ABSTRACT The role of archaeal ammonia oxidizers often exceeds that of bacterial ammonia oxidizers in marine and terrestrial environments but has been understudied in permafrost, where thawing has the potential to release ammonia. Here, three thaumarchaea genomes were assembled and annotated from metagenomic data sets from carbon-poor Canadian High Arctic active-layer cryosols.

Recent studies have shown that ammonia-oxidizing archaea (AOA) often outnumber ammonia-oxidizing bacteria (AOB) (1–3) and that AOA do the lion’s share of ammonia oxidation in terrestrial soils (3–9). However, their potential role in both fixed nitrogen losses and ozone-depleting nitric oxide or nitrous oxide release and their taxonomic diversity are poorly understood in terrestrial permafrost ecosystems (10–12). Here, we report three draft genome sequences of thaumarchaea that are potential chemolithoautotrophic ammonia oxidizers and may play a potentially significant role in the nitrogen cycle of the Canadian High Arctic.

These three genomes were constructed through the analysis of 21 metagenomes from mineral cryosols at 5-cm depth retrieved in our previously published studies (13, 14). Metagenomic libraries were prepared using the Illumina Nextera DNA library preparation kit (Illumina, Inc., San Diego, CA), followed by 100-bp paired-end DNA sequencing on an Illumina HiSeq 2000 platform (13). A total of 498,483,227 forward and reverse reads were filtered separately for quality using tools available on the Princeton University Galaxy server as follows: reads with 90% of the bases with a Phred score of 30 were removed using “filter by quality” v1.0.0; Nextera transposase adaptor sequences were trimmed using cutadapt v1.6; FASTQ Trimmer v1.0.0 was used to remove the last five bases at the 3′ end; and trimmed reads with fewer than 50 nucleotides (nt) were removed. One of the paired reads may be discarded, and the remaining read is referred to as a single read. Default parameters were used for all software unless otherwise specified.

Individual metagenomes were assembled using IDBA-UD v1.1.1 (15), and the scaffolds were sorted into taxonomic bins using MetaBAT v0.32.4 in “very specific” mode (16). The bins recovered from each of the 21 metagenomes were analyzed using CheckM v1.0.7 to assess the quality and taxonomy (17). The taxa appearing multiple times across the 21 metagenomes included 10 Nitrosopumilales bins within the phylum Thaumarchaeota (completeness, 3.88 to 97.73%; contamination, 0 to 2.02%). Of these, three Nitrosopumilales bins were 71.62 to 97.73% complete, so the quality reads from the three corresponding metagenomes (NCBI Sequence Read Archive accession num-

Citation Sun EW-H, Hajirezaie S, Dooner M, Vishnivetskaya TA, Layton A, Chauhan A, Pfiffner SM, Whyte LG, Onstott TC, Lau MCY. 2020. Thaumarchaea genome sequences from a High Arctic active layer. Microbiol Resour Announc 9:e00326-20. https://doi.org/10.1128/MRA.00326-20.}

Editor John J. Dennehy, Queens College
Copyright © 2020 Sun et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.
Address correspondence to Emily Wei-Hsin Sun, esun@princeton.edu, or Maggie C. Y. Lau, maglau@idsse.ac.cn.
Received 28 March 2020
Accepted 27 April 2020
Published 21 May 2020
bers SRR1586318, SRR1586310, and SRR1586268), which were mapped onto their respective genome bins using Bowtie 2 v2.3.2 (18), were extracted and reassembled separately using IDBA-UD v1.1.1. The number of mapped reads was used to calculate the depth of the coverage. The quality and taxonomy of the draft genomes were evaluated using CheckM v1.0.7. The statistics of these three Nitrosopumilales draft genome sequences are presented in Table 1.

Prodigal (Prokaryotic Dynamic Programming Genefinding Algorithm) v2.6.3 (19) was employed for gene prediction. Genes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline v4.10 (PGAP) (20), and the metabolic pathways were analyzed using the KEGG Automatic Annotation Server (KAAS) v2.1 (21). The results revealed genes belonging to the 3-hydroxypropanoate/4-hydroxybutanoate carbon fixation pathway, ammonia monooxygenase subunits A, B, and C (amoCAB gene), and nitrite reductase (nirK gene). Like in the ammonia-oxidizer Nitrosopumilus maritimus SCM1 (22), the gene encoding hydroxylamine oxidoreductase (HAO), the key enzyme in the conversion of hydroxylamine to nitrite, is absent in all three genomes. The enzyme 4-hydroxybutyryl-coenzyme A dehydratase (4-HCD) identified in the three thaumarchaeota genomes has been proposed as an environmental marker for autotrophic thaumarchaeota (23).

Data availability. This genome assembly project and the three high-completion genome sequences have been deposited at NCBI GenBank under the accession numbers WJXC00000000, WJXE00000000, and WJXD00000000 (BioSample numbers SAMN11973980, SAMN11973981, and SAMN11973982 and BioProject number PRJNAS48371). The versions described in this paper are the first versions, WJXC01000000, WJXD01000000, and WJXE01000000. The raw reads were deposited at the NCBI Sequence Read Archive under the accession number SRR047512 (13).

ACKNOWLEDGMENTS
The original sequencing data were supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science Program (DE-SC0004902) to S.M.P. This research was supported by a National Science Foundation Dimensions grant (DEB-1442059) to T.C.O.

### Table 1: Genome statistics for AHI_AL_Thaum01 to AHI_AL_Thaum03

| Characteristic                                      | Data for strain:             |
|-----------------------------------------------------|------------------------------|
|                                                     | AHI_AL_Thaum01 | AHI_AL_Thaum02 | AHI_AL_Thaum03 |
| Metagenome data set from which the draft genome originated (SRA accession no.) | SRR1586268 | SRR1586310 | SRR1586318 |
| No. of paired-end reads in the original data set | 150,218,712 | 34,785,144 | 38,551,334 |
| Quality-filtered reads available for the initial assembly | 24,966,724\(a\) (96) | 13,300,133 (93) | 18,935,562 (88) |
| No. of single reads (average read length [nt]) | 31,932,629\(a\) (96) | 9,420,119 (93) | 9,420,119 (89) |
| Quality-filtered reads extracted for reassembly | 132,263 (96) | 195,388 (93) | 184,721 (88) |
| Genome size (bp) | 1,442,500 | 2,538,969 | 1,925,816 |
| No. of contigs | 578 | 521 | 602 |
| GC content (%) | 39.6 | 39.3 | 39.3 |
| Minimum contig length (bp) | 1,000 | 1,000 | 1,003 |
| \(N_{50}\) (bp) | 2,965 | 6,780 | 3,871 |
| Completeness (%) | 62.73 | 98.71 | 84.47 |
| Contamination (%) | 2.18 | 2.43 | 2.99 |
| No. of predicted coding genes\(b\) | 1,495 | 2,587 | 2,014 |
| Mean base coverage (×) | 12.07 | 8.48 | 9.49 |

\(a\) Sample SRR1586268 was assembled using only the paired reads because the inclusion of single reads failed to run due to an IDBA-UD segmentation fault that was not resolved.

\(b\) Annotated using the NCBI Prokaryotic Genome Annotation Pipeline v4.10.
We thank the staff at Research Computing, Office of Information Technology, Princeton University, for their technical support with the computational analyses. We also thank the GEO523 2016 class for filtering the metagenomic reads on the Galaxy platform.

E.W.-H.S., S.H., M.D., and M.C.Y.L conceived the analysis. T.A.V. performed the total DNA extraction and submitted it to A.L. for sequencing. E.W.-H.S., S.H., and M.D. assembled the sequenced reads and performed mapping, binning, reassembly, gene prediction, and annotation with consultation from M.C.Y.L. E.W.-H.S., S.H., M.D., M.C.Y.L., and T.C.O. contributed to the interpretation of the data and production of the manuscript.

REFERENCES

1. Beman MJ, Francis CA. 2006. Diversity of ammonia-oxidizing archaea and bacteria in the sediments of a hypernitrified subtropical estuary: Bahia del Tóbari, Mexico. Appl Environ Microbiol 72:7767–7777. https://doi.org/10.1128/AEM.00946-06.

2. Wuchter C, Abbas B, Coolen MJL, Herfort L, van Bleijswijk J, Timmers P, Strous M, Teira E, Hermil GJ, Middelburg JJ, Schouten S, Sinninghe Damsté JS. 2006. Archaeal nitrification in the ocean. Proc Natl Acad Sci U S A 103:12317–12322. https://doi.org/10.1073/pnas.0600756103.

3. Leininger S, Urich T, Schlüter M, Schwark L, Qi J, Nicol GW, Prosser JI, Schuster SC, Schleper C. 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442:806–809. https://doi.org/10.1038/nature04983.

4. Stemphlhuber B, Engel M, Fischer D, Neskovick-Prit G, Wubet T, Schöning I, Gubry-Rangin C, Kubitšik S, Schlüter-Hai B, Kati N, Welzl G, Nicol GW, Schrumpf M, Buscot F, Prosser JI, Schlüter M. 2015. pH as a driver for ammonia-oxidizing archaea in forest soils. Microb Ecol 69:879–883. https://doi.org/10.1007/s00248-014-0548-5.

5. Offer P, Prosser JI, Nicol GW. 2009. Growth of ammonia-oxidizing archaea in soil microcosms is inhibited by acetylene. FEMS Microbiol Ecol 70:99–108. https://doi.org/10.1111/j.1574-6941.2009.00725.x.

6. Gubry-Rangin C, Nicol GW, Prosser JI. 2010. Archaea rather than bacteria control nitrification in two agricultural acidic soils. FEMS Microbiol Ecol 76:7626–7634.

7. Leininger S, Urich T, Schloter M, Schwark L, Qi J, Nicol GW, Prosser JI, Schuster SC, Schleper C. 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442:806–809. https://doi.org/10.1038/nature04983.

8. Stemphlhuber B, Engel M, Fischer D, Neskovick-Prit G, Wubet T, Schöning I, Gubry-Rangin C, Kubitšik S, Schlüter-Hai B, Kati N, Welzl G, Nicol GW, Schrumpf M, Buscot F, Prosser JI, Schlüter M. 2015. pH as a driver for ammonia-oxidizing archaea in forest soils. Microb Ecol 69:879–883. https://doi.org/10.1007/s00248-014-0548-5.

9. Prosser JI, Nicol GW. 2012. Archaeal and bacterial ammonia-oxidisers in soil: the quest for niche specialization and differentiation. Trends Microbiol 20:523–531. https://doi.org/10.1016/j.tim.2012.08.001.

10. Alves RE, Wanek W, Zappe A, Richter A, Svenning MM, Schleper C, Urich T. 2013. Nitrification rates in Arctic soils are associated with functionally distinct populations of ammonia-oxidizing archaea. ISME J 7:1620–1631. https://doi.org/10.1038/ismej2012.45.

11. Daebeler A, Bodelier PLE, Yan Z, Hefting MM, Jia Z, Laanbroek HJ. 2014. Interactions between Thaumarchaea, Nitrosira and methanotrophs modulate autotrophic nitrification in volcanic grassland soil. ISME J 8:2397–2410. https://doi.org/10.1038/ismej2014.81.

12. Ermel M, Behrendt T, Oswald R, Derstroff B, Wu D, Hohlmann S, Stößer C, Pommerehn-Röser A, Konecke M, Williams J, Meixner FX, Andreæ MO, Trebs I, Sörgel M. 2018. Hydroxylamine released by nitrifying microorganisms is a precursor for HONO emission from drying soils. Sci Rep 8:1877. https://doi.org/10.1038/s41598-018-20170-1.

13. Chauhan A, Layton AC, Vishnevskaya TA, Williams D, Pfiffner SM, Rekepalli B, Stackhouse B, Lau MCY, Phelps TJ, Mykytczuk N, Ronholm J, Whyte L, Onstott TC, Sayler GS. 2014. Metagenomes from thawing low-soil-organic-carbon mineral cryosols and permafrost of the Canadian High Arctic. Genome Announcement 2:e01217-14. https://doi.org/10.1128/genomeA.01217-14.

14. Stackhouse BT, Vishnevskaya TA, Layton A, Chauhan A, Pfiffner S, Mykytczuk NC, Sanders R, Whyte LG, Hedin L, Saad N, Myreni S, Onstott TC. 2015. Effects of simulated spring thaw of permafrost from mineral cryosol on CO2 emissions and atmospheric CH4 uptake. J Geophys Res Biogeosci 120:1764–1784. https://doi.org/10.1002/2015JG003004.

15. Peng Y, Leung HCM, Yiu SM, Chin FYL. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428. https://doi.org/10.1093/bioinformatics/bts174.

16. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3:e1165. https://doi.org/10.7717/peerj.1165.

17. Imelfort M, Skennerton CT, Parks DH, Tyson GW, Hugenholtz P. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.

18. Langmued B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

19. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:19. https://doi.org/10.1186/1471-2105-11-119.

20. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Ostell J. 2016. NCBI taxonomy database project. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1038/nrmicro2365.

21. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 40:109–114.

22. Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, Ronholm J, Trebs I, Sörgel M. 2018. Hydroxylamine released by nitrifying microorganisms is a precursor for HONO emission from drying soils. Sci Rep 8:1877. https://doi.org/10.1038/s41598-018-20170-1.

23. Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, Alber BE, Fuchs G. 2010. Autotrophic carbon fixation in archaea. Nat Rev Microbiol 8:447–460. https://doi.org/10.1038/nrmicro2365.