenomic DNA was isolated using a MO BIO Power Soil DNA Kit (MO BIO Laboratories, Qiagen, United States).

Sequencing and analysis of 16S rRNA gene fragments. PCR amplification of 16S rRNA gene fragments including the hypervariable regions V3–V6 was performed using the universal primers 341F (5'-CC-TAYGGGDBGCWSCAG-3') and 806R (5'-GGAC-TACNVGGGTHTCTAAT-3') (Frey et al., 2016). The obtained PCR fragments were barcoded using a Nextera XT Index Kit v.2 (Illumina, United States) and purified using Agencourt AMPure beads (Beckman Coulter, United States); the amount of DNA was determined using a Qubit dsDNA HS Assay Kit (Invitrogen, United States). Next, the amplicons were sequenced using Illumina MiSeq (paired-end reads, 2×300 bp). Overlapping reads were merged using FLASH v.1.2.11 (Magoč and Salzberg, 2011). Filtering by quality and clustering of sequences into operational taxonomic units (OTUs) at the level of 97% sequence identity was performed using the Usearch software (Edgar, 2010). Chimeric sequences and singletons were removed during clustering by the Usearch algorithm. To calculate the relative abundance of OTUs, all reads (including singletons and low-quality ones) were mapped using Usearch on OTU sequences with an identity threshold of 97%.

Taxonomic identification of OTUs was carried out by searching the SILVA v.132 rRNA sequence database using the VSEARCH algorithm (Rognes et al., 2016).

Sequencing of metagenomic DNA, assembly of contigs, and their clustering to obtain MAGs. Metagenomic DNA was sequenced using Illumina HiSeq2500 according to the manufacturer's instructions (Illumina, United States). Sequencing of the TruSeq DNA library (paired-end reads, 2×150 bp) yielded 232885794 pairs of reads. Removal of adapters and exclusion of low quality sequences ($Q < 30$) were performed using Cutadapt v.1.8.3 (Martin, 2011) and Sickle v.1.33 (<https://github.com/najoshi/sickle>), respectively. The processed paired-end reads were merged using FLASH v.1.2.11 (Magoč and Salzberg, 2011).

Metagenomic DNA was additionally sequenced on a MinION (Oxford Nanopore, United Kingdom) using a 1D Genomic DNA by Ligation kit (SQK-LSK108). Sequencing this library using a MinION device with an R9.4 flowcell (FLO-MIN106) yielded 8280228 reads with a total length of 16.12 billion bp.

All Illumina (about 25 billion bp in total) and Nanopore reads obtained were *de novo* assembled into contigs using the metaSPAdes hybrid assembler v.3.13.0 software (Nurk et al., 2017). Contigs longer than 1500 bp were binned into clusters representing

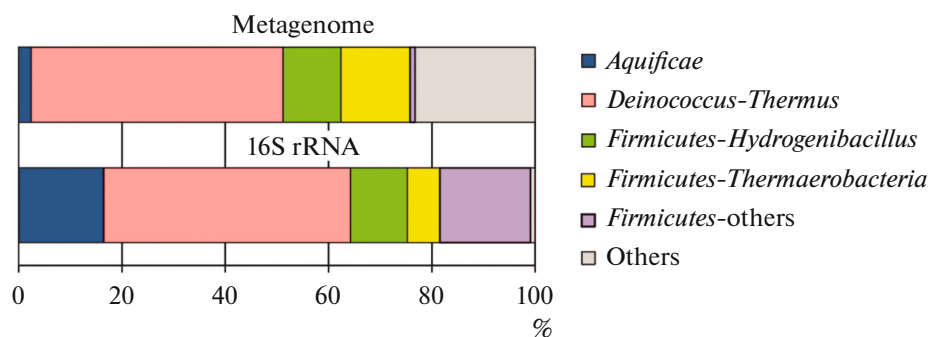


Fig. 1. Composition of the microbial community based on the analysis of 16S rRNA genes and metagenomic sequencing. The category Others in the case of the metagenome comprises sequences not included in the MAGs assigned to the indicated groups.

MAGs using MetaBAT v.2.12.1 (Kang et al., 2015). To improve MAG assembly, MinION reads were mapped to the contigs included in MAG using the BWA v.0.7.15 software (Li and Durbin, 2010). Next, Npscarf v.1.0 (Cao et al., 2017) was used to form chains of contigs (scaffolds) and to fill the gaps between contigs using Illumina consensus sequences from the metaSPAdes assembly graph.

In addition, MinION reads were *de novo* assembled into contigs using Flye v.2.7 (Kolmogorov et al., 2019). Contig sequences were corrected with Pilon v.1.2.2 (Walker et al., 2014) as a result of two iterations of mapping the Illumina reads onto the assembled contig sequences using the Bowtie 2 software (Langmead and Salzberg, 2012). The resulting contigs were binned into MAGs using MetaBAT v.2.12.1 (Kang et al., 2015).

MAG annotation and analysis. The completeness of MAGs and their possible contamination (i.e., possible presence of contigs representing other genomes due to incorrect binning) were assessed using CheckM v.1.05 (Parks et al., 2015). The assembled MAGs were taxonomically classified using the Genome Taxonomy Database Toolkit (GTDB-Tk) v.0.3.2 (Chaumeil et al., 2020) and the Genome Taxonomy Database (GTDB) (Parks et al., 2018).

Gene search and MAG annotation were performed using the NCBI Prokaryotic Genome Annotation Pipeline (Tatusova et al., 2016) or the RAST server 2.0 (Brettin et al., 2015) with subsequent correction of the annotation by comparing the predicted protein sequences to the databases of the National Center for Biotechnology Information (NCBI). N-terminal signal peptides were predicted using Signal P v.5.0, and the presence of transmembrane domains was predicted using the TMHMM v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>).

Evaluation of genome similarity and phylogenetic analysis based on complete genome data. The levels of average nucleotide identity (ANI) and average amino acid identity (AAI) between selected genomes were

calculated using scripts from the Enveomics Collection toolbox (Rodriguez-R and Konstantinidis, 2016).

GTDB-Tk v.0.3.2 software was used to search for single-copy marker genes in a given MAG and to construct multiple alignments of concatenated single-copy marker gene sequences from this MAG and all species from GTDB. A part of the multiple alignment created using the GTDB-Tk software was used to build a phylogenetic tree using PhyML v.3.3 (Guindon et al., 2010) with default parameters. Internal branching support was assessed using a Bayesian test in the PhyML.

Deposition of nucleotide sequences. The primary data obtained by sequencing 16S rRNA gene fragments and by metagenome sequencing were deposited in the NCBI Sequence Read Archive (SRA) under accession numbers SRX10881305, SRX10881306 and SRX10881307. The annotated MAG sequences were deposited in the GenBank database and are available via BioProject PRJNA728906.

RESULTS

Composition of the microbial community based on analysis of 16S rRNA gene amplicons. A ground sample was collected from a burning coal mining dump with numerous vents releasing combustion products to the surface. The ground temperature at the sampling site was 72°C.

The composition of the microbial community was characterized based on 57 213 sequences of 16S rRNA gene fragments. As a result of clustering these sequences, 29 OTUs were identified with an identity threshold of 97%. All identified OTUs belonged to bacteria; no archaea were detected. The results of the taxonomic classification of the OTUs are shown in Fig. 1.

The community was dominated by members of three phyla: *Deinococcus-Thermus* (47.7% of the total number of 16S rRNA gene sequences), *Firmicutes* (34.9%), and *Aquificae* (16.5%). Members of the phyla *Bacteroidetes* (0.48%), *Proteobacteria* (0.27%), *Chlo-*

roflexi (0.02%), and *Actinobacteria* (0.01%) were found in minor amounts.

The phylum *Deinococcus-Thermus* was represented by two OTUs, one of which dominated the community, accounting for 47.4% of the total number of 16S rRNA gene sequences. This OTU belonging to the genus *Thermus* had 98.7% sequence identity to the thermophilic heterotrophic bacterium *Thermus antranikianii* isolated from hot springs in Iceland (Chung et al., 2000).

The phylum *Firmicutes* was represented by 12 OTUs. The most abundant of them (14.1%) belonged to the genus *Brevibacillus* and had 98.73% 16S rRNA gene sequence identity with *Brevibacillus borstelensis*, a moderately thermophilic heterotrophic spore-forming bacterium found in soils and hydrothermal habitats (Shida et al., 1995; Khalil et al., 2018). Approximately 11% of the sequences represented one OTU belonging to the genus *Hydrogenibacillus* (order *Thermicanales*, family *Thermicanaceae* according to the Genome Taxonomy Database). A cultured representative of this genus, *Hydrogenibacillus schlegelii*, is a facultative chemolithotrophic aerobic thermophilic bacterium capable of oxidizing hydrogen and, presumably, CO (Schenk and Aragno, 1979; Kämpfer et al., 2013). About 6.3% of the community belonged to an OTU phylogenetically distant from the cultured members of the phylum *Firmicutes* (<87% 16S rRNA gene sequence identity).

The phylum *Aquificae* was represented by one OTU phylogenetically close to the genus *Hydrogenobacter* of the family *Aquificaceae*. Members of *Hydrogenobacter*, typical inhabitants of high-temperature thermal springs, are chemolithoautotrophic thermophiles capable of aerobic hydrogen oxidation (Reysenbach et al., 2000; Takacs-Verbach et al., 2013).

Metagenome sequencing and MAGs assembly. To obtain the genomes of the microbial community members, we sequenced the metagenome of RBS10 sample using a combination of Illumina and Oxford Nanopore technologies. The assembled contigs were binned into 18 MAGs with completeness of more than 80% and contamination (redundancy) of less than 10% as assessed by CheckM based on the analysis of a set of conserved single-copy marker genes (Table 1). Altogether, these MAGs represented approximately 80% of the total metagenome of the community. The taxonomic affiliation of the obtained MAGs was determined based on phylogenetic analysis of concatenated sequences of conserved marker genes according to the Genome Taxonomy Database (Parks et al., 2018).

The taxonomic classification of the MAGs identified the same main bacterial phyla that were found using 16S rRNA profiling. The relative abundances of some groups differed between the pool of 16S rRNA gene sequences and the complete metagenome (Fig. 1), probably due to the different number of 16S

rRNA gene copies in the genomes and different sizes of the genomes. Nearly a half of the total metagenome was represented by a single MAG, RBS10-92, assigned to the phylum *Deinococcus-Thermus*. About 26% of the metagenome were represented by ten MAGs of the phylum *Firmicutes*; three MAGs were attributed to *Aquificae* (2.4% of the metagenome); three belonged to *Chloroflexi* (3.3% of the metagenome); two were assigned to *Proteobacteria* (0.5% of the metagenome), and 0.02% of the metagenome was the MAG of the candidate phylum WOR-3 (Table 1).

Using the long reads obtained by nanopore sequencing, the complete circular sequences of four genomes were assembled: *Hydrogenobacter thermophilus* (MAG RBS10-74), candidate genus UBA11096 of the phylum *Aquificae* (MAG RBS10-58), *Thermoflexus hugenholtzii* of the phylum *Chloroflexi* (MAG RBS10-4), and *Thermus antranikianii* (MAG RBS10-92).

To characterize the metabolic capabilities of the dominant members of the microbial community, we performed a detailed analysis of several MAGs.

Genomes of members of the phylum *Firmicutes*. Two bacteria of the phylum *Firmicutes*, represented by MAG RBS10-35 and MAG RBS10-49, were abundant members of the microbial community. The RBS10-35 genome represented 11.0% of the total metagenome and was identified only at the class level as belonging to *Thermaerobacteria* according to the genomic taxonomy system. Probably, this genome corresponded to the OTU that constituted 6.3% of the 16S rRNA sequences and was assigned to unclassified *Firmicutes*. A GenBank search for closest relatives of RBS10-35 revealed only one 16S rRNA gene sequence (FN687452) with 91% identity found in a thermal aerobic bioreactor for processing activated sludge waste (Hayes et al., 2011). Analysis of the RBS10-35 genome showed that this bacterium had a complete aerobic respiratory chain and CO dehydrogenase, which indicated the possibility of CO oxidation.

The organism represented by the genome RBS10-49 was the second most abundant among *Firmicutes* (10.7% of the metagenome). This genotype was identified as *Hydrogenibacillus schlegelii* based on 97.68% AAI with *H. schlegelii* strain MA48 (Maker et al., 2017). Analysis of the RBS10-49 genome revealed the presence of the genes for the aerobic respiratory chain, membrane-bound uptake [NiFe]-hydrogenase of group 1d, and respiratory CO dehydrogenase, which indicates the ability of this bacterium to obtain energy by oxidation of coal gas components, hydrogen and CO. The genome also encoded a complete Calvin cycle, which can enable autotrophic carbon fixation.

It should be noted that among the assembled MAGs, there was no genome belonging to the genus *Brevibacillus*, although its members accounted for about 15% of the 16S rRNA gene sequences. Probably,

Table 1. Principal characteristics of MAGs

MAG ID	Completeness/ contamination, %	Genome size, bp	Number of contigs	GC, %	Share in the metagenome, %	Taxonomy*
2	97.63/1.36	1699170	143	41.6	0.08	p__Aquificota; g__Hydrogenobacter
74	99.59/0.51	1794458	1**	44	0.17	p__Aquificota; s__Hydrogenobacter thermophilus
58	99.59/0.41	1722082	1**	43.1	2.13	p__Aquificota; g__UBA11096
4	97.27/0.91	3316678	1**	67.6	3.17	p__Chloroflexota; s__Thermoflexus hugenoltzii
7	85.55/1.87	2784472	202	65.1	0.03	p__Chloroflexota; g__Thermomicrobium
30	96.23/0	3717411	433	65.6	0.11	p__Chloroflexota; f__UBA6265
92	100/0	2424424	1**	64.9	48.83	p__Deinococcota; s__Thermus antranikianii
77	93.16/1.82	2294372	228	45	0.16	p__Firmicutes; f__Amphibacillaceae
39	98.09/0.68	2502348	201	37.5	0.12	p__Firmicutes_A; s__Caldanaerobacter subterraneus
35	94.06/0.2	2368455	42	63.5	11.01	p__Firmicutes_E; c__Thermaerobacteria
62	92.57/13.9	4585313	204	68.9	2.34	p__Firmicutes_E; g__Thermaerobacter
82	92.25/2.97	3307331	619	58.5	0.09	p__Firmicutes_G; o__DTU080
36	96.15/2.95	4121736	396	53.4	0.25	p__Firmicutes_I; s__Bacillus_BB thermozeamaize
40	84.97/4.92	2109662	127	64	0.37	p__Firmicutes_I; s__Brockia lithotrophica
8	91.28/1.72	2288228	67	61.8	0.48	p__Firmicutes_I; g__Hydrogenibacillus
49	92.44/8.08	2964420	72	66.1	10.75	p__Firmicutes_I; s__Hydrogenibacillus schlegelii
91	99.19/2.07	4672685	147	66.4	0.18	p__Proteobacteria; g__Paracoccus
48	99.38/1.91	3146152	24	70.2	0.29	p__Proteobacteria; g__Lysobacter
29	92.37/1.69	1580388	114	27.6	0.02	p__WOR-3_A; o__LBFQ01

* Taxonomic affiliation according to the Genome Taxonomy Database, version R89. The phylum and the lowest-rank taxon up to which classification was possible are indicated. ** Complete circular genome.

Brevibacillus were represented by several closely related phylotypes, which made it difficult to assemble long contigs and, accordingly, to obtain their MAGs.

Complete genome of *Thermus antranikianii* RBS10-92. As a result of metagenome sequencing, the complete genome of the bacterium RBS10-92 was assembled. This genome had 97.79% AAI with *Thermus antranikianii* DSM 12462 (GCF_000423905),

and therefore was assigned to this species. The genome of *T. antranikianii* RBS10-92 was sequenced with 5030-fold average coverage and was 2424424 bp long. It is the first known complete genome of *T. antranikianii*. The relative abundance of RBS10-92 in the metagenome was 48.8%, which was in good agreement with the 47.64% share of the corresponding OTU among the 16S rRNA gene reads. Annotation of the

RBS10-92 genome identified 2636 potential protein-coding genes. The functions were predicted for only a half of the genes revealed. Two copies of the rRNA operon and 50 transfer RNA (tRNA) genes were identified in the genome.

Analysis of the genome of *T. antranikianii* RBS10-92 showed that this bacterium is probably an aerobic heterotroph capable of hydrolyzing various carbohydrates. This microorganism also has the ability to grow under anaerobic conditions using nitrate as an electron acceptor. These predictions are consistent with the microbiological characteristics of *T. antranikianii* isolates (Chung et al., 2000).

Complete genome of *Hydrogenobacter thermophilus* RBS10-74. Three genomes of representatives of the phylum *Aquificae* were assembled with completeness of more than 97% and contamination of less than 2%. Complete circular genomic sequences were obtained for two of these organisms. RBS10-74, which accounts for 0.17% of the total metagenome, was classified as *Hydrogenobacter thermophilus* based on 95.7% ANI with *H. thermophilus* TK-6 (GCA_000010785) isolated from a hot spring in Japan (Arai et al., 2010).

Analysis of the genome of *H. thermophilus* RBS10-74 predicted that, like cultured strains of *H. thermophilus*, it is an obligate autotrophic organism capable of oxidizing hydrogen and fixing CO₂ through the reverse tricarboxylic acid cycle. When growing under anaerobic conditions, this bacterium can use nitrate as an electron acceptor. RBS10-74 can also oxidize elemental sulfur or thiosulfate.

Complete genome of MAG RBS10-58, a member of the candidate genus UBA11096. The second complete genome of a member of the phylum *Aquificae* was assigned to the genus UBA11096 of the family *Aquificaceae* according to the genomic taxonomy system. To date, no cultured members of this genus have been isolated; for this reason, its definition is based on several draft genomes assembled from metagenomes that have not been analyzed previously.

The RBS10-58 genome was sequenced with 309-fold average coverage and assembled into a 1 722 082-bp-long circular chromosome. The relative abundance of this genotype in the metagenome was 2.13%. Annotation of the RBS10-58 genome identified 1848 potential protein-coding genes, as well as one 16S–23S–5S rRNA operon and 42 tRNA genes; gene functions were predicted for a half of the protein-coding genes revealed.

To determine the phylogenetic position of the bacterium RBS10-58, a phylogenetic tree was constructed based on concatenated amino acid sequences of conserved marker genes belonging to members of all genera of the family *Aquificaceae* determined in the genomic taxonomy system. It was shown that RBS10-58, together with several other MAGs, represented a separate genus-level lineage, along with the genera *Aquifex*, *Hydrogenivirga*, *Hydrogenobacter*, and

Thermocrinis, as well as *Thermocrinis minervae*, which represents a separate genus *Thermocrinis_A* in the genomic taxonomy (Fig. 2).

Analysis of the RBS10-58 genome revealed a complete set of genes for the reverse tricarboxylic acid cycle, which is used in *Aquificae* for autotrophic CO₂ fixation (Hügler et al., 2007), as well as genes of the gluconeogenesis pathway. The pentose phosphate pathway was absent in RBS10-58. The genome contained a complete set of genes for oxidative phosphorylation, including NADH dehydrogenase, succinate dehydrogenase, cytochrome *bc1* complex, and several cytochrome *c* oxidases. The presence of membrane-bound uptake [NiFe]-hydrogenases of the groups 2a and 1d indicates the possibility of using molecular hydrogen as an electron donor during respiration, which is typical for members of the phylum *Aquificae*. Formate can also serve as energy source, as evidenced by the presence of membrane-bound formate dehydrogenase.

Under anaerobic conditions, nitrate can serve as an electron acceptor, as indicated by the presence of membrane-bound nitrate reductase, which reduces nitrate to nitrite. Further, nitrite can be reduced by cytochrome *cd1* nitrite reductase to form nitric oxide(II). Subsequent denitrification steps can be carried out by nitric oxide reductase and nitrous oxide reductase to form molecular nitrogen.

The bacterium RBS10-58 can also use sulfur compounds as an electron donor. Oxidation of hydrogen sulfide can be carried out by sulfide:quinone oxidoreductase and flavocytochrome *c*:sulfide dehydrogenase. The genome also encoded a Sox-Hdr-Soe variant of the pathway for the oxidation of sulfur compounds to sulfate (Watanabe et al., 2019), including a cluster of *sox* genes *soxYZAXB*, sulfite dehydrogenase genes *soeABC*, and a gene cluster encoding subunits of heterosulfide reductase *hdrCBAhyphdrCB*.

Thus, the bacterium RBS10-58 is a chemolithoautotroph capable of obtaining energy by oxidation of molecular hydrogen and sulfur compounds through aerobic respiration, while under anaerobic conditions it can perform all stages of denitrification.

DISCUSSION

Despite the fact that one of the first described thermophiles, *Thermoplasma acidophilum*, was isolated from coal waste samples by Thomas Brock's group in the late 1960s (Darling et al., 1970), the composition of the microbial communities of these ecosystems remains poorly understood. The composition of soil microbial communities in areas of underground coal burning was studied in Pennsylvania (United States) (Tobin-Janzen et al., 2005; Lee et al., 2017). Analysis of microbial communities by 16S rRNA gene sequencing revealed the presence of archaea of the phylum *Crenarchaeota*, as well as bacteria of the phyla *Chloro-*

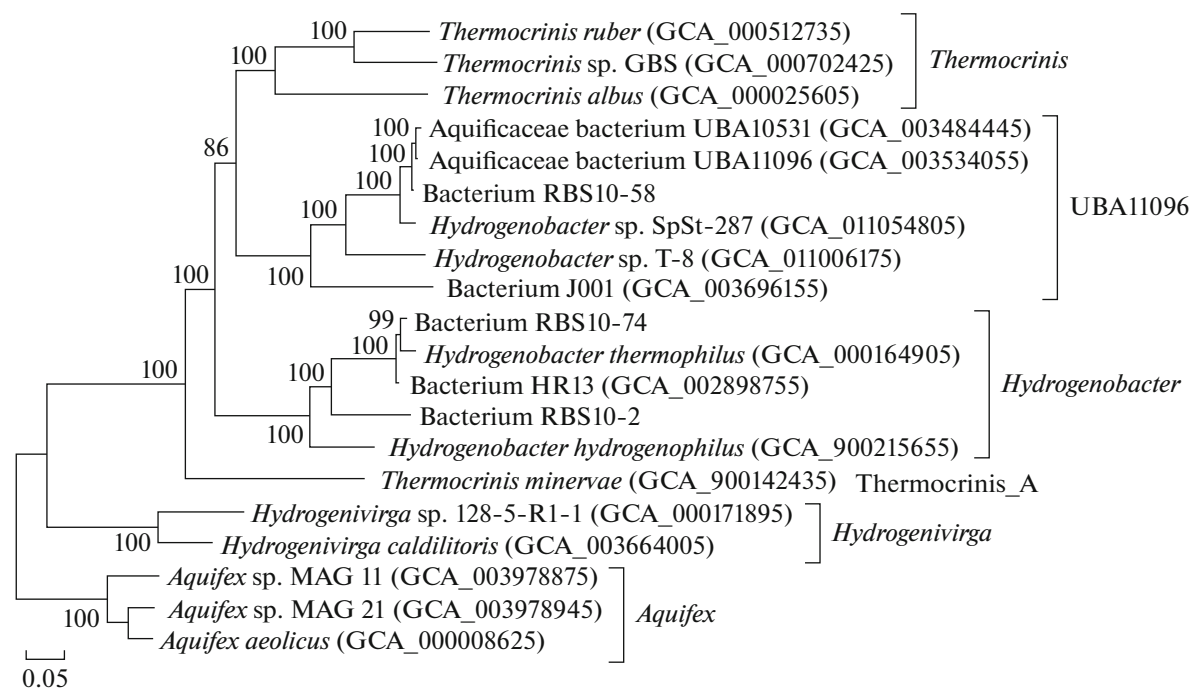


Fig. 2. Phylogenetic position of RBS10-58 in the family *Aquificaceae*. The maximum likelihood phylogenetic tree was based on concatenated sequences of conserved marker genes. The branching support levels were determined using a Bayesian test in the PhyML software. The taxonomy is shown according to the GTDB system.

flexi, *Acidobacteria*, *Proteobacteria*, *Bacteroidetes*, *Elusimicrobia*, and *Gemmatimonadetes*. Further metagenomic studies showed that microorganisms with smaller cell and genome sizes predominate in areas of soil with higher temperatures (Sorensen et al., 2019).

Microbial communities of heated soil near hot coal gas release sites in Xinjiang (China) were investigated using T-RFLP analysis and 16S rRNA gene clone libraries (Zhang et al., 2013). Predominant groups of microorganisms included members of the phyla *Firmicutes*, *Proteobacteria*, *Acidobacteria*, *Bacteroidetes*, *Planctomycetes*, and *Actinobacteria*. The most abundant group was *Firmicutes*, mainly represented by the genera *Bacillus* and *Paenibacillus*.

In both these cases, the objects of study were “ordinary,” organic-rich soils heated due to release of hot gases, which led to predominance of soil groups of microorganisms in the communities. In the present work, the object of study was not soils, but dumps of open-pit coal mining composed of coal-bearing rocks. It is assumed that coal gases can sustain development of specific communities of thermophiles in such ecosystems. The closest analog of the object studied in the present work was the microbial community of heated rocks in the area of underground coal burning in the Altai Mountains (Russia) (Kadnikov et al., 2018). It had a simple composition and included only three dominant phylotypes, all of which represented the phylum *Firmicutes*. It was the aerobic heterotroph

Ca. Carbobacillus altaicus, the anaerobic chemolithoautotroph *Brockia lithotrophica*, and the aerobic bacterium *Hydrogenibacillus schlegelii* capable of both using organic compounds and growing autotrophically. All these microorganisms can obtain energy by oxidation of molecular hydrogen, and some of them can also oxidize CO (Kadnikov et al., 2018). This community was also found to contain uncultured lineages of *Firmicutes* related to *Thermaerobacteria*.

Another close analog was the burning dumps of a coal mine near the city of Kiselevsk, Kemerovo region (Russia). The microbial community of the surface ground layer heated to 58°C was dominated by members of *Ktedonobacteria* (phylum *Chloroflexi*) capable of oxidizing hydrogen and CO, while thermophilic hydrogenotrophic *Firmicutes* constituted a minor share of the community (Kadnikov et al., 2021). The differences between this community and the object studied in the present work were probably due to differences in temperature and humidity.

In the RBS10 community studied in the present work, *Firmicutes* accounted for about one third of the community and were represented by thermophilic groups. *H. schlegelii* and uncultured *Thermaerobacteria* were among the dominant groups, while *Brockia lithotrophica* was found in small numbers. However, along with *Firmicutes*, the dominant groups in the RBS10 community were members of the genus *Thermus* and the family *Aquificaceae*, which are character-

istic of hydrothermal ecosystems (Counts et al., 2017; Bonch-Osmolovskaya, 2020). Previously, their presence in soils of underground coal burning areas has not been reported. *Aquificaceae* are capable of autotrophic carbon fixation and energy production through oxidation of hydrogen and sulfur compounds; along with *Firmicutes*, they represent the autotrophic component of the community sustained by coal gases. In turn, the organic substances produced by autotrophic *Firmicutes* and *Aquificaceae* are used as substrates by *T. antranikianii*, which constitutes about a half of the microbial community.

The thermal ecosystem associated with burning coal dumps that was studied in this work, as well as other similar objects in the Kemerovo region and Altai Mountains, formed no more than a few decades ago; they are young in comparison to geothermal objects. Previously, it was suggested that thermophilic *Firmicutes* with spores that can spread over long distances (Bonjour et al., 1988; Aullo et al., 2013) may be the first colonizers of such novel thermal ecological niches (Kadnikov et al., 2018). The discovery of members of *Thermus* and *Aquificaceae* indicates that the spread of non-spore-forming thermophiles, the source of which may be hot springs and other geothermal objects found in southern Siberia, can also be fast.

FUNDING

This work was partially supported by the Russian Science Foundation (project no. 19-74-00142).

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.

AUTHOR CONTRIBUTION

Sampling and DNA isolation were carried out by V.V. Kadnikov. Sequencing of 16S rRNA genes and metagenomic DNA was carried out by V.V. Kadnikov and A.V. Mardanov. Bioinformatic analysis of sequencing results was carried out by A.V. Beletsky. Data analysis and preparation of the article were carried out by V.V. Kadnikov, O.V. Karnachuk and N.V. Ravin. All authors participated in the discussion of the results.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the

article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

- Arai, H., Kanbe, H., Ishii, M., and Igarashi, Y., Complete genome sequence of the thermophilic, obligately chemolithoautotrophic hydrogen-oxidizing bacterium *Hydrogenobacter thermophilus* TK-6, *J. Bacteriol.*, 2010, vol. 192, pp. 2651–2652.
- Aullo, T., Ranchou-Peyruse, A., Ollivier, B., and Magot, M., *Desulfotomaculum* spp. and related gram-positive sulfate-reducing bacteria in deep subsurface environments, *Front. Microbiol.*, 2013, vol. 4, p. 362.
- Bonch-Osmolovskaya, E., *Aquificales*, *eLS*, 2020, vol. 1, pp. 433–441.
- Bonjour, F., Graber, A., and Aragno, M., Isolation of *Bacillus schlegelii*, a thermophilic, hydrogen oxidizing, aerobic autotroph, from geothermal and nongeothermal environments, *Microb. Ecol.*, 1988, vol. 16, pp. 331–337.
- Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A. III, Stevens, R., Vonstein, V., Wattam, A.R., and Xia, F., RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes, *Sci. Rep.*, 2015, vol. 5, art. 8365. <https://doi.org/10.1038/srep08365>
- 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, pp. 1–10.
- Chaumeil, P.A., Mussig, A.J., Hugenholtz, P., and Parks, D.H., GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database, *Bioinformatics*, 2020, vol. 36, pp. 1925–1927.
- Chung, A.P., Rainey, F.A., Valente, M., Nobre, M.F., and da Costa, M.S., *Thermus igniterrae* sp. nov. and *Thermus antranikianii* sp. nov., two new species from Iceland, *Int. J. Syst. Evol. Microbiol.*, 2000, vol. 50, pp. 209–217.
- Counts, J.A., Zeldes, B.M., Lee, L.L., Straub, C.T., Adams, M.W.W., and Kelly, R.M., Physiological, metabolic and biotechnological features of extremely thermophilic microorganisms, *Wiley Interdiscip. Rev. Syst. Biol. Med.*, 2017, vol. 9, e1377.
- Darling, G., Brock, T.D., Samsonoff, W., and Conti, S.F., A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile, *Science*, 1970, vol. 170, pp. 1416–1418.
- Edgar, R.C., Search and clustering orders of magnitude faster than BLAST, *Bioinformatics*, 2010, vol. 26, pp. 2460–2461.

- Engle, M.A., Radke, L.F., Heffern, E.L., O'Keefe, J.M., Hower, J.C., Smeltzer, C.D., Hower, J.M., Olea, R.A., Eatwell, R.J., Blake, D.R., Emsbo-Mattingly, S.D., Stout, S.A., Queen, G., Aggen, K.L., Kolker, A., et al., Gas emissions, minerals, and tars associated with three coal fires, Powder River Basin, USA, *Sci. Total Environ.*, 2012, vol. 420, pp. 146–159.
- Frey, B., Rime, T., Phillips, M., Stierli, B., Hajdas, I., Widmer, F., and Hartmann, M., Microbial diversity in European alpine permafrost and active layers, *FEMS Microbiol. Ecol.*, 2016, vol. 92, fiw018.
- 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.
- Hayes, D., Izzard, L., and Seviour, R., Microbial ecology of autothermal thermophilic aerobic digester (ATAD) systems for treating waste activated sludge, *Syst. Appl. Microbiol.*, 2011, vol. 34, pp. 127–138.
- Hügler, M., Huber, H., Molyneaux, S.J., Vetriani, C., and Sievert, S.M., Autotrophic CO₂ fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum *Aquificae*: evidence for two ways of citrate cleavage, *Environ. Microbiol.*, 2007, vol. 9, pp. 81–92.
- Kadnikov, V.V., Mardanov, A.V., Beletsky, A.V., Grigoriev, M.A., Karnachuk, O.V., and Ravin, N.V., Thermophilic *Chloroflexi* dominate in the microbial community associated with coal-fire gas vents in the Kuznetsk Coal Basin, Russia, *Microorganisms*, 2021, vol. 9, art. 948.
- Kadnikov, V.V., Mardanov, A.V., Ivasenko, D.A., Antsiferov, D.V., Beletsky, A.V., Karnachuk, O.V., and Ravin, N.V., Lignite coal burning seam in the remote Altai Mountains harbors a hydrogen-driven thermophilic microbial community, *Sci. Rep.*, 2018, vol. 8, pp. 1–12.
- Kämpfer, P., Glaeser, S.P., and Busse, H.J., Transfer of *Bacillus schlegelii* to a novel genus and proposal of *Hydrogenibacillus schlegelii* gen. nov., comb. nov., *Int. J. Syst. Bacteriol.*, 2013, vol. 63, pp. 1723–1727.
- Kang, D.D., Froula, J., Egan, R., and Wang, Z., MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities, *PeerJ*, 2015, vol. 3, e1165.
- Khalil, A.B., Sivakumar, N., Arslan, M., Saleem, H., and Qarawi, S., Insights into *Brevibacillus borstelensis* AK1 through whole genome sequencing: a thermophilic bacterium isolated from a hot spring in Saudi Arabia, *Biomed. Res. Int.*, 2018, art. 5862437.
- Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A., Assembly of long, error-prone reads using repeat graphs, *Nat. Biotechnol.*, 2019, vol. 37, pp. 540–546.
- Langmead, B. and Salzberg, S.L., Fast gapped-read alignment with Bowtie 2, *Nat. Methods*, 2012, vol. 9, pp. 357–359.
<https://doi.org/10.1038/nmeth.1923>
- Lee, S.H., Sorensen, J.W., Grady, K.L., Tobin, T.C., and Shade, A., Divergent extremes but convergent recovery of bacterial and archaeal soil communities to an ongoing subterranean coal mine fire, *ISME J.*, 2017, vol. 11, pp. 1447–1459.
- Li, H. and Durbin, R., Fast and accurate long-read alignment with Burrows–Wheeler transform, *Bioinformatics*, 2010, vol. 26, pp. 589–595.
- Magoč, T. and Salzberg, S.L., FLASH: fast length adjustment of short reads to improve genome assemblies, *Bioinformatics*, 2011, vol. 27, pp. 2957–2963.
- Maker, A., Hemp, J., Pace, L.A., Ward, L.M., and Fischer, W.W., Draft genome sequence of *Hydrogenibacillus schlegelii* MA48, a deep-branching member of the *Bacilli* class of *Firmicutes*, *Genome Announc.*, 2017, vol. 5, e00380-16.
- Martin, M., Cutadapt removes adapter sequences from high-throughput sequencing reads, *EMBnet. J.*, 2011, vol. 17, pp. 10–12.
- Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A., metaSPAdes: a new versatile metagenomic assembler, *Genome Res.*, 2017, vol. 27, pp. 824–834.
- Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarszewski, 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.
- Pone, J.D.N., Hein, K.A., Stracher, G.B., Annegarn, H.J., Finkleman, R.B., Blake, D.R., McCormack, J.K., and Schroeder, P., The spontaneous combustion of coal and its by-products in the Witbank and Sasolburg coalfields of South Africa, *Int. J. Coal Geol.*, 2007, vol. 72, pp. 124–140.
- Rattigan, J.H., Phenomena about Burning Mountain, Wingen, NSW, *Aust. J. Sci.*, 1967, vol. 30, pp. 183–184.
- Reysenbach, A.L., Ehringer, M., and Hershberger, K., Microbial diversity at 83°C in Calcite Springs, Yellowstone National Park: another environment where the *Aquificales* and “Korarchaeota” coexist, *Extremophiles*, 2000, vol. 4, pp. 61–67.
- Rodriguez-R, L.M. and Konstantinidis, K.T., The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes, *PeerJ Preprints*, 2016, e1900v1.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F., VSEARCH: A versatile open source tool for metagenomics, *PeerJ*, 2016, vol. 4, e2584.
- Schenk, A. and Aragno, M., *Bacillus schlegelii*, a new species of thermophilic, facultatively chemolithoautotrophic bacterium oxidizing molecular hydrogen, *Microbiology (SGM)*, 1979, vol. 115, pp. 333–341.
- Shafirovich, E. and Varma, A., Underground coal gasification: a brief review of current status, *Ind. Eng. Res.*, 2009, vol. 48, pp. 7865–7875.
- Shida, O., Takagi, H., Kadowaki, K., Udaka, S., Nakamura, L.K., and Komagata, K., Proposal of *Bacillus reuszeri* sp. nov., *Bacillus formosus* sp. nov., nom. rev., and *Bacillus borstelensis* sp. nov., nom. rev., *Int. J. Syst. Bacteriol.*, 1995, vol. 45, pp. 93–100.

- Sorensen, J.W., Dunivin, T.K., Tobin, T.C., and Shade, A., Ecological selection for small microbial genomes along a temperate-to-thermal soil gradient, *Nat. Microbiol.*, 2019, vol. 4, pp. 55–61.
- Stracher, G.B. and Taylor, T.P., Coal fires burning out of control around the world: Thermodynamic recipe for environmental catastrophe, *Int. J. Coal Geol.*, 2004, vol. 59, pp. 7–17.
- Takacs-Verbach, C., Inskeep, W.P., Jay, Z.J., Herrgard, M.J., Rusch, D.B., Tringe, S.G., Kozubal, M.A., Hamamura, N., Macur, R.E., Fouke, B.W., Reysenbach, A.L., McDermott, T.R., Jennings, R.deM., Hengartner, N.W., and Xie, G., Metagenome sequence analysis of filamentous microbial communities obtained from geochemically distinct geothermal channels reveals specialization of three *Aquificales* lineages, *Front. Microbiol.*, 2013, vol. 4, p. 84.
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M., and Ostell, J., NCBI prokaryotic genome annotation pipeline, *Nucleic Acids Res.*, 2016, vol. 44, pp. 6614–6624.
- Tobin-Janzen, T., Shade, A., Marshall, L., Torres, K., Beblo, C., Janzen, C., Lenig, J., Martinez, A., and Ressler, D., Nitrogen changes and domain bacteria ribo-
type diversity in soils overlying the Centralia, Pennsylvania underground coal mine fire, *Soil Sci.*, 2005, vol. 170, pp. 191–201.
- Urbiet, M.S., Donati, E.R., Chan, K.G., Shahar, S., Sin, L.L., and Goh, K.M., Thermophiles in the genomic era: Biodiversity, science, and applications, *Biotechnol. Adv.*, 2015, vol. 33, pp. 633–647.
- 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, e112963.
- Watanabe, T., Kojima, H., Umezawa, K., Hori, C., Takasuka, T.E., Kato, Y., and Fukui, M., Genomes of neutrophilic sulfur-oxidizing chemolithoautotrophs representing 9 proteobacterial species from 8 genera, *Front. Microbiol.*, 2019, vol. 10, art. 316.
- Zhang, T., Xu, J., Zeng, J., and Lou, K., Diversity of prokaryotes associated with soils around coal-fire gas vents in MaNasi county of Xinjiang, China, *Antonie van Leeuwenhoek*, 2013, vol. 103, pp. 23–36.

Translated by A. Bulaev