

1 **Metagenomic analyses of the late Pleistocene permafrost – additional tools for**
2 **reconstruction of environmental conditions**

3

4 Elizaveta Rivkina¹, Lada Petrovskaya², Tatiana Vishnivetskaya³, Kirill Krivushin¹, Lyubov
5 Shmakova¹, Maria Tutukina^{4,7}, Arthur Meyers³, Fyodor Kondrashov^{4,5,6}

6

7 ¹Institute of Physicochemical and Biological Problems in Soil Science, Russian Academy of
8 Sciences, Pushchino, Russia

9 ²Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of
10 Sciences, Moscow, Russia

11 ³University of Tennessee, Center for Environmental Biotechnology, Knoxville, USA

12 ⁴ Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG),
13 Barcelona, Spain

14 ⁵ Universitat Pompeu Fabra (UPF), Barcelona, Spain

15 ⁶Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

16 ⁷Institute of Cell Biophysics, Russian Academy of Sciences, Pushchino, Russia

17

18 *Corresponding author E-mail: elizaveta.rivkina@gmail.com*

19 Institute of Physicochemical and Biological Problems in Soil Science, Pushchino, Moscow
20 region, Russia, 142292

21

22 **Key words:** permafrost, Kolyma Lowland, methane, metagenome, methanogens,
23 methanotrophs, late Pleistocene Ice Complex, lake sediments

24 **Abstract**

25 A comparative analysis of the metagenomes from two 30,000 year-old permafrost samples,
26 one of lake-alluvial origin and the other from late Pleistocene Ice Complex sediments,
27 revealed significant differences within microbial communities. The late Pleistocene Ice
28 Complex sediments (which have been characterized by the absence of methane with lower
29 values of redox-potential and Fe^{2+} content) showed both a low abundance of methanogenic
30 archaea and enzymes from the carbon, nitrogen and sulfur cycles. The metagenomic and
31 geochemical analyses described in the paper provide evidence that the formation of the
32 sampled late Pleistocene Ice Complex sediments likely took place under much more aerobic
33 conditions than lake-alluvial sediments.

34

1. Introduction

Permafrost, including constantly frozen sediments of the Arctic, is a unique subsurface complex environment where microorganisms retain viability over a long period of time, from thousands to millions of years (Gilichinsky and Rivkina 2011). The impact of climate change on permafrost stability has recently been discussed widely by the scientific community (Zimov et al 2006, Walter et al 2007, Anthony et al 2014). The permafrost deposits of the North-East Siberia, which did not thaw during the Holocene climatic optimum, have attracted particular interest, especially the late Pleistocene Ice Complex deposits (Yedoma Suite) that are widespread on the Kolyma-Indigirka lowland (152-162°E, 68-72°N) (Schirrmeister et al 2011). Earlier we found that the epigenetically (consecutively) frozen sediments of both lake and marine origin (independent of age) contain biogenic methane, whereas methane was either absent or present at trace concentrations in samples from the sincryogenic (simultaneously frozen) late Pleistocene Ice Complex (Rivkina et al 2007, Rivkina and Kraev 2008). However, low concentrations of methane have been detected in the Yedoma deposits located both far south in boreal forest (Brouchkov and Fukuda 2002) and near the Arctic coast in Lena Delta (Brouchkov and Fukuda 2002, Bischoff et al 2013). The presence of biogenic methane in the Ice Complex deposits located near the coast of the Arctic Ocean was also confirmed by our research conducted in 2013 on the Gydan Peninsula (unpublished data).

Anaerobic microcosm incubation of thawed permafrost samples in a carbon dioxide- and hydrogen- enriched atmosphere showed methanogenic activity in epicryogenic sediments only, while this process was not observed in samples from the sincryogenic late Pleistocene Ice Complex (Rivkina and Kraev 2008). Similar results were obtained during experiments with radioactively labeled substrates (Rivkina et al 2002, Rivkina et al 2004, Rivkina et al

2007), showing an absence of methanogenic activity in the sincryogenic late Pleistocene Ice Complex samples with this process evidently taking place in epicryogenic sediments of lake or lake-alluvial origin. Moreover, only from the latter sediments methanogenic archaea were isolated in pure culture (Rivkina et al 2007, Krivushin et al 2010, Shcherbakova et al 2011). Viable methanogens were also isolated from the permafrost-affected soils located in the Lena River Delta region (Morozova et al 2007, Wagner et al 2013).

The specific distribution of methane and methane-producing microorganisms in permafrost raises number of questions; for example, why the sediments of the late Pleistocene Ice Complex in Kolyma–Indigirka Lowland do not contain methane or methanogenic activity. To answer this question, it is thought that the application of new methodologies such as metagenomic analyses is required (Graham et al 2012, Jansson and Tas 2014). Until recently, the determination of microbial diversity in low biomass environments, including permafrost, was problematic. The microbial cell abundances in the ancient permafrost is 10–100 times lower than that in the active layer samples, thereby resulting in low yields of the total community genomic DNA (gDNA) (Yergeau et al 2010). However, using appropriate DNA extraction kits (Vishnivetskaya et al 2014) and the whole-community genome amplification technique (Yergeau et al 2010), a sufficient amount of gDNA can be obtained for next-generation sequencing technologies, producing sequences on an unprecedented scale. Indeed, the first metagenomic analyses of permafrost samples became available recently. Specifically, the analyses of the metagenomes from active layer soil and two-meter deep permafrost samples collected in the Canadian High Arctic and Alaska identified signature genes responsible for hydrogenotrophic and acetoclastic methanogenesis, methylotrophic methane oxidation, nitrification, and carbohydrate degradation (Yergeau et al 2010, Mackelprang et al 2011).

Here we report results of the comparative metagenomic analyses of the two ancient permafrost samples similar in age (*ca.* 30,000 years old), however of different origins (epigenetically frozen lake sediments versus syncryogenic sediments from the late Pleistocene Ice Complex). The results based on comparative analyses of microbial communities from sediments of different origin may be used for evaluation of responses associated with permafrost thawing due to global warming.

2. Materials and Methods

Sample collection and description

Samples were collected within the Kolyma-Indigirka Lowland in northeast Siberia (69°29'N, 156°59'E) during the summer field season of 2007 (Fig. 1A). Permafrost sediments were sampled using drilling equipment that operates without fluids and prevents down-hole contamination. The sampling technique was tested and described previously (Shi et al 1997). Briefly, the surfaces of the 20- to 30-cm-long cores were cleaned immediately by shaving melted layers out with an ethyl alcohol-sterilized knife and then the frozen internal part of the core was split into 5-cm-long segments; these were placed into sterile aluminum containers and kept frozen during storage in field and transportation to the Institute of Physicochemical and Biological Problems in Soil Science, Pushchino.

The physicochemical characteristics of the samples are presented in Table 1. Methane content was measured as described previously (Rivkina et al 2007, Rivkina and Kraev 2008).

Sample IC4 corresponded to the permafrost sediment of lake origin from the floodplain of the Ambolikha River, borehole DH-4/07, depth of 22.5 m (Fig. 1B). Total carbon concentration was ~1.1% (w/w). Methane content of this sample was 1.2 mmol/kg,

$\delta^{13}\text{C} = -85\text{‰}$ indicative of biogenic origin. The radiocarbon age of this sample was $30,696 \pm 394$ years (J-5829) (Kraev et al 2013).

Sample IC8 represented a permafrost soil from the late Pleistocene Ice Complex (Omolon River), borehole DH-2/07, depth of 16 m. Total carbon concentration was $\sim 1.1\%$ (w/w). Methane levels in all samples tested from this borehole were non-detectable (Fig. 1B). The age of this sample was estimated to be $\sim 32,000$ years, based on the age determination for the same outcrop which was described recently (Legendre et al 2014).

2.1 DNA extraction and sequencing

In the laboratory, material from the inner part of the permafrost cores was subsampled aseptically for DNA isolation. The gDNA was extracted from eight replicates of ~ 0.5 g each randomly taken from ~ 50 g of permafrost core collected at corresponding depth (Figure 1) using the PowerSoil[®] DNA Extraction Kit (MO BIO Laboratories, Inc., USA). Due to low yield ($3\text{--}6 \text{ ng } \mu\text{l}^{-1}$), gDNAs from eight replicates were combined, then purified and concentrated using Genomic DNA Clean & Concentrator[®] Kit (Zymo Research Corporation, USA).

The gDNA sequencing libraries were prepared using NEBNext[®] reagents (New England BioLabs Inc., USA), according to protocol recommended by the manufacturer, having an estimated peak insert size of 150 nt. Metagenome sequencing was performed at the CRG Genomics Core Facility (Centre for Genomic Regulation, Barcelona, Spain) on an Illumina HiSeq 2000[™] machine using Flow Cell v3 with TruSeq SBS v3 reagents and a 2 x 100 cycle sequencing protocol.

2.2 MG-RAST Analysis

Raw sequencing data, i.e., 19.8 Gb representing 143.7 M sequences with an average length of 138 bp for IC4 and 19.7 Gb representing 131.7 M sequences with an average length of 150 bp for IC8, were uploaded to the MG-RAST server (Meyer et al 2008) for gene calling and annotation under ID 4606864.3 for IC4 and 4606865.3 for IC8. A total of 6.6% (IC4) and 3.4% (IC8) sequences failed to pass the quality control (QC) pipeline, whereas 0.3% of total sequences in both data sets were assigned to ribosomal RNA genes.

For functional assignment, protein sequences of putative ORF were searched against the M5NR non-redundant protein database (Wilke et al 2012) with an e-value threshold of $1e^{-5}$, minimum percentage identity of 60%, and minimum alignment length of 15 aa. The taxonomic assignments of Illumina reads were performed against M5NR and M5RNA databases at default parameters. The best-hit classification method was used in both cases for match assessment.

2.3 Comparative metagenomic data analyses

Statistical analyses of the metagenomes were performed in order to compare community composition and functional profiles of the permafrost samples. Statistical significance was calculated using two-sided Fisher's exact test. The differences between proportions were analyzed in STAMP (Parks and Beiko 2010) using MG-RAST taxonomic profile and the Newcombe-Wilson method (Newcombe 1998) at a 95% confidence interval and with Storey's FDR correction (Storey and Tibshirani 2003, Storey et al 2004). Original data sets were deposited at the NCBI Sequence Read Archive (SRA) under the accession numbers SRX763249, SRX751044 (Krivushin et al 2015).

3. Results and Discussion

3.1 Community description

The gDNA yield was higher in IC8 sample, with an average of $0.5 \mu\text{g g}^{-1}$ of wet sediment, in comparison to $0.37 \mu\text{g g}^{-1}$ in the IC4 sample. Efficiency of gDNA extraction from permafrost for the PowerSoil[®] DNA Extraction Kit (MO BIO Laboratories, Inc., USA) was estimated at $16 \pm 8\%$ (Vishnivetskaya et al 2014). Based on the metagenomics data (Angly et al 2009, Raes et al 2007) giving an average genome length of 4.7 Mb for the soil bacterial/archaeal population and an estimated weight of 4.05 fg (Ellenbroek and Cappenberg 1991) for a genome of this size, the theoretical level of the prokaryotic cell populations calculated from the total gDNA recovered accounting average extraction efficiency were 4.3×10^8 for IC4 and 5.9×10^8 for IC8, including a reduction of the total cell population by the eukaryotic component equal to 25% (Raes et al 2007).

The PowerSoil[®] DNA Extraction Kit (MO BIO Laboratories, Inc., USA) was selected based on its performance and cost per sample. For extractions of gDNA from permafrost, this kit provided similar microbial community structure results with minimal variation among triplicates ($P > 0.05$) (Vishnivetskaya et al 2014). Analyses of metagenomes of the two permafrost samples showed that bacterial genes were dominant and 96.4% and 97.7% of sequences were assigned to the domain Bacteria in the IC4 and IC8 samples, respectively. Archaea were the second dominant domain followed by Eukaryotes, while viruses comprised only 0.06% in IC4 and 0.03% in IC8 samples. Archaea and Eukaryotes were more abundant in the IC4 sample compared to the IC8 sample, i.e., 2.4% versus 1.3% and 1.0 versus 0.9%, respectively. While Bacteria were significantly more abundant than Eukaryotes, the functional diversity was higher among the Eukaryotic organisms, which play essential roles in the biology and fertility of soils. A number of detected unparalleled

genes of Eukaryotic organisms including plants, algae, fungi, and protists was 3.4 and 2.6 times higher in comparison to a number of Bacterial unparalleled genes detected in the IC4 and IC8 samples, respectively. The most abundant microorganisms were comprised of nine bacterial and one Archaeal phyla, with Proteobacteria being the most plentiful, followed by Actinobacteria, Firmicutes, Bacteroidetes, Chloroflexi, Planctomycetes, Euryarchaeota, Acidobacteria, Cyanobacteria, and Verrucomicrobia in that order (Fig. 2). The quantities of all other phyla were less than 1%. Sequences affiliated with Actinobacteria ($P < 1e^{-15}$), Firmicutes ($P < 1e^{-15}$), Chloroflexi ($P < 1e^{-15}$), Cyanobacteria ($P < 1e^{-15}$), Acidobacteria ($P < 1e^{-15}$), Verrucomicrobia ($P < 1e^{-15}$) were dominant within IC8, whereas α -, β -, γ -, and δ -Proteobacteria ($P < 1e^{-15}$), Bacteroidetes ($P < 1e^{-15}$), Planctomycetes ($P < 1e^{-15}$), and Euryarchaeota ($P < 1e^{-15}$) dominated within IC4. The microorganisms that are involved into methane formation and methane metabolism, for example methanogenic archaea (phylum Euryarchaeota) or methanotrophic and methylotrophic bacteria (phylum Proteobacteria), were more abundant in more reduced IC4 sample (Table 1). Microorganisms without strong association to methane metabolism (e.g. phyla Actinobacteria and Chloroflexi) dominated in IC8.

3.2 Dependence of the microbial community structure on permafrost characteristics as detected from genus taxonomical level

3.2.1 Methanogenic community. As was shown in previous research (Kraev et al 2013, Legendre et al 2014), methane was not detected in the IC8 sample, but its level in the IC4 sample reached 1-2 mmol/kg; hence, one would expect to find a greater abundance of methanogenic species in the IC4 sample. Indeed, the phylum Euryarchaeota, which contains methanogens, was twice as prevalent in IC4 sample, and methanogenic Archaea constituted as much as 1.5% of the microbial communities in IC4 (compared to 0.5% in IC8). The most

abundant methane-producing species was *Methanosarcina* (0.32% in IC4 versus 0.14% in IC8), followed by *Methanoregula* (0.28% versus 0.03%), *Methanoculleus* (0.1% versus 0.05%), *Methanosphaerula* (0.1% versus 0.03%), *Methanospirillum* (0.1% versus 0.03%), *Methanosaeta* (0.1% versus 0.03%). Twelve other methanogenic species were detected at <0.1% in IC4 and at an even lower percentage in IC8.

3.2.2 Methanotrophic bacteria. In the vicinity of the environments where methane is produced, methane-oxidizing (methanotrophic) bacteria can be found (Khmelenina et al 2002). Thus, we analyzed the presence of methanotrophic DNA in our data. To date two types of methane-oxidizing bacteria are recognized, Type I methanotrophs belonging to γ -Proteobacteria and type II methanotrophs from α -Proteobacteria. Indeed, the phylum Proteobacteria clearly dominated in the IC4 sample (50.0% versus 26.5% in IC8) with the α -Proteobacteria being more abundant in IC4 (32.7%) in comparison to IC8 (12.3%). On a more refined taxonomical scale the most abundant order of α -Proteobacteria was Rhizobiales (24.2% in IC4 versus 7.7% in IC8), which contains both methanotrophic and nitrogen-fixing bacteria. The α -Proteobacteria Type II methanotrophs (3.32% versus 1.06%) were represented by the genera *Methylocella* (0.53% versus 0.14%), *Methylosinus* (0.24% versus 0.07%), *Methylocystis* (0.22% versus 0.06%), and *Methylobacterium* (2.32% versus 0.8%) in IC4 and IC8. The last genus (*Methylobacterium*) is a facultative methylotroph, however, some species are capable of growth on methane. The class γ -Proteobacteria was the most diverse Proteobacteria class (184 species); nonetheless, it was six-fold less plentiful in comparison to α -Proteobacteria. The γ -Proteobacteria Type I methanotrophs, such as *Methylococcus* (0.13% versus 0.11%), *Methylobacter* (0.12% versus 0.05%), and *Methylophaga* (0.02% versus 0.01%) were again more abundant in the IC4 than in the IC8 sample. In this Siberian permafrost, Type II methanotrophs dominated over Type I methanotrophs, which is similar to methanotrophic bacteria abundance and diversity in

Canadian high Arctic permafrost (Lau et al 2015). Methylophages, as well as the subset methanotrophs, play an essential role in the carbon cycle. Interestingly, obligate methylophagic bacteria belonging to the β -Proteobacteria, such as *Methylobium* (0.17% versus 0.16%), *Methylobacillus* (0.07% versus 0.08%), *Methylovorus* (0.04% versus 0.04%), and *Methylothermus* (0.06% versus 0.05%), were detected in both the IC4 and IC8 samples at equally low levels.

3.2.3 Bacteria of nitrogen cycle. Another important metabolic process in an environment is the nitrogen cycle. Nitrogen-fixing bacteria were more abundant in the IC4 of the following genera: *Bradyrhizobium* (1.85% versus 0.5%), *Sinorhizobium* (1.3% versus 0.6%), *Rhizobium* (0.82% versus 0.4%), *Rhodospirillum* (0.63% versus 0.3%), *Azospira* (0.4% versus 0.12%), the γ -Proteobacteria nitrogen-fixing species. Other species involved in the nitrogen cycle from the *Hyphomicrobium* (2.5% versus 0.23%, capable of denitrification with methanol), *Nitrobacter* (1.9% versus 0.6%, capable of oxidizing nitrite into nitrate) and *Rhodopseudomonas* (3.8% versus 1.2%, capable of carbon dioxide and nitrogen fixation) genera were also more abundant in the IC4 sample. However, ammonia-oxidizing and nitrifying bacteria of the class β -Proteobacteria, such as *Nitrosomonas* (~0.1%), *Nitrosospora* (~0.17%), were detected in both samples at the similar level. Other bacteria involved in the nitrogen cycle are members of the phylum Planctomycetes, many of which conduct anaerobic ammonium oxidation or so-called "anammox" metabolism, a process of ammonia oxidation by nitrite involvement to yield nitrogen gas. Four planctomycetes genera were more abundant in IC4 compared to the IC8 sample with *Planctomyces* (0.63% versus 0.34%) being the most abundant, followed by *Pirellula* (0.61% versus 0.28%), *Blastopirellula* (0.57% versus 0.24%), and *Isosphaera* (0.16% versus 0.13%). Some planctomycetes, e.g., *Pirellula*, are able to live in environments with high inorganic sulfate concentrations (Glockner et al 2003). Nitrogen-fixing cyanobacteria slightly dominated in IC8 (0.82%) in

comparison to IC4 (0.71%); however, the proportion of nitrogen-fixing to total amount of cyanobacteria was higher in the IC4 (55.5%) than in the IC8 (48.6%) sample. Another nitrogen-fixing bacterium dominating in IC8 (2.6%) in comparison to IC4 (0.8%) was the actinobacterium *Frankia*, which is characterized by the ability to engage in a symbiotic relationship with plants, producing nitrogen-fixing root nodules.

3.2.4 Sulfate-reducing bacteria. The δ -Proteobacteria were also more abundant in IC4 (5.8%) than in IC8 (4.5%). Sulfate-reducing bacteria, namely, *Desulfovibrio* (0.49% versus 0.41%), *Desulfatibacillum* (0.22% versus 0.08%), *Desulfococcus* (0.18% versus 0.07%), *Desulfobacterium* (0.14% versus 0.04%), *Desulfomicrobium* (0.11% versus 0.06%), and metal-reducing bacteria of the genus *Geobacter* (1.13% versus 0.82%) were more plentiful in the IC4 sample. Two species from the order Syntrophobacterales were more abundant in IC4; these included the strictly anaerobic, sulfate-reducing, propionate-degrading bacterium *Syntrophobacter* (0.33% versus 0.18% in IC8) and the benzoate-degrading bacterium *Syntrophus* (0.62% versus 0.13%). During growth on certain compounds, both of these organisms are known to form syntrophic associations with methanogens, e.g., *Methanospirillum hungatei*, facilitating methane production (Harmsen et al 1998, Jackson et al 1999). The sulfate-reducing bacterium *Desulfotomaculum* from the phylum Firmicutes was found at similar concentration in both samples (i.e., 0.32-0.33%). However, another strictly anaerobic bacterium *Desulfitobacterium*, which is capable of using a wide variety of electron acceptors, such as nitrate, sulfite, metals, humic acids, and halogenated organic compounds can use hydrogen gas as an electron donor at extremely low concentrations to facilitate sulfate reduction and methanogenesis (Villemur et al 2006) was twice as abundant in the IC4 (0.2%) than in the IC8 sample (0.1%). Sulfate-reducing *Thermodesulfovibrio* species from the Nitrospirae division were found at 0.09% and 0.06% in the IC4 and IC8 samples, respectively.

3.3 Similarities and dissimilarities in the microbial communities based on functional annotation

In contrast to the taxonomical assignment, the functional annotation of the metagenomes exhibited a similar structure for the microbial communities. Essential portions of sequencing data ranging from 51.9% in IC4 to 49.1% in IC8 were assigned to unknown proteins, while 37.6% (IC4) and 44.1% (IC8) were annotated proteins. A comparison of the metagenomes at the function level using the SEED - the database for bioinformatic research (Overbeek et al 2005) genome annotations and profile scatter plot showed that the IC4 and IC8 metagenomes possess >83% similarity at the function level (Fig. 1S). Functional analysis of the IC4 and IC8 metagenomes demonstrated that among the annotated protein sequences, the most abundant groups represented housekeeping functions, including carbohydrate metabolism (16.6% and 17.7%), amino acid biosynthesis (13.8% and 14.8%), and vitamin and protein metabolism (12.1% and 13.2%). Approximately 4.1% and 3.7% of annotated reads were related to stress response including a high proportion of oxidative stress- (52% and 48%) and osmotic stress- (13% and 12%) related sequences (Fig. 3). The functional potential documented through the metagenome provides the information on how many different genes may exist in a sample, but it does not give any indication if these genes are expressed in permafrost. If slow metabolic activity at subfreezing temperatures in permafrost exists as shown in the laboratory studies (Rivkina et al 2000) then such a community may develop adaptation directed by sediments' physicochemical properties.

3.3.1 Methane metabolism. Analyses using KEGG (Kyoto Encyclopedia of Genes and Genomes) database (Kanehisa and Goto 2000) demonstrated that several functional genes directly related to methanogenesis are abundant in the IC4 metagenome (Fig. 4), including genes coding for F420-dependent methylene-H4 MPT reductase (Enzyme Commission

number, EC 1.5.99.11), formylmethanofuran dehydrogenases (*fmd*) (EC 1.2.99.5), CoB-CoM heterodisulfide reductases (EC 1.8.98.1), F420-reducing hydrogenases (EC 1.12.98.1), and methylenetetrahydromethanopterin dehydrogenases (*mer*) (EC 1.5.99.9). All of these genes, with exception of F420-dependent methylene-H4 MPT reductase, were also present in the IC8 metagenome though in lower amounts. For example, the gene coding for the enzyme catalyzing the last step of methanogenesis, methyl-coenzyme M reductase (MCR) (EC 2.8.4.1), was represented by 1333 hits in IC4 but only by nine hits in the IC8 metagenome.

3.3.2 Nitrogen metabolism. The abundance of genes associated with nitrogen metabolism was higher in the IC4 metagenome than in that of IC8 (2.2% versus 1.7%). The percentage of genes related to nitrogen metabolism detected in permafrost was slightly higher than the relative abundance (1.3%) of genes responsible for nitrogen metabolism reported in metagenomes examined from glacier ice of the Northern Schneeferner (Simon et al 2009). Notably, in the IC8 metagenome the gene *nifH* coding for a nitrogenase (EC 1.18.6.1), which performs nitrogen fixation, was represented only by seven sequences. These genes were associated mainly with Proteobacteria (43.7%) represented by *Bradyrhizobium* sp., *Azoarcus* sp., *Mesorhizobium loti*, and Actinobacteria (*Frankia* sp., 41%). In the IC4 metagenome 90 features corresponding to *nifH* gene were detected, and these nitrogenase sequences were linked predominantly with Proteobacteria (33.5%). Nitrogenase genes related to Firmicutes (primarily *Desulfitobacterium hafniense* and '*Alkaliphilus metalliredigens*') and Actinobacteria (*Frankia* species) constituted 17.6% and 15%, respectively. Interestingly, *nifH* genes of cyanobacterial origin (primarily linked to *Nostoc* species) were detected exclusively within the IC4 metagenome (12.9%).

321 Genes connected with the denitrification processes, such as nitrate reductase (EC 1.7.99.4)
322 and nitrite reductase (EC 1.7.2.1) were found in both metagenomes (Fig. 4). The nitrate
323 reductase (*narG*, EC 1.7.99.4) sequences predominating within the IC4 metagenome came
324 from more diverse phylogenetic groups in contrast to *narG* gene from IC8. By contrast, even
325 though nitrite reductase (*nirS*) genes were significantly overrepresented in IC8, their
326 presence was detected in similar phylogenetic groups in both metagenomes (Fig. 5). The
327 sequences related to both nitrite reductase (EC 1.7.1.4) and nitric oxide reductase (EC
328 1.7.2.5) were found in similar phylogenetic groups in both metagenomes with prevalence in
329 IC4 (Fig. 5).

330 The ammonium oxidation pathway was represented by a few sequences related to
331 hydroxylamine oxidase (EC 1.7.3.6) genes in known nitrifying bacteria such as
332 *Nitrosomonas eutropha*, *Nitrosococcus oceani*, *Nitrosospira multiformis* and some others.
333 Genes coding for ammonia monooxygenase (EC 1.14.99.39) were not detected in either
334 metagenome by a search with its EC number. However, a search for ammonia
335 monooxygenase using functional hierarchies such as KEGG orthologs yielded ten hits in
336 IC4 and four hits in IC8, while SubSystems showed presence of 509 hits in IC4 and 324 hits
337 in IC8. The ammonia monooxygenase sequences were annotated as methane
338 monooxygenase (EC1.14.13.25). It should be noted that the particulate methane
339 monooxygenase and ammonia monooxygenase are related and occur in both methanotrophs
340 and ammonia oxidizers (Holmes et al 1995). These enzymes have wide substrate specificity
341 catalyzing the oxidation of various substrates including ammonia, methane, halogenated
342 hydrocarbons, and aromatic molecules (Arp et al 2002). Overall, the low abundance of
343 ammonia-oxidizers presumably represents the rare utilization of this pathway by permafrost
344 bacteria in studied environments. Similar results were also reported for metagenomes from

other cold environments, e.g., high Arctic hypersaline subzero spring (Lay et al 2013) and Arctic snow packs (Larose et al 2013).

3.3.3 Sulfur metabolism. Sequences associated with sulfur metabolism were present in both IC4 and IC8 metagenomes and related to both reduction and oxidation (Fig. 6). Genes coding for sulfate reduction were more abundant in the IC8 metagenome including genes for sulfate adenylyltransferase (EC 2.7.7.4), phosphoadenylyl-sulfate reductase (EC 1.8.4.8), and ferredoxin-sulfite reductase (EC 1.8.7.1). Taxonomic distribution of associated species was similar in both metagenomes with the exception of sulfate adenylyltransferase, which was represented in IC4 by sequences related mainly to Proteobacteria and Actinobacteria. In the IC8 sample this gene was of more diverse phylogenetic origin (Fig. 6). A few sulfur oxidation genes detected were associated with *Renibacterium salmoninarum* and *Gordonia bronchialis* in IC8 and *Mycobacterium* species and *Sinorhizobium meliloti* in IC4.

3.3.4 Stress response. Genes associated with stress response were detected in both of the metagenomes (33683 hits in IC4 metagenome and 28557 in IC8). The three most abundant groups present corresponded to oxidative stress, heat shock, and osmotic stress response genes. Sequences related to oxidative stress originated principally from Proteobacteria, Actinobacteria and Firmicutes (5148 features in IC4 and 3832 features in IC8) and included genes for catalase (EC 1.11.1.6), peroxidase (EC 1.11.1.7), and different superoxide dismutases (EC 1.15.1.1.). Their occurrence is presumably explained by increased oxygen solubility at low temperatures and associated increase of reactive oxygen species concentration (Chattopadhyay 2006). Sequences related to osmotic stress were represented by the genes involved in the synthesis and uptake of compatible solutes including choline, betaine, periplasmic glucan, and ectoine. Genes for osmoprotectant ABC transporters were also detected. Choline dehydrogenase (EC 1.1.99.1) (222 features in IC4 and 213 features in

IC8 mainly from Proteobacteria and Actinobacteria) and betaine-aldehyde dehydrogenase (EC 1.2.1.8) (166 features in IC4 and 186 features in IC8 from Proteobacteria, Actinobacteria and Firmicutes) were the most abundant enzymes of this class. This emphasizes the importance of betaine osmolyte for the osmoprotection of members in microbial communities from subfreezing environments. The genes encoded the heat shock proteins were mainly represented by the chaperone protein DnaK (816 hits in IC4 and 54 in IC8) and its interacting protein DnaJ (759 hits in IC4 and 67 in IC8). These proteins are among the most plentiful chaperons in the bacterial cell and often prevalent in microorganisms from cold environments (D'Amico et al 2006).

3.4. Implication of metagenomic sequencing data for the analysis of the samples origin and evolution.

The application of biological markers for paleo-reconstructions in various environmental sites has been used occasionally. For example, utilization of lipid analyses for petroleum reservoirs formation and maturation (Seifert and Moldowan 1981); analysis of fossil chironomid assemblages in the Holocene lake-sediment cores (south-central Alaska) for evaluating anthropogenic climatic changes and quantitative paleo-temperature reconstructions (Clegg et al 2010); and analyses of fossil ostracodal assemblages from the Arctic seas for reconstruction of coastline and interpretation of environmental differences in Arctic areas (Stepanova et al 2010). To track the occurrence and distribution of microorganisms in the environment, the gDNA and DNA fragments amplified with PCR can be employed as biological markers. Therefore, we anticipated that integration of the next generation sequencing capabilities and approaches of microbial ecology (such as linking microbial community composition and environmental processes involved in carbon, nitrogen and sulfur cycling) could be used for paleo-reconstructions.

393 The late Pleistocene Ice Complex or Yedoma deposits are widely distributed in northeast
394 Siberia and because of their wide occurrence on the Kolyma lowland, these deposits may
395 play a significant role in climate warming, permafrost degradation and greenhouse gases
396 emission. A question, which processes formed Yedoma has been under dispute in the last
397 several decades. Several hypotheses have been proposed about the origin of the late
398 Pleistocene Ice Complex, including eolian (Tomirdiaro et al 1984, Tomirdiaro and
399 Chernen'k'ii 1987), alluvial (Rozenbaum 1981), and polygenetic (Konishchev and
400 Kolesnikov 1981, Sher et al 1987) formation. Different opinions on the origin of these
401 deposits have been summarized in the recent publications of Lutz Schirrmeister and co-
402 authors (Schirrmeister et al 2011, Schirrmeister et al 2013). The researchers suggested that
403 the ice rich syngenetic permafrost of the late Pleistocene Ice Complex was developed under
404 a cold-arid climate at less hydromorphic conditions than the lake and lake-alluvial
405 sediments. In general, Yedoma sediments have been characterized by the absence of
406 methane (Rivkina et al 2007, Rivkina and Kraev 2008) and much lower redox-potential and
407 iron (Fe^{2+}) content in comparison to permafrost layers of lake and lake-alluvial origin
408 (Rivkina et al 2006).

409 A comparison of the two late Pleistocene permafrost metagenomes of different genesis, IC4
410 and IC8, revealed differences in the composition of the microbial community that reflects
411 the conditions under which these deposits were formed. These data uncovered significant
412 distinctions in microbial community compositions between Yedoma and lake-alluvial
413 sediments. The relatively low abundance of methanogenic archaea, limited presence of
414 enzymes from the carbon, nitrogen, and sulfur cycles, as well as the presence of
415 methanotrophic bacteria could explain the absence of methane in Yedoma deposits and
416 provide evidence that the formation of these sediments took place under much more aerobic
417 conditions. In other words, we may assume that during the late Pleistocene period, nearly

30,000 years ago, different environmental conditions predetermined biogeochemical regimes and composition of microbial communities of the studied samples. Similarly to our study the recent survey of soil taxonomic, phylogenetic, and functional diversity have demonstrated that metagenomic approaches can be used to build a predictive understanding of variations in microbial diversity and functions across terrestrial biomes (Fierer et al 2012). Thus, the involvement of metagenomic analyses, along with geological and biogeochemical methods, may be used for characterization of the permafrost microbial community, its contribution during climate warming and permafrost thawing, as well as it may become an additional instrument in paleo-reconstructions.

4. Conclusions

In the current study, we performed a comparative analysis of the two permafrost samples isolated from lake sediments and ice complex with utilization of metagenomic sequencing approach. We have shown that the presence of methane in one sample and its absence in the second are associated with the prevalence of the certain groups of microorganisms and their corresponding genes. The possible scenario of the development of both Yedoma and lake sediments can be proposed on the basis of these findings. Analysis of metagenome from IC4, a sample of the lake sediments, demonstrated occurrence of many physicochemical reactions such as denitrification, iron reduction and sulfate reduction, which could reduce environmental redox potential and ultimately create favorable conditions for development of methanogenic community and methanogenesis. As reflected from the composition of IC8 metagenome, the sporadic occurrence of such physicochemical reactions brought to deficiency of methanogenic activity and lack of biogenic methane in the late Pleistocene ice complex on Kolyma-Indigirka lowland.

The obtained results demonstrate that the metagenomic analysis of permafrost microbial communities can represent a valuable instrument for paleo-reconstruction of conditions under which the permafrost sediments were formed in geological perspective. However more studies would be needed before a predictive capability can be achieved with respect to microbial community function after permafrost thaw.

Author contribution:

E.R., L.P. and F.K. designed research; L.S., M.T., L.P., E.R. performed research; L.S. and M.T. extracted DNA; F.K. performed DNA sequencing; E.R., L.P., K.K., T.V., A.M. analyzed data; E.R. and F.K. contributed new reagents and analytical tools; E.R., L.P., T.V. wrote the paper.

Acknowledgements

This work was supported by grants from the Russian Scientific Fund (14-14-01115) to E.R.; from the National Science Foundation (DEB-1442262) to T.V.; and by HHMI International Early Career Scientist Program (55007424), the EMBO Young Investigator Programme, MINECO (BFU2012-31329 and Sev-2012-0208), AGAUR program (2014 SGR 0974) to F.K. The Russian Scientific Fund (14-14-01115) support was allocated for sample collection, gDNA isolation, and analysis of metagenomic data.

462 **References**

- 463 Angly, F. E., Willner, D., Prieto-Davó, A., Edwards, R. A., Schmieder, R., Vega-Thurber,
464 R., Antonopoulos, D. A., Barott, K., Cottrell, M. T., Desnues, C., Dinsdale, E. A., Furlan,
465 M., Haynes, M., Henn, M. R., Hu, Y., Kirchman, D. L., McDole, T., McPherson, J. D.,
466 Meyer, F., Miller, R. M., Mundt, E., Naviaux, R. K., Rodriguez-Mueller, B., Stevens, R.,
467 Wegley, L., Zhang, L., Zhu, B., and Rohwer, F.: The GAAS metagenomic tool and its
468 estimations of viral and microbial average genome size in four major biomes, PLoS Comput
469 Biol. 5, e1000593, doi:10.1371/journal.pcbi.1000593, 2009.
- 470 Anthony Walter, K. M., Zimov, S. A., Grosse, G., Jones, M. C., Anthony, P. M., Chapin III,
471 F. S., Finlay, J. C., Mack, M. C., Davydov, S., Frenzel, P., and Frolking, S.: A shift of
472 thermokarst lakes from carbon sources to sinks during the Holocene epoch, Nature, 511,
473 452–456, 2014.
- 474 Arp DJ, Sayavedra-Soto LA, Hommes NG: Molecular biology and biochemistry of
475 ammonia oxidation by *Nitrosomonas europaea*. Arch Microbiol 178, 250-255, 2002.
476 Bischoff, J., Mangelsdorf, K., Gattinger, A., Schlöter, M., Kurchatova, A., Herzsich, U.,
477 Wagner, D.: Response of methanogenic archaea to Late Pleistocene and Holocene climate
478 changes in the Siberian Arctic, Global Biogeochem Cy, 27, 305–317, 2013.
479
- 480 Brouchkov A., Fukuda M.: Preliminary Measurements on Methane Content in Permafrost,
481 Central Yakutia, and some Experimental Data 21. Permafrost Periglac, 13, 187–197, 2002.
- 482 Chattopadhyay M.K.: Mechanism of bacterial adaptation to low temperature. J Biosciences,
483 31, 157-165, 2006.
- 484 Clegg, B. F., Clarke, G. H., Chipman, M. L., Chou, M., Walker, I. R., Tinner, W., and Hu F.
485 S.: Six millennia of summer temperature variation based on midge analysis of lake
486 sediments from Alaska, Quaternary Sci Rev, 29, 3308–3316, 2010.
- 487 D'Amico S, Collins T, Marx JC, Feller G, Gerday C.: Psychrophilic microorganisms:
488 challenges for life. EMBO Rep, 7, 385-389, 2006.
- 489 Ellenbroek, F. M. and Cappenberg, T.E.: DNA-synthesis and tritiated-thymidine
490 incorporation by heterotrophic fresh-water bacteria in continuous culture, Appl. Environ
491 Microb, 57, 1675–1682, 1991.
- 492 Fierer N., Leff J.W., Adams B.J., Nielsen U.N., Bates S.T., Lauber C.L., Owens, S., Gilbert
493 J.A., Caporaso J.G.: Cross-biome metagenomic analyses of soil microbial communities and
494 their functional attributes, Proc Natl Acad Sci USA, 109, 21390-21395, 2012.
- 495 Gilichinsky, D. and Rivkina, E.: Permafrost microbiology, in: Encyclopedia of Geobiology,
496 edited by: Reitner, J. and Thiel, V., Springer, Dordrecht, the Netherlands, 726–732, 2011.
- 497 Glockner, F. O., Kube, M., Bauer, M., Teeling, H., Lombardot, T., and Ludwig, W., Gade,
498 D., Beck, A., Borzym, K., Heitmann, K., Rabus, R., Schlesner, H., Amann, R., and
499 Reinhardt, R.: Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1,
500 Proc Natl Acad Sci USA, 100, 8298–8303, 2003.

- 501 Graham, D. E., Wallenstein, M. D., Vishnivetskaya, T. A., Waldrop, M. P., Phelps, T. J.,
502 Pfiffner, S. M., Onstott, T. C., Whyte, L. G., Rivkina, E. M., Gilichinsky, D. A., Elias, D.
503 A., Mackelprang, R., VerBerkmoes, N. C., Hettich, R. L., Wagner, D., Wulfschleger, S. D.,
504 and Jansson, J. K.: Microbes in thawing permafrost: the unknown variable in the climate
505 change equation, *ISME J*, 6, 709–712, 2012.
- 506 Harmsen, H. J. M., Van Kuijk, B. L. M., Plugge, C. M., Akkermans, A. D. L., De Vos, W.
507 M., and Stams, A. J. M.: *Syntrophobacter fumaroxidans* sp. nov., a syntrophic propionate-
508 degrading sulfate-reducing bacterium, *Int J Syst Bacteriol*, 48, 1383–1387, 1998.
- 509 Holmes, A. J., Costello, A., Lidstrom, M. E., and Murrell, J. C.: Evidence that
510 particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily
511 related, *FEMS Microbiol Lett*, 132, 203–208, 1995.
- 512 Jackson, B. E., Bhupathiraju, V. K., Tanner, R. S., Woese, C. R., and McInerney, M. J.:
513 *Syntrophus aciditrophicus* sp. nov., a new anaerobic bacterium that degrades fatty acids
514 and benzoate in syntrophic association with hydrogen-using microorganisms, *Arch*
515 *Microbiol*, 171, 107–114, 1999.
- 516 Jansson, J. K. and Tas, N.: The microbial ecology of permafrost, *Nature Rev Microbiol*, 12,
517 414–425, 2014.
- 518 Kanehisa M., Goto S.: KEGG: Kyoto Encyclopedia of Genes and Genomes, *Nucleic Acids*
519 *Res*, 28, 27–30, 2000.
- 520 Khmelenina, V. N., Makutina, V. A., Kalyuzhnaya, M. G., Rivkina, E. M., Gilichinsky, D.
521 A., and Trotsenko, Y.: Discovery of viable methanotrophic bacteria in permafrost sediments
522 of northeast Siberia, *Dokl Biol Sci*, 384, 235–237, 2002.
- 523 Konishchev, V. N. and Kolesnikov, S. F.: Specialities of structure and composition of late
524 Cenozoic deposits in the section of Oyogossky Yar, in: *Problems of Cryolithology*, MGU
525 Publishing, Moscow, 107–117, 1981 (in Russian).
- 526 Kraev, G. N., Schultze, E. D., and Rivkina, E. M.: Cryogenesis as a factor of methane
527 distribution in layers of permafrost, *Dokl Earth Sci*, 451, 882–885, 2013.
- 528 Krivushin, K. V., Shcherbakova, V. A., Petrovskaya, L. E., and Rivkina, E. M.:
529 *Methanobacterium veterum* sp nov., from ancient Siberian permafrost, *Int. J. Syst. Evol.*
530 *Micr.*, 60, 455–459, 2010.
- 531 Krivushin, K., Kondrashov, F., Shmakova, L., Tutukina, M., Petrovskaya, L., and Rivkina,
532 E.: Two metagenomes from Late Pleistocene northeast Siberian permafrost, *Genome*
533 *Announc*, 3, e01380–14, doi:10.1128/genomeA.01380-14, 2015.
- 534 Larose, C., Dommergue, A., and Vogel, J. G.: Microbial nitrogen cycling in Arctic
535 snowpacks, *Environ Res Lett*, 8, 035004, 2013.
- 536 Lau, M. C. Y., Stackhouse, B. T., Layton, A. C., Chauhan, A., Vishnivetskaya, T. A.,
537 Chourey, K., Ronholm, J., Mykytczuk, N. C. S., Bennett, P. C., Lamarche-Gagnon, G.,
538 Burton, N., Pollard, W. H., Omelon, C. R., Medvigy, D. M., Hettich, R. L., Pfiffner, S. M.,

539 Whyte, L. G., and Onstott, T. C.: An active atmospheric methane sink in high Arctic mineral
540 cryosols, *ISME J.*, 9, 1880–1891, 2015.

541 Lay C.Y., Mykytczuk N.C., Yergeau E., Lamarche-Gagnon G., Greer C.W., Whyte L.G.:
542 Defining the functional potential and active community members of a sediment microbial
543 community in a high-arctic hypersaline subzero spring, *Appl Environ Microb*, 79, 3637–
544 3648, 2013.

545 Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., Lescot, M.,
546 Poirot, O., Bertaux, L., Bruley, C., Couté, Y., Rivkina, E., Abergel, C., and Claverie, J-M.:
547 Thirty thousand-year-old distant relative of giant icosahedral DNA viruses with a
548 pandoravirus morphology, *Proc Natl Acad Sci USA*, 111, 4274–4279, 2014.

549 Mackelprang, R., Waldrop, M. P., DeAngelis, K. M., David, M. M., Chavarria, K. L.,
550 Blazewicz, S. J., Rubin, E. M., and Jansson, J. K.: Metagenomic analysis of a permafrost
551 microbial community reveals a rapid response to thaw, *Nature*, 480, 368–371, 2011.

552 Meyer, F., Paarmann, D., D’Souza, M., Olson, R., Glass, E., Kubal, M., Paczian, T.,
553 Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., and Edwards, R. A.: The
554 metagenomics RAST server – a public resource for the automatic phylogenetic and
555 functional analysis of metagenomes, *BMC Bioinformatics*, 9, 386, doi:10.1186/1471-2105-
556 9-386, 2008.

557 Morozova, D., Wagner D.: Stress response of methanogenic archaea from Siberian
558 permafrost compared with methanogens from nonpermafrost habitats, *FEMS Microbiol*
559 *Ecol*, 61, 16–25, 2007.

560
561 Newcombe, R. G.: Two-sided confidence intervals for the single proportion: comparison of
562 seven methods, *Stat Med*, 17, 857–872, 1998.

563 Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H-Yu., Cohoon, M.,
564 Vale’rie de Cre’cy-Lagard, Diaz N., Disz, T., Edwards, R., Fonstein, M., Frank, E.D.,
565 Gerdes, S., Glass, E.M., Goesmann, A., Hanson, A. Iwata-Reuy, D., Jensen, R., Jamshidi, N.,
566 Krause L., Kuba, M., Larsen, L., Linke, B., McHardy, A.C., Meyer, F., Neuweiger, H.,
567 Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G.D., Rodionov, D.A., Ruckert, C.,
568 Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Yu., Zagnitko, O. and Vonstein, V. : The
569 subsystems approach to genome annotation and its use in the project to annotate 1000
570 genomes. *Nucleic Acids Res*, 33, 5691-5702, 2005.

571 Parks, D. H. and Beiko, R. G.: Identifying biologically relevant differences between
572 metagenomic communities, *Bioinformatics*, 26, 715–721, 2010.

573 Raes, J., Korb, J. O., Lercher, M. J., von Mering, C., and Bork, P.: Prediction of effective
574 genome size in metagenomic samples, *Genome Biol*, 8, R10, doi:10.1186/gb-2007-8-1-r10,
575 2007.

576 Rivkina, E. and Kraev, G.: Permafrost degradation and influx of biogeogases into the
577 atmosphere, Ninth International Conference on Permafrost, Fairbanks, USA, University of
578 Alaska Fairbanks, 29 June–03 July, 1499–1504, 2008.

- 579 Rivkina, E., Laurinavichius, K., McGrath, J., Tiedje, J., Shcherbakova, V., and Gilichinsky,
580 D.: Microbial life in permafrost, *Adv Space Res*, 33, 1215–1221, 2004.
- 581 Rivkina, E., Kraev, G., Krivushin, K., Laurinavichius, K., Fyodorov-Davydov, D.,
582 Kholodov, A., Shcherbakova, V., and Gilichinsky, D.: Methane in permafrost of
583 northeastern Arctic, *Earth Cryosphere* (Russian), 10, 23–41, 2006.
- 584 Rivkina, E., Shcherbakova, V., Laurinavichius, K., Petrovskaya, L., Krivushin, K., Kraev,
585 G., Pecheritsina, S., and Gilichinsky, D.: Biogeochemistry of methane and methanogenic
586 archaea in permafrost, *FEMS Microbiol Ecol*, 61, 1–15, 2007.
- 587 Rivkina, E. M., Friedmann, E. I., McKay, C. P. and Gilichinsky, D. A.: Metabolic activity of
588 permafrost bacteria below the freezing point, *Appl Environ Microbiol*, 66, 3230–3233,
589 2000.
- 590 Rozenbaum, G. E.: Special features of lithogenesis of the alluvial planes in the Eastern
591 Subarctic as related to the problem of the Ice (Yedoma) Complex, in: *Problems of*
592 *Cryolithology*, MSU Press, Moscow, 87–100, 1981 (in Russian).
- 593 Schirrmeister, L., Kunitsky, V., Grosse, G., Wetterich, S., Meyer, H., Schwamborn, G.,
594 Babi, O., Derevyagin, A., and Siegert, C.: Sedimentary characteristics and origin of the
595 Late Pleistocene Ice Complex on North-East Siberian Arctic coastal lowlands and islands –
596 a review, *Quatern Int*, 241, 3–25, 2011.
- 597 Schirrmeister, L., Froese, D., Tumskey, V., Grosse, G., and Wetterich, S.: Yedoma: late
598 Pleistocene ice-rich syngenetic permafrost of Beringia, in: *Encyclopedia of Quaternary*
599 *Science*, 2nd edn., edited by: Elias, S., Mock, C., and Murton, J., Elsevier, Amsterdam, 542–
600 552, 2013.
- 601 Seifert, W. K. and Moldovan, J. M.: Paleoreconstruction by biological markers, *Geochim*
602 *Cosmochim Acta*, 45, 783–794, 1981.
- 603 Shcherbakova, V., Rivkina, E., Pecheritsyna, S., Laurinavichius, K., Suzina, N., and
604 Gilichinsky, D.: *Methanobacterium arcticum* sp. nov., a methanogenic archaeon from
605 Holocene Arctic permafrost, *Int J Syst Evol Microbiol*, 61, 144–147, 2011.
- 606 Sher, A., Kaplina, T., and Ovander, M.: Unified regional stratigraphic chart for the
607 Quaternary deposits in the Yana-Kolyma Lowland and its mountainous surroundings,
608 Explanatory Note, Decisions of Interdepartmental Stratigraphic Conference on the
609 Quaternary of the Eastern USSR, Magadan, 1982, USSR Academy of Sciences, Far-Eastern
610 Branch, North-Eastern Complex Research Institute, Magadan, USSR, 29–69, 1987 (in
611 Russian).
- 612 Shi, T., Reeves, R. H., Gilichinsky, D. A., and Friedmann, E. I.: Characterization of viable
613 bacteria from Siberian permafrost by 16S rDNA sequencing, *Microb Ecol*, 33, 169–179,
614 1997.
- 615 Simon, C., Wiezer, A., Strittmatter, A. W., and Daniel, R.: Phylogenetic diversity and
616 metabolic potential revealed in a glacier ice metagenome, *Appl Environ Microb*, 75, 7519–
617 7526, 2009.

- 618 Stepanova, A. Y., Taldenkova, E. E., and Bauch, H. A.: Arctic quaternary ostracods and
619 their use in paleoreconstructions, *Paleontol J*, 44, 41–48, 2010.
- 620 Storey, J. D. and Tibshirani, R.: Statistical significance for genome wide studies, *Proc Nat*
621 *Acad Sci USA*, 100, 9440–9445, 2003.
- 622 Storey, J. D., Taylor, J. E., and Siegmund, D.: Strong control, conservative point estimation
623 and simultaneous conservative consistency of false discovery rates: a unified approach, *J*
624 *Royal Stat Soc B*, 66, 187–205, 2004.
- 625 Tomirdiaro, S. V. and Chernen'k'ii, B. I.: Cryogenic Deposits of East Arctic and Sub Arctic,
626 ANSSSR Far-East-Science Center, Magadan, 1987.
- 627 Tomirdiaro, S. V., Arslanov, K. A., Chernen'kii, B. I., Tertychnaya, T. V., and Prokhorova,
628 T. N.: New data on formation of loess-ice sequences in Northern Yakutia and ecological
629 conditionsof mammoth fauna in the Arctic during the late Pleistocene, *Reports Academy of*
630 *Sciences USSR* 278, 1446–1449, 1984 (in Russian).
- 631 Villemur, R., Lanthier, M., Beaudet, R., and Lepine, F.: The *Desulfitobacterium* genus,
632 *FEMS Microbiol Rev*, 30, 706–733, 2006.
- 633 Vishnivetskaya, T. A., Layton, A. C., Lau, M. C. Y., Chauhan, A., Cheng, K. R., Meyers, A.
634 J., Murphy, J. R., Rogers, A.W., Saarunya, G. S., Williams, D. E., Pfi ner, S. M., Biggersta ,
635 J. P.,Stackhouse, B. T., Phelps, T. J., Whyte, L., Sayler G. S., and Onstott, T. C.:
636 Commercial DNA extraction kits impact observed microbial community composition in
637 permafrost samples, *FEMS Microbiol Ecol*, 87, 217–230, 2014.
- 638 Walter, K. M., Edwards, M. E., Grosse, G., Zimov, S. A., and Chapin, F. S.: Thermokarst
639 lakes as a source of atmospheric CH₄ during the last deglaciation, *Science*, 318, 633–636,
640 2007.
- 641 Wilke, A., Harrison, T., Wilkening, J., Field, D., Glass, E. M., Kyrpides, N., Mavrommatis,
642 K.,and Meyer, F.: The M5nr: a novel non-redundant database containing protein sequences
643 and annotations from multiple sources and associated tools, *BMC Bioinformatics*, 13,
644 141,doi:10.1186/1471-2105-13-141, 2012.
- 645 Yergeau, E., Hogues, H., Whyte, L. G., and Greer, C. W.: The functional potential of high
646 Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses,
647 *ISME J*, 4, 1–9, 2010.
- 648 Zimov, S. A., Schuur, E. A. G., and Chapin, F. S.: Permafrost and the global carbon budget,
649 *Science*, 312, 1612–1613, 2006.

650 Table 1. Physicochemical characteristic of two permafrost soils
651
652

Samples	Grain-size distribution			Ice content	Eh	C _{tot}	Water extract (1:5)								
	Sand	Silt	Clay				Dry residue	pH	HCO ₃ ⁻	Cl ⁻	SO ₄ ²⁻	Ca ²⁺	Mg ²⁺	K ⁺	Na ⁺
	%						%	mV	%	%	Meq/100g dry soil				
IC4	79.44	15.84	4.72	17.5	45	1.1	0.130	7.22	0.12	1.40	0.14	0.32	0.42	0.06	1.33
IC8	40.72	52.72	6.56	27.6	167	1.2	0.135	8.16	0.57	0.32	0.48	0.75	0.35	0.02	0.56

653

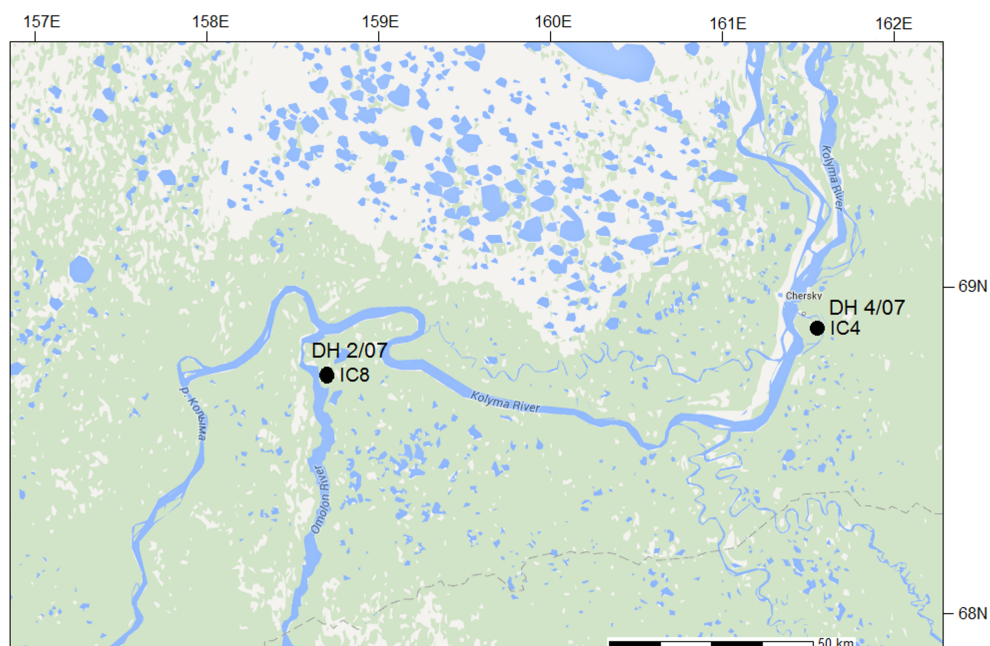
654 **Table 2.** Number of sequences showing homologies to genes associated with KEGG pathways
655 within the IC4 and IC8 metagenomes. Data were generated by searching in Hierarchical
656 classification table based upon SubSystems annotation source.
657

Enzyme Nomenclature	Function and Enzyme Profile	IC4	IC8
	<i>Methane metabolism</i>		
EC 1.5.99.11	F420-dependent methylene-H ₄ MPT reductase	3649	0
EC 1.12.98.1	Coenzyme F420 hydrogenase	1441	370
EC 1.8.98.1	CoB-CoM heterodisulfide reductase	24601	3982
EC 1.2.99.5	Formylmethanofuran dehydrogenase	10757	3012
EC 1.5.99.9	Methylenetetrahydromethanopterin dehydrogenases	370	3
EC 2.8.4.1	Methyl-coenzyme M reductase	1333	9
EC 2.3.1.101	Formylmethanofuran--tetrahydromethanopterin N-formyltransferase	3056	1642
EC 2.1.1.86	Tetrahydromethanopterin S-methyltransferase	1997	9
	<i>Nitrogen metabolism</i>		
EC 1.18.6.1	Nitrogenase (molybdenum-iron) alpha chain nifH	2519	42
EC 1.7.2.1	Copper-containing nitrite reductase	1824	11405
EC 1.7.2.5 ^a	Nitric-oxide reductase	3440	1838
EC 1.7.1.4	Nitrite reductase [NAD(P)H] small subunit	20713	13041

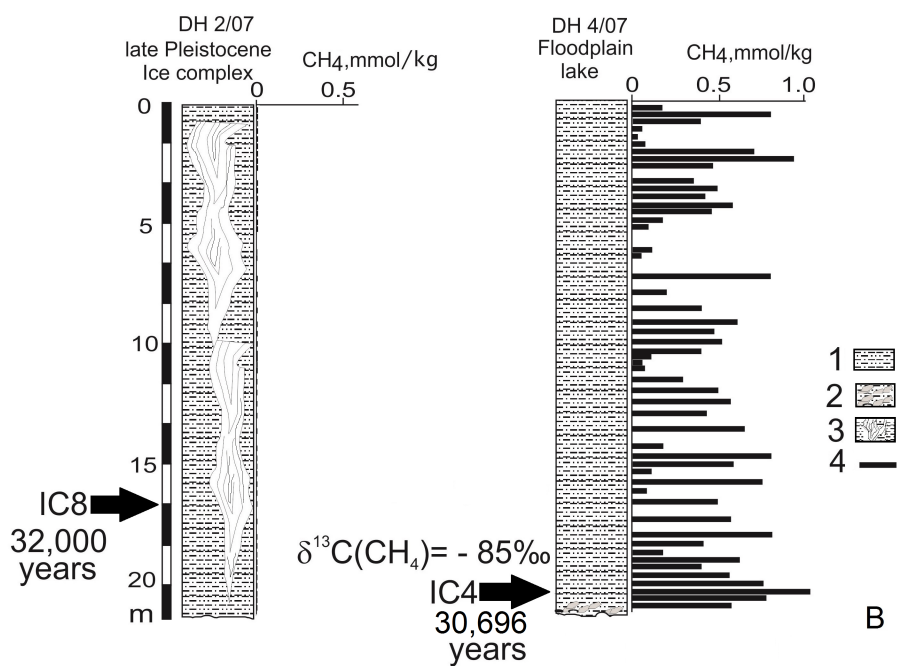
EC 1.7.99.4	Respiratory nitrate reductase alpha chain narG	36476	28169
EC	Ammonia monooxygenase	509	324
EC 1.7.3.4	Hydroxylamine oxidase	517	103
EC 1.7.99.6	Nitrous oxide reductase	61	46
	<i>Sulfur metabolism</i>		
EC 1.8.4.8	Phosphoadenylyl-sulfate reductase [thioredoxin]	7521	11165
EC 1.8.7.1	Ferredoxin--sulfite reductase	2587	8920
EC 2.7.7.4	Sulfate adenylyltransferase subunit 2	15449	23435
EC 2.7.1.25	Adenylylsulfate kinase	22898	22418
EC 1.8.3.1	Sulfite oxidase	14353	13968
	<i>Stress response</i>		
EC 1.11.1.6	Catalase	74622	41693
EC 1.15.1.1	Manganese superoxide dismutase	8922	13302
EC 1.11.1.7	Peroxidase	35552	19698
EC 1.15.1.1	Superoxide dismutase [Cu-Zn] precursor	2855	1327
EC 1.15.1.1	Superoxide dismutase [Fe]	7221	3109
EC 1.1.99.1	Choline dehydrogenase	25244	18018
EC 1.2.1.8	Betaine-aldehyde dehydrogenase	10022	12782
NR	Chaperone protein DnaK	78365	89944
NR	Interacting protein DnaJ	32897	45343

658

659 ^aThe current KEGG number has been created for enzyme EC 1.7.99.7.



A



B

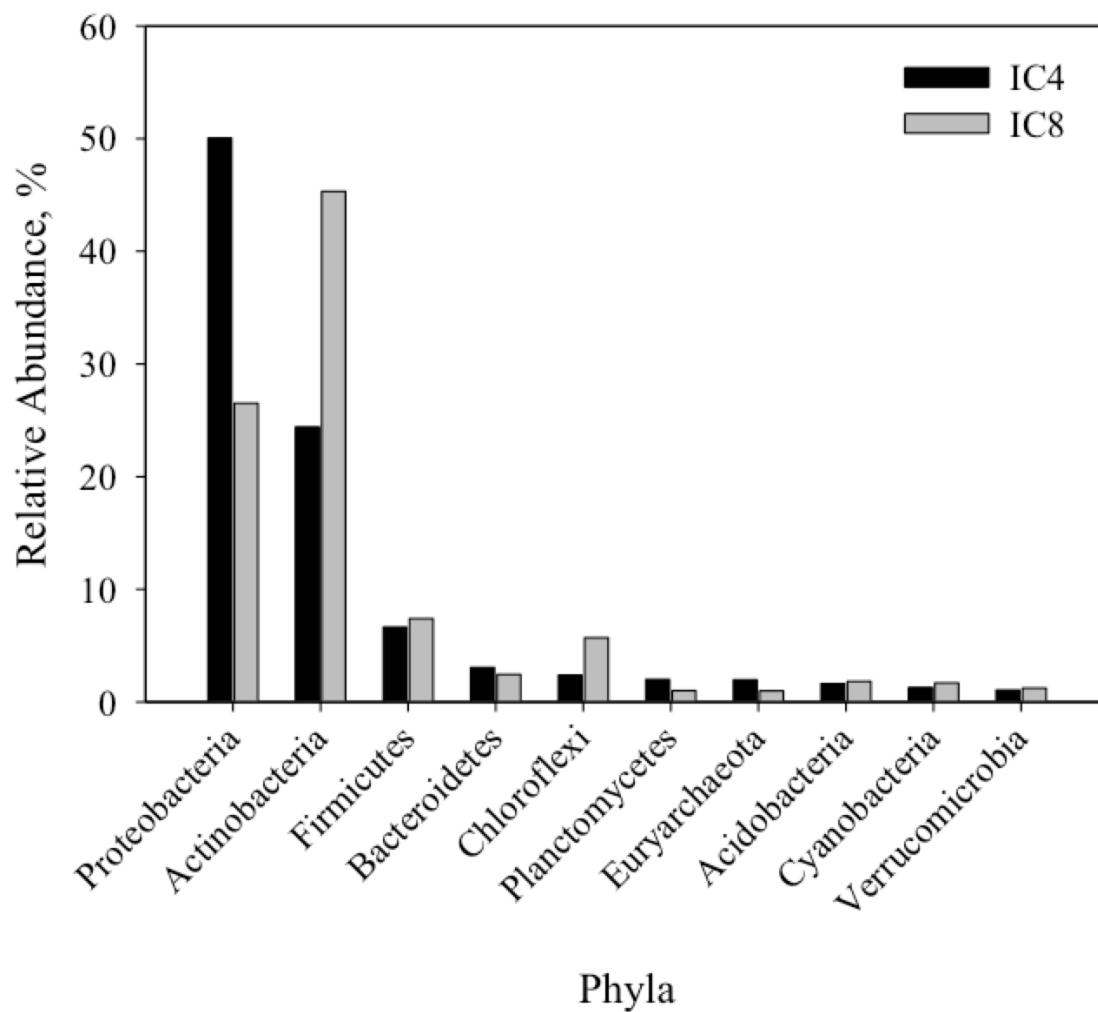
660

661 Fig. 1. Location of sampling sites on the Kolyma lowland (A). The position of samples IC4 and IC8
 662 in drilling holes and methane content from different depth of these boreholes (B): 1- sandy-loam,
 663 2 - sand with gravel, 3 - ice wedges, 4 - methane content

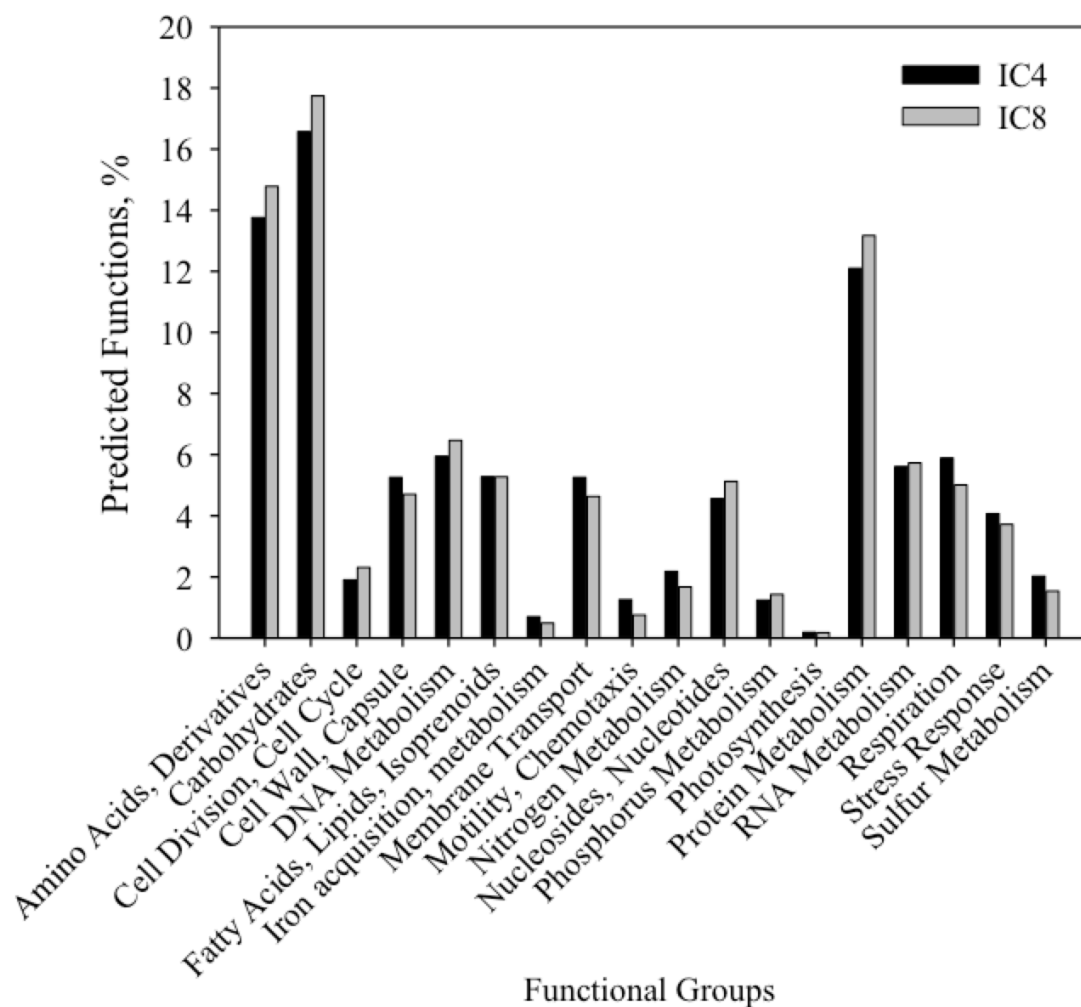
664

665

666

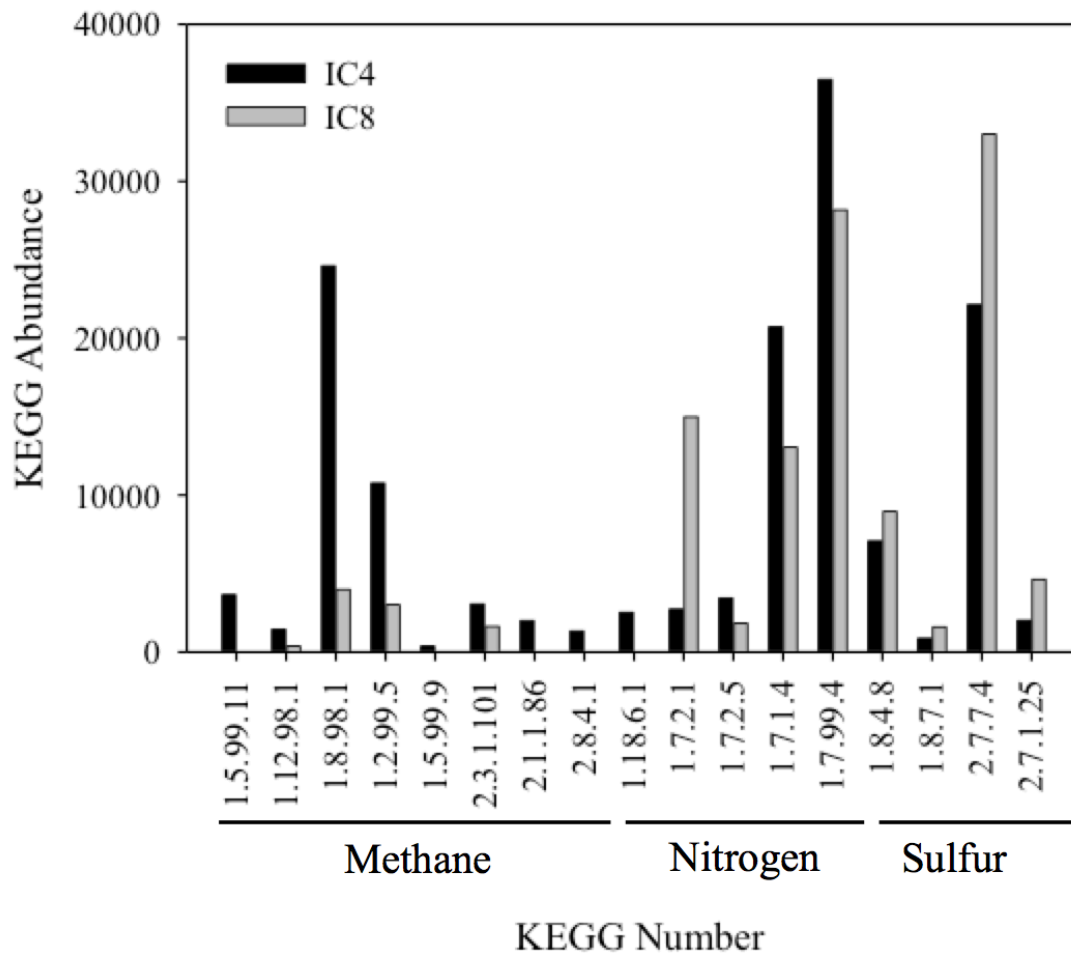


667
 668 Fig. 2. Community analyses of the IC4 and IC8 metagenomes at phylum level. The phyla present in
 669 metagenomes at a >1% level are shown.



670

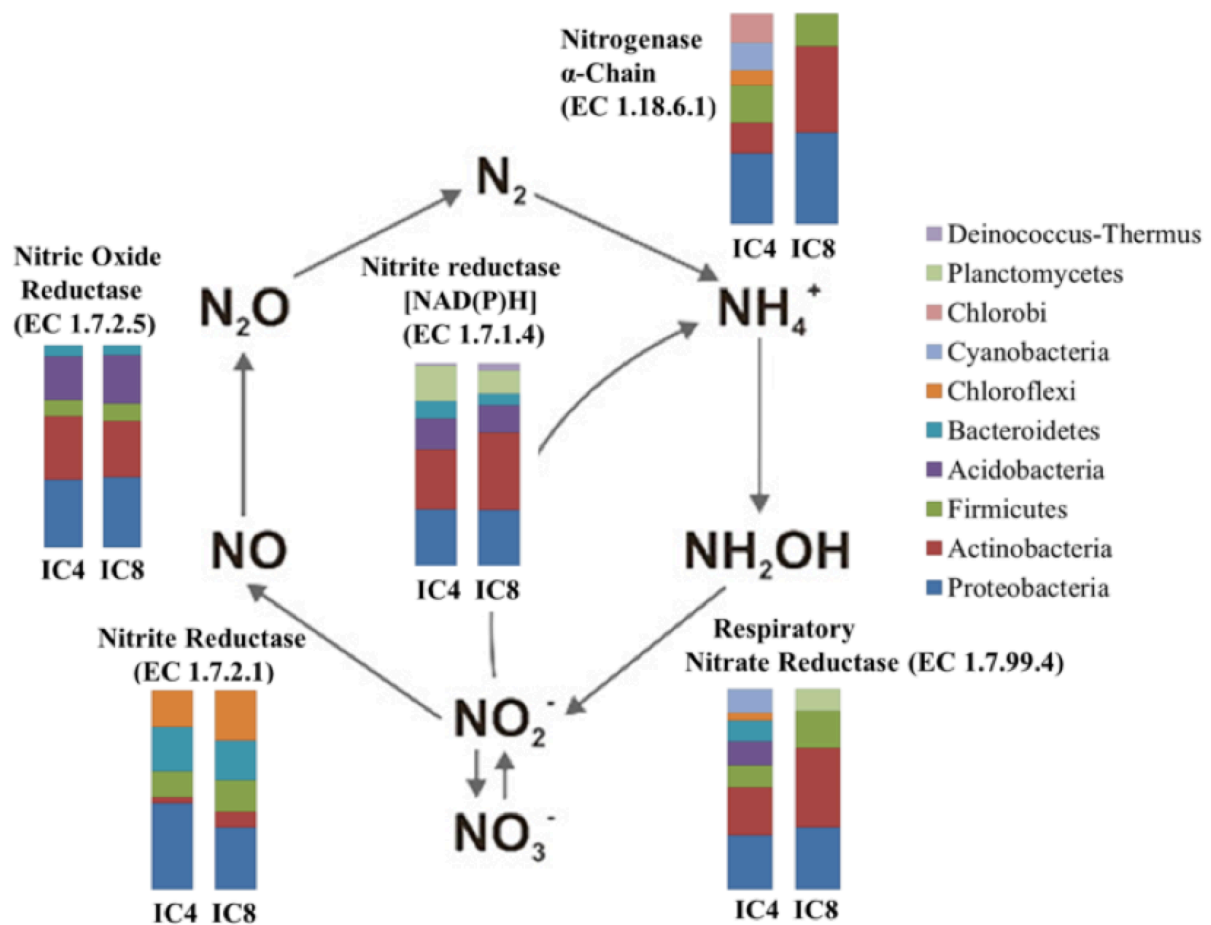
671 Fig. 3. Percentage of gene sequences associated with different functions in the annotated protein
 672 sequences within IC4 and IC8 metagenomes.



673

674 Fig. 4. KEGG abundance of the selected functional gene sequences found within IC4 and IC8
 675 metagenomes. Genes found in low abundance were not included in the figure. Genes from methane,
 676 nitrogen, and sulfur metabolic pathways are underlined. The enzyme name that corresponds to each
 677 KEGG number is given in Table 2.

678



679

680

681 Fig 5. Phylogenetic distribution of the sequences related to nitrogen metabolism within the IC4 and

682 IC8 metagenomes.

683

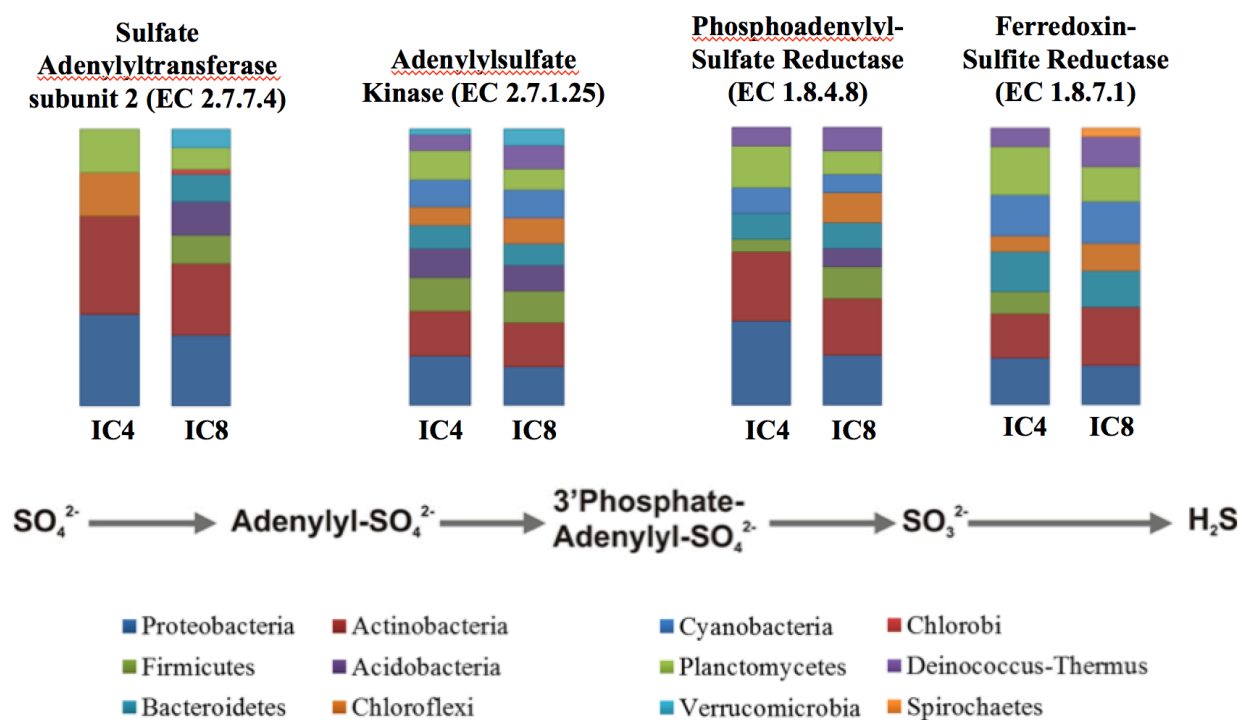


Fig. 6. Phylogenetic distribution of the sequences related to sulfur metabolism (sulfur reduction) within the IC4 and IC8 metagenomes.