AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome of *Sphingomonas aerolata* PDD-32b-11, Isolated from Cloud Water at the Summit of Puy de Dôme, France

Domitille Jarrige,[a] Thierry Nadalig,[a] Muriel Joly,[b] Martine Sancelme,[b] Stéphane Vuilleumier,[a] Pierre Amato,[b] Françoise Bringel[a]

[a]Génétique Moléculaire, Génomique, Microbiologie (GMGM), Université de Strasbourg, Strasbourg, France

[b]Université Clermont Auvergne, Clermont Auvergne Institut National Polytechnique (INP), Centre National de la Recherche Scientifique (CNRS), Institut de Chimie de Clermont-Ferrand (ICCF), Clermont-Ferrand, France

ABSTRACT The complete genome of *Sphingomonas aerolata* PDD-32b-11, a bacterium isolated from cloud water, was sequenced. It features four circular replicons, a chromosome of 3.99 Mbp, and three plasmids. Two putative rhodopsin-encoding genes were detected which might act as proton pumps to harvest light energy.

Metabolically active microorganisms in clouds participate in atmospheric chemistry (1, 2). Genome analysis of atmospheric bacteria may help identify key functions involved in the maintenance of activity under harsh conditions (e.g., low nutrient and water levels, rapid variations in temperature, and UV exposure), and their impact on atmospheric functioning. Members of the *Sphingomonas* genus (3) have been isolated from air in both natural and anthropized environments (4–7). Representatives were regularly isolated from clouds sampled at the summit of puy de Dôme (1465 m above sea level) (8). Among those, strain PDD-32b-11 was isolated in 2009 (8).

The complete genome of strain PDD-32b-11 was sequenced by single-molecule real-time long reads using a PacBio Sequel II sequencer (Pacific Biosciences, Menlo Park, CA, USA; Gentyane sequencing platform, Clermont-Ferrand, France). Genomic DNA was extracted from an aerobic culture grown at 17°C in R2A medium (Millipore) using the MasterPureTM

TABLE 1 *Sphingomonas aerolata* PDD-32b-11 genome assembly and annotation statistics

Feature	*Sphingomonas aerolata* PDD-32b-11
Sequencing coverage depth	119×
Assembly size (Mbp)	4.21
No. of scaffolds/contigs	4
GC content (%)	66.09
Chromosome contig (bp)	3,991,634
Plasmids contigs (bp)	38,945; 77,112; 102,246
Annotation statistics	
No. of predicted CDS[a]	4,181
No. of predicted tRNAs	57 (chromosome-encoded)
No. of predicted rRNAs	12 (chromosome-encoded)
Average CDS length (bp)	919.67
BUSCO v5.2.2 assembly completeness assessment (%)	
Complete	99.5
Single	99.3
Duplicated	0.2
Fragmented	0.2
Missing	0.3

[a] CDS, Coding sequence.

Editor J. Cameron Thrash, University of Southern California

Address correspondence to Françoise Bringel, francoise.bringel@unistra.fr.

The authors declare no conflict of interest.

Received 7 July 2022
Accepted 5 September 2022
Published 15 September 2022

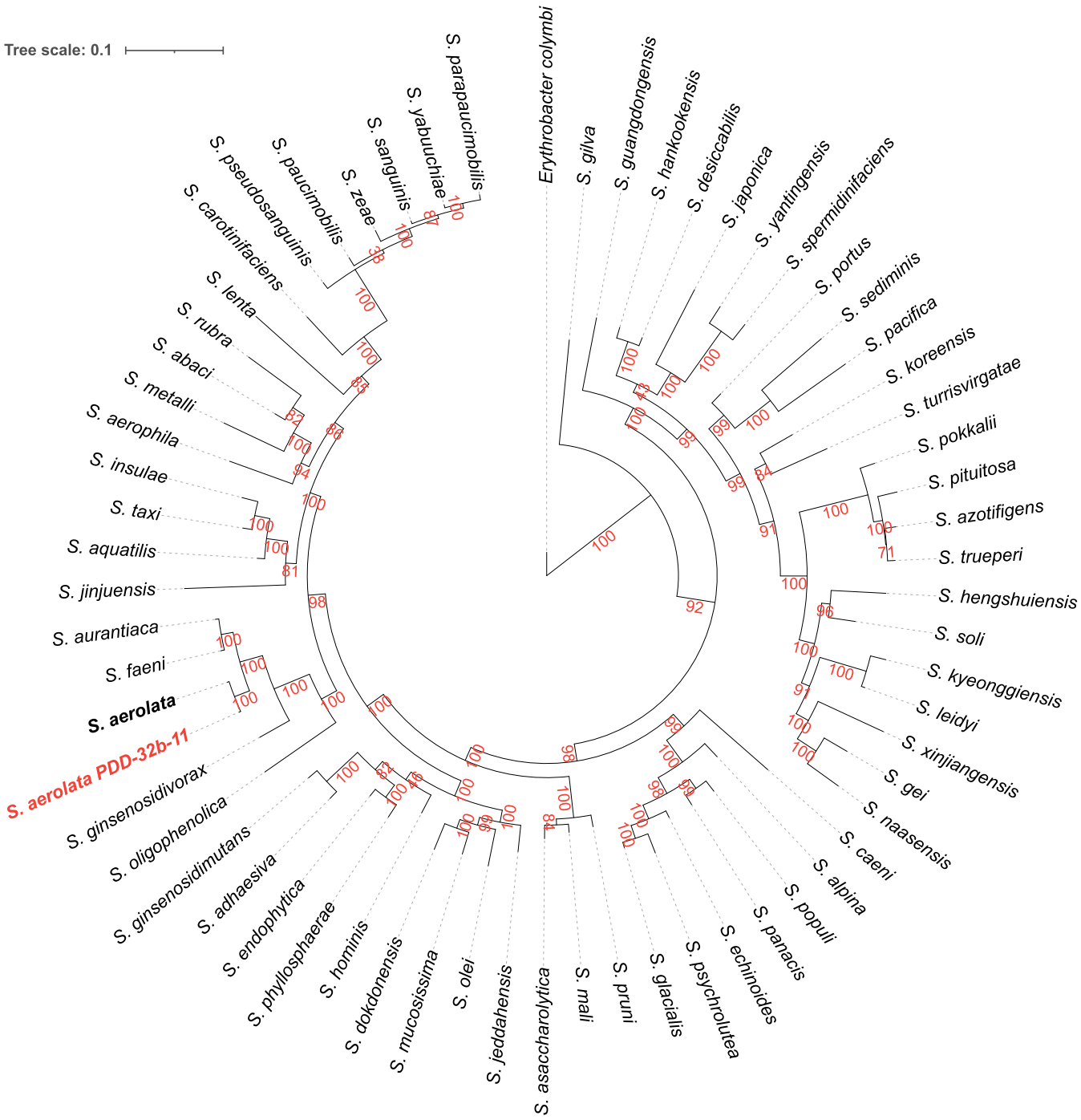

FIG 1 Phylogenetic tree of the *Sphingomonas* genus. The approximately maximum likelihood tree was built using GTDB-Tk (11) with 118 marker genes, pruned with ETE v3 (16), and visualized with ITOL v 6.5.8 (17). *Erythrobacter colymbi* was chosen as the outgroup. Bootstrap support values in percent are indicated in red.

complete DNA & RNA purification kit (Lucigen). A SMRTbell express 2 template prep kit was used for library preparation from sheared DNA fragments of approximately 10 kb on average. A SMRTBell polymerase complex was obtained using Binding kit 2.2 (Pacific Biosciences) and primer v5 was sequenced using a PacBio SMRTcell 8 M and a sequencing plate 2.0. Default parameters were used with all software unless otherwise specified. Raw sequencing data (CCS reads) were treated using the library smrttools10.1.0.119588 (Pacific Biosciences) (3 minimum passes for "ccs" tool; split-bam-named, ccs and peek-guess for "lima" tool). The resulting 70391 reads (mean read length 7119 nt) were assembled with Flye v2.9 (9)

(parameters: pacbio-hifi, genome-size 5 Mbp), yielding a genome of 4,209,937 bp in total, with four circular contigs (Table 1). The chromosome (3.99 Mbp) harbors all rRNA and tRNA encoding genes. The three other contigs (39, 77, and 102 kb) have a lower GC content (59 to 62%) compared to the chromosome (66%) and encode replication initiator proteins typical of plasmid replication. Genome assembly completeness was assessed with BUSCO v5.2.2 (10) (reference data set sphingomonadales_odb10 2021-02-23; 124 genomes; 1018 BUSCOs). Strain PDD-32b-11 was assigned to the species *Sphingomonas aerolata* using GTDB-Tk v2.0.0 (GTDB r207) (11) (Fig. 1).

The obtained genome assembly was annotated using the NCBI Prokaryotic Genome Annotation Pipeline v6.1 (12). Additional putative genes were annotated on the MicroScope platform v3.15.4 (13). Genes were identified for both tetrahydrofolate- (FolD and Fhs) and glutathione-dependent (FrmA and FrmC) pathways for oxidation of formaldehyde, a key compound of atmospheric chemistry (2). A partial photosynthesis gene cluster (PGC) (*bchI*, *bchD*, *bchO*, *bchF*, *bchG*, *ppaA*, *ppsR*, and *tspO*) typical of aerobic anoxygenic phototrophs (14) was also found. In addition, two putative homologs of the functional proton-pumping DTG/DTS rhodopsin of *Pseudomonas putida* (15) were identified, each adjacent to a gene encoding beta-carotene 15,15′-dioxygenases associated with biosynthesis of the retinal cofactor of light-driven rhodopsin proton pumps.

Data availability. The *Sphingomonas aerolata* PDD-32b-11 complete genome project was deposited in DDBJ/ENA/GenBank under accession numbers CP098762 to CP098765 within BioProject PRJNA847404. Raw reads were deposited in the Sequence Read Archive under accession number SRR19594883. Genome sequence and annotation are accessible on the MicroScope platform (https://mage.genoscope.cns.fr/microscope/genomic/overview.php?O_id=17210).

ACKNOWLEDGMENTS

We thank Amandine Moreno for her help with DNA extraction.

This study was funded by the French National Research Agency Agence Nationale de la Recherche (ANR) (contract ANR-19-CE01-0004-02), including a postdoctoral grant to D.J. LABGeM (Commissariat à l'énergie atomique et aux énergies alternatives [CEA]/Genoscope and CNRS UMR8030), France Génomique, and French Bioinformatics Institute national infrastructures (funded as part of Investissement d'Avenir program, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are acknowledged for support of the MicroScope annotation platform.

We declare no conflict of interest.

REFERENCES

1. Vaïtilingom M, Amato P, Sancelme M, Laj P, Leriche M, Delort A-M. 2010. Contribution of microbial activity to carbon chemistry in clouds. Appl Environ Microbiol 76:23–29. https://doi.org/10.1128/AEM.01127-09.
2. Vaïtilingom M, Deguillaume L, Vinatier V, Sancelme M, Amato P, Chaumerliac N, Delort A-M. 2013. Potential impact of microbial activity on the oxidant capacity and organic carbon budget in clouds. Proc Natl Acad Sci U S A 110:559–564. https://doi.org/10.1073/pnas.1205743110.
3. Yabuuchi E, Yano I, Oyaizu H, Hashimoto Y, Ezaki T, Yamamoto H. 1990. Proposals of *Sphingomonas paucimobilis* gen. nov. and comb. nov., *Sphingomonas parapaucimobilis* sp. nov., *Sphingomonas yanoikuyae* sp. nov., *Sphingomonas adhaesiva* sp. nov., *Sphingomonas capsulata* comb, nov., and two genospecies of the genus *Sphingomonas*. Microbiol Immunol 34:99–119. https://doi.org/10.1111/j.1348-0421.1990.tb00996.x.
4. Busse H-J, Denner EBM, Buczolits S, Salkinoja-Salonen M, Bennasar A, Kämpfer P. 2003. *Sphingomonas aurantiaca* sp. nov., *Sphingomonas aerolata* sp. nov. and *Sphingomonas faeni* sp. nov., air- and dustborne and Antarctic, orange-pigmented, psychrotolerant bacteria, and emended description of the genus *Sphingomonas*. Int J Syst Evol Microbiol 53:1253–1260. https://doi.org/10.1099/ijs.0.02461-0.
5. Lee H, Kim D-U, Lee S, Yun J, Park S, Yoon J-H, Park SY, Ka J-O. 2017. *Sphingomonas carri* sp. nov., isolated from a car air-conditioning system. Int J Syst Evol Microbiol 67:4069–4074. https://doi.org/10.1099/ijsem.0.002250.
6. Kim S-J, Moon J-Y, Lim J-M, Ahn J-H, Weon H-Y, Ahn T-Y, Kwon S-W. 2014. *Sphingomonas aerophila* sp. nov. and *Sphingomonas naasensis* sp. nov., isolated from air and soil, respectively. Int J Syst Evol Microbiol 64:926–932. https://doi.org/10.1099/ijs.0.055269-0.
7. Brimblecombe P, Blades N, Camuffo D, Sturaro G, Valentino A, Gysels K, Grieken R, Busse H-J, Kim O, Ulrych U, Wieser M. 1999. The indoor environment of a modern museum building, The Sainsbury Centre for Visual Arts, Norwich, UK. Indoor Air 9:146–164. https://doi.org/10.1111/j.1600-0668.1999.t01-1-00002.x.
8. Vaïtilingom M, Attard E, Gaiani N, Sancelme M, Deguillaume L, Flossmann AI, Amato P, Delort A-M. 2012. Long-term features of cloud microbiology at the puy de Dôme (France). Atmos Environ 56:88–100. https://doi.org/10.1016/j.atmosenv.2012.03.072.
9. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
10. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol 38:4647–4654. https://doi.org/10.1093/molbev/msab199.
11. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848.
12. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the Prokaryotic Genome

Annotation Pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105.
13. Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, Burlot L, Bussell X, Fouteau S, Gautreau G, Lajus A, Langlois J, Planel R, Roche D, Rollin J, Rouy Z, Sabatet V, Médigue C. 2019. MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. Nucleic Acids Res 48:579–589. https://doi.org/10.1093/nar/gkz926.
14. Kopejtka K, Tomasch J, Zeng Y, Selyanin V, Dachev M, Piwosz K, Tichý M, Bína D, Gardian Z, Bunk B, Brinkmann H, Geffers R, Sommaruga R, Koblížek M. 2020. Simultaneous presence of bacteriochlorophyll and xanthorhodopsin genes in a freshwater bacterium. mSystems 5:e01044-20. https://doi.org/10.1128/mSystems.01044-20.
15. Suzuki K, del Carmen Marín M, Konno M, Bagherzadeh R, Murata T, Inoue K. 2022. Structural characterization of proton-pumping rhodopsin lacking a cytoplasmic proton donor residue by X-ray crystallography. J Biol Chem 298:101722. https://doi.org/10.1016/j.jbc.2022.101722.
16. Huerta-Cepas J, Serra F, Bork P. 2016. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. Mol Biol Evol 33:1635–1638. https://doi.org/10.1093/molbev/msw046.
17. Letunic I, Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23:127–128. https://doi.org/10.1093/bioinformatics/btl529.