



Complete Genome Sequence of *Francisella* sp. Strain LA11-2445 (FDC406), a Novel *Francisella* Species Isolated from a Human Skin Lesion

 Caroline Öhrman,^a Ingrid Uneklint,^a Linda Karlsson,^a Laurel Respcio-Kingry,^b Mats Forsman,^a Jeannine M. Petersen,^b
 Andreas Sjödin^a

^aSwedish Defence Research Agency, CBRN Defence and Security, Umeå, Sweden

^bCenters for Disease Control and Prevention, Division of Vector-Borne Diseases, Bacterial Diseases Branch, Fort Collins, Colorado, USA

ABSTRACT Here, we present the 2,139,666-bp circular chromosome of *Francisella* sp. strain LA11-2445 (FDC406), a proposed novel species of *Francisella* that was isolated from a human cutaneous lesion and is related to *Francisella* species from marine environments.

In 2011, a novel *Francisella* species was isolated from a lesion on the left ankle of a 69-year-old man in Louisiana (1). A bacterial culture swab taken from the underlying granular tissue yielded a pure isolate of a Gram-negative coccobacillus that grew on blood and chocolate agar. The isolate was initially suspected to be *Francisella tularensis* on the basis of growth characteristics, colony morphology, Gram staining, and biochemical testing; however, the results of *F. tularensis*-specific testing (direct fluorescent antibody and PCR assays) were negative. At the Centers for Disease Control and Prevention (CDC), the isolate, designated LA11-2445, was identified as a novel *Francisella* species by multilocus sequence comparison, with the greatest nucleotide similarities to *Francisella* organisms associated with marine environments (1).

The isolate was grown on cysteine heart agar with 9% chocolated sheep blood for 48 h at 35°C in an ambient atmosphere. DNA extracted using the QIAamp DNA minikit (Qiagen, Hilden, Germany) was sent from the CDC to the Swedish Defense Research Agency and designated collection number FDC406. A short-read library was prepared with the Nextera XT kit (Illumina, San Diego, CA, USA) and sequenced on an Illumina MiSeq system (500 cycles), generating 1,354,128 reads with an N_{50} value of 251 bp. For long-read sequencing, the DNA was amplified with the multiple-displacement amplification (MDA) REPLI-g midikit (Qiagen), and a library was prepared with a 1D ligation sequencing kit (SQK-LSK108) with no shearing or size selection, barcoded with a 1D native barcoding kit (EXP-NBD103), and sequenced with an Oxford Nanopore Technologies (Oxford, UK) MinION system (R9.4). Nanopore reads were demultiplexed using Albacore v2.1.3 (Oxford Nanopore Technologies), generating 63,244 reads with an N_{50} value of 5,027 bp. Illumina reads were trimmed using Trimmomatic v0.38 (2) and Nanopore reads using Porechop v0.2.3_seqan2.1.1 (3); quality was ensured with FastQC v0.11.9 (4) and NanoPlot v1.20.0 (5). A one-sequence complete circular chromosome was generated with long and short reads as inputs using the hybrid assembler Unicycler v0.4.7 with default settings (6), including two rounds of polishing with Pilon v1.23 (7) and rotation to *dnaA* as the start at position 463527. No incongruences were found in the assembly when it was compared, using DNAdiff v1.3 (8), with a short-read assembly generated with ABySS v2.1.4 (9). To verify circularization and to find possible scaffolding errors, Nanopore reads were mapped to the assembly using minimap2 v2.15 (10), followed by variation calling using Sniffles v1.0.10 (11). The scaffold ends were spanned by $>10\times$ coverage, and no incongruences were found. The genome was annotated with NCBI Prokaryotic Genome

Citation Öhrman C, Uneklint I, Karlsson L, Respcio-Kingry L, Forsman M, Petersen JM, Sjödin A. 2021. Complete genome sequence of *Francisella* sp. strain LA11-2445 (FDC406), a novel *Francisella* species isolated from a human skin lesion. Microbiol Resour Announc 10:e01233-20. <https://doi.org/10.1128/MRA.01233-20>.

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2021 Öhrman 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 Andreas Sjödin, andreas.sjodin@foi.se.

Received 29 October 2020

Accepted 13 December 2020

Published 14 January 2021

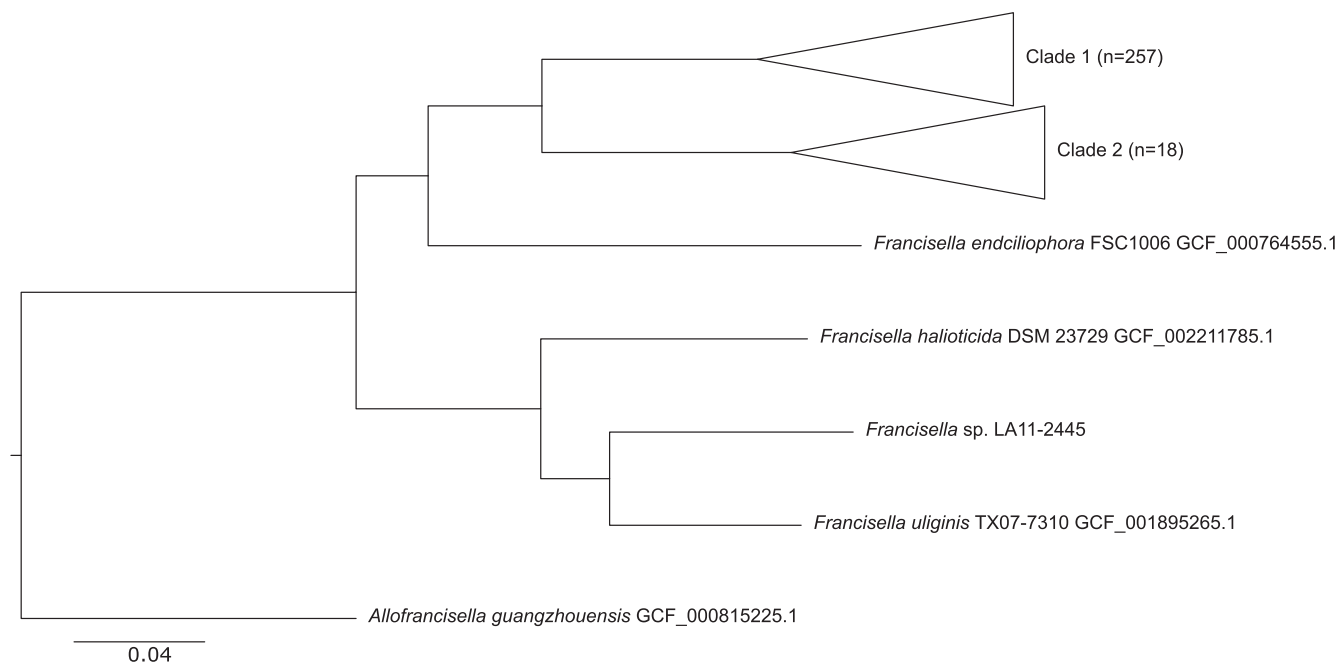


FIG 1 Whole-genome NJ phylogeny with 280 genomes, showing the relation of LA11-2445 to 279 other genomes within the genus *Francisella*. The nomenclature of clades 1 and 2 is from a previous publication (21). Public genomes are labeled using the RefSeq assembly accession number used.

Annotation Pipeline (PGAP) v4.11 (12, 13). A total of 279 genomes of the genus *Francisella* according to GTDB taxonomy (14), rooted using one genome of the nearest neighbor *Allofrancisella* (GenBank accession number [GCF_000815225.1](https://doi.org/10.1093/bioinformatics/bty149)), were pairwise aligned to the reference strain SCHUS4 (GenBank accession number [GCF_000008985.1](https://doi.org/10.1093/bioinformatics/bty149)) using progressiveMauve vsnapshot_2015_02_13 (15, 16). A whole-genome neighbor-joining (NJ) phylogenetic tree, using a Jukes-Cantor model, was calculated with FastTree v2.1.10 (17, 18) and visualized with FigTree v1.4.3 (19) (Fig. 1). The average nucleotide identity (ANI) was calculated with pyani v0.2.10 (20). All software was executed using default settings unless otherwise stated. This work was approved by the institutional review board at the CDC.

The *Francisella* sp. LA11-2445 circular chromosome is 2,139,666 bp in size, has a GC content of 31.84%, and consists of 1,978 protein-coding sequences. The phylogeny is in agreement with that in the case report (1) (Fig. 1). An ANI of 83.8% with respect to the closest neighbor, TX07-7310, an isolate from seawater from the Gulf of Mexico, is consistent with LA11-2445 being a novel species of *Francisella*.

Data availability. The complete genome sequence for LA11-2445 (FDC406) is the first version and has been deposited in GenBank under the accession number [CP041030](https://doi.org/10.1093/mgen.0.000132), and the reads have been deposited in the SRA under accession numbers [SRR11853262](https://doi.org/10.1093/mgen.0.000132) and [SRR11853263](https://doi.org/10.1093/mgen.0.000132).

REFERENCES

1. Respcio-Kingry LB, Byrd L, Allison A, Brett M, Scott-Waldron C, Galliher K, Hannah P, Mead P, Petersen JM. 2013. Cutaneous infection caused by a novel *Francisella* sp. *J Clin Microbiol* 51:3456–3460. <https://doi.org/10.1128/JCM.01105-13>.
2. 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>.
3. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genomics* 3:e000132. <https://doi.org/10.1099/mgen.0.000132>.
4. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
5. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
6. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
7. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
8. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol* 14:e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>.

9. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res* 19:1117–1123. <https://doi.org/10.1101/gr.089532.108>.
10. Li H. 2018. minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
11. Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, Schatz MC. 2018. Accurate detection of complex structural variations using single-molecule sequencing. *Nat Methods* 15:461–468. <https://doi.org/10.1038/s41592-018-0001-7>.
12. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.
13. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
14. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. 2020. A complete domain-to-species taxonomy for bacteria and archaea. *Nat Biotechnol* 38:1079–1086. <https://doi.org/10.1038/s41587-020-0501-8>.
15. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
16. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. <https://doi.org/10.1371/journal.pone.0011147>.
17. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26:1641–1650. <https://doi.org/10.1093/molbev/msp077>.
18. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2: approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
19. Rambaut A. 2016. FigTree v1.4.3. <https://github.com/rambaut/figtree>.
20. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 8:12–24. <https://doi.org/10.1039/C5AY02550H>.
21. Sjödin A, Svensson K, Ohrman C, Ahlinder J, Lindgren P, Duodo S, Hnath J, Burans JP, Johansson A, Colquhoun DJ, Larsson P, Forsman M. 2012. Genome characterisation of the genus *Francisella* reveals insight into similar evolutionary paths in pathogens of mammals and fish. *BMC Genomics* 13:268. <https://doi.org/10.1186/1471-2164-13-268>.