



# Metagenomes from Late Pleistocene Ice Complex Sediments of the Siberian Arctic

 Tatiana Vishnivetskaya,<sup>a,b</sup> Elena Spirina,<sup>a</sup> Lyubov Shmakova,<sup>a</sup> Maria Tutukina,<sup>c</sup> Zhou Li,<sup>b,d</sup> Xiaofen Wu,<sup>b</sup> Elizaveta Rivkina<sup>a</sup>

<sup>a</sup>Institute of Physicochemical and Biological Problems in Soil Science, Russian Academy of Sciences, Pushchino, Russia

<sup>b</sup>University of Tennessee, Knoxville, Tennessee, USA

<sup>c</sup>Institute of Cell Biophysics, Russian Academy of Sciences, Pushchino, Russia

<sup>d</sup>Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

**ABSTRACT** The late Pleistocene Ice Complex (also known as Yedoma) encompasses ice-rich permafrost formed when alluvial and/or aeolian sediments accumulated under cold climatic conditions. Three metagenomes obtained from Yedoma deposits continually frozen for periods up to 60,000 years are reported here.

The late Pleistocene Ice Complex contains massive polygonal ice wedges and segregated ice (1) that formed during syngenetic deposition when freezing and ground ice accumulation occurred simultaneously with sedimentation (2). The Yedoma sediments are widely distributed in northeastern Siberia (3), representing important archives of paleoenvironmental conditions (4), and are vulnerable to thawing (5). These deposits contain about  $10^{15}$  kg of organic carbon, with 10% of this carbon being decomposable (2). A release of this carbon due to microbial decomposition upon permafrost thaw may increase greenhouse gas (GHG) emissions to the atmosphere (5, 6). However, there remain substantial uncertainties associated with the microbial component of the Yedoma permafrost environment.

Yedoma deposits located on the Bykovskiy Peninsula (BP) and Omolon River (OR) mouth were frozen under cold, arid climatic conditions about 50,000 to 60,000 and 25,000 to 28,000 years ago (3, 4), respectively, while Yedoma deposits from the Gydanskiy Peninsula (GP) represent slightly saline, silty loams with peat intercalations and a permafrost age of 34,300 years (7). Samples collected from the BP (at a depth of 24.3 m) and the OR (at a depth of 26.5 m) had  $\sim 1.1\%$  total organic carbon, and methane was not detected (3, 4). On the other hand, the GP sample (from a depth of 4.5 m) had  $<1\%$  total organic carbon and biogenic methane in concentrations up to  $0.4 \text{ mmol kg}^{-1}$  (7). Permafrost cores were obtained using a portable gasoline-powered drill that operates without a drilling fluid so as to ensure aseptic sampling (8). A peripheral portion of the extracted permafrost cores was removed with a sterile knife (9), and the inner uncontaminated frozen part of the cores was immediately sectioned, placed in Whirl-Pak bags, transported to a mobile  $-20^\circ\text{C}$  freezer (Waeco CoolFreeze CF-50), and stored frozen until performing genetic analyses. The total community genomic DNA was isolated from  $\sim 5$  g of frozen permafrost sample using the PowerSoil kit (Mo Bio Laboratories, Qiagen Company, USA), followed by the Genomic DNA Clean & Concentrator kit (Zymo Research, USA). Metagenome sequencing was performed at the Centre for Genomic Regulation sequencing facility (Barcelona, Spain), employing the Illumina TruSeq SBS kit (v.3) and the Illumina HiSeq 2000 machine with a  $2 \times 250$ -cycle sequencing protocol. Metagenome analysis was carried out by first checking the quality of paired reads with Illumina FastQC v.0.11.5 (10) before removing adapters and low-quality reads using Trimmomatic v.0.36 (11), with options LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. Taxonomic classification on trimmed reads was per-

**Citation** Vishnivetskaya T, Spirina E, Shmakova L, Tutukina M, Li Z, Wu X, Rivkina E. 2019. Metagenomes from late Pleistocene Ice Complex sediments of the Siberian Arctic. *Microbiol Resour Announc* 8:e01010-19. <https://doi.org/10.1128/MRA.01010-19>.

**Editor** Frank J. Stewart, Georgia Institute of Technology

**Copyright** © 2019 Vishnivetskaya et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Tatiana Vishnivetskaya, [tvishniv@utk.edu](mailto:tvishniv@utk.edu), or Elizaveta Rivkina, [elizaveta.rivkina@gmail.com](mailto:elizaveta.rivkina@gmail.com).

**Received** 20 August 2019

**Accepted** 26 September 2019

**Published** 24 October 2019

formed with a novel metagenome classifier, Kaiju v.1.7.2 (12), to find matches on the protein level. The reads were compared to the NCBI nr+euk database in greedy mode (minimum match length, 15; minimum match score, 75; and allowed mismatches, 3) (12). Trimmed reads were also individually assembled using MEGAHIT v.1.0.3, with k-mer sizes of 41, 51, 61, 71, 81, and 91 (13). The subsequent assembled contigs were evaluated using QUAST v.5.0.2 (14). Prodigal v.2.6.3 (15) was used to find coding genes on contigs of  $\geq 1,000$  bp with a meta flag. The amino acid sequences from predicted genes were subsequently uploaded to GhostKOALA (<https://www.kegg.jp/ghostkoala/>) for KEGG annotation using the genus\_prokaryotes and family\_eukaryotes databases (16). After the preprocessing of raw paired reads, we obtained 56 to 77 million trimmed reads per sample, which were assembled into 11,143 to 207,624 contigs of  $\geq 1,000$  bp in length, representing 3.81% to 9.42% mapped reads to  $\geq 1,000$ -bp contigs.

Based upon the taxonomic analysis on trimmed reads, *Actinobacteria* was the most abundant phylum in both the BP (22.4% of reads) and the OR (36.3%) samples, whereas *Proteobacteria* dominated in the GP (17.7%) samples. Archaea were more abundant in the GP samples (3.4% of reads) than in the BP (0.5%) and the OR (1.6%) samples. Of all reads assigned to Archaea in the GP samples, 61% belonged to Euryarchaeota, with which many methanogens are taxonomically affiliated. Genes encoding methyl-coenzyme M reductase (based upon the presence of *mcrABG* genes) that can catalyze the final step of methanogenesis were identified only within the GP samples, which was consistent with the presence of biogenic methane (17). While metagenomes from both the BP and OR samples contained genes assigned to methane/ammonia monooxygenase (based upon the presence of the *pmoABC-amoABC* genes), suggesting that microorganisms in these two locations have the potential to oxidize methane to methanol and/or ammonia to hydroxylamine. In addition, enzymes such as endoglucanase (based upon the presence of Enzyme Commission number [EC] 3.2.1.4) and xylan 1,4-beta-xylosidase (based upon the presence of the *xynB* gene) involved in the degradation of complex polysaccharides were identified in all three metagenomes.

Organic-bearing ice-rich Yedoma sediments are vulnerable to environmental changes, including thawing, and these sediments might be considered a potential GHG source in permafrost. A comprehensive investigation of Yedoma metagenomes could provide a basis for more reliable predictions of future biochemical reactions within these sediments in response to climate change.

**Data availability.** The raw sequences for these metagenomes were deposited to the NCBI Sequence Read Archive (SRA) under the accession numbers [SRX2163490](#), [SRX2163491](#), and [SRX2163492](#).

## ACKNOWLEDGMENTS

This study was supported by the Russian Federal Target Program under grant 0191-2019-0044, the Russian Foundation for Basic Research (RFBR) under grants 19-04-01240 and 19-29-05003mk, and the National Science Foundation under grant DEB-1442262.

We are grateful to Arthur Meyers for the English proofreading.

## REFERENCES

- Schirmer L, Froese D, Tumskey V, Grosse G, Wetterich S. 2013. Yedoma: late Pleistocene ice-rich syngenetic permafrost of Beringia, p 542–552. In Elias S, Mock C, Murton J (ed), *Encyclopedia of quaternary science*, 2nd ed. Elsevier, Amsterdam, the Netherlands.
- Strauss J, Schirmer L, Grosse G, Fortier D, Hugelius G, Knoblach C, Romanovsky V, Schädel C, Schneider von Deimling T, Schuur EAG, Shmelev D, Ulrich M, Veremeeva A. 2017. Deep Yedoma permafrost: a synthesis of depositional characteristics and carbon vulnerability. *Earth Sci Rev* 172:75–86. <https://doi.org/10.1016/j.earscirev.2017.07.007>.
- Schirmer L, Kunitsky V, Grosse G, Wetterich S, Meyer H, Schwamborn G, Babiy O, Derevyagin A, Siegert C. 2011. Sedimentary characteristics and origin of the late Pleistocene Ice Complex on north-east Siberian Arctic coastal lowlands and islands—a review. *Quat Int* 241:3–25. <https://doi.org/10.1016/j.quaint.2010.04.004>.
- Rivkina E, Petrovskaya L, Vishnivetskaya T, Krivushin K, Shmakova L, Tutukina M, Meyers A, Kondrashov F. 2016. Metagenomic analyses of the late Pleistocene permafrost—additional tools for reconstruction of environmental conditions. *Biogeosciences* 13:2207–2219. <https://doi.org/10.5194/bg-13-2207-2016>.
- Turetsky MR, Abbott BW, Jones MC, Anthony KW, Olefeldt D, Schuur EAG, Koven C, McGuire AD, Grosse G, Kuhry P, Hugelius G, Lawrence DM, Gibson C, Sannel A. 2019. Permafrost collapse is accelerating carbon release. *Nature* 569:32–34. <https://doi.org/10.1038/d41586-019-01313-4>.
- Mackelprang R, Waldrop MP, DeAngelis KM, David MM, Chavarria KL, Blazewicz SJ, Rubin EM, Jansson JK. 2011. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* 480:368. <https://doi.org/10.1038/nature10576>.
- Demidov NE, Baranskaya AV, Durdenko EV, Zanina OG, Karaevskaia ES,

- Pushina ZV, Rivkina EM, Spirina EV, Spencer R. 2016. Biogeochemistry of the permanently frozen sediments on the Arctic coast of the Gydan Peninsula. *Probl Arct Antarct* 3:34–49. (In Russian.)
8. Shi T, Reeves RH, Gilichinsky DA, Friedmann EI. 1997. Characterization of viable bacteria from Siberian permafrost by 16S rDNA sequencing. *Microb Ecol* 33:169–179. <https://doi.org/10.1007/s002489900019>.
  9. Bang-Andreasen T, Schostag M, Prieme A, Elberling B, Jacobsen CS. 2017. Potential microbial contamination during sampling of permafrost soil assessed by tracers. *Sci Rep* 7:43338. <https://doi.org/10.1038/srep43338>.
  10. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
  11. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
  12. Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>.
  13. Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
  14. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34:i142–i150. <https://doi.org/10.1093/bioinformatics/bty266>.
  15. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
  16. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28:27–30. <https://doi.org/10.1093/nar/28.1.27>.
  17. Rivkina E, Shcherbakova V, Laurinavichius K, Petrovskaya L, Krivushin K, Kraev G, Pecheritsina S, Gilichinsky D. 2007. Biogeochemistry of methane and methanogenic archaea in permafrost. *FEMS Microbiol Ecol* 61:1–15. <https://doi.org/10.1111/j.1574-6941.2007.00315.x>.