





Complete Genome Sequences of *Allofrancisella inopinata* SYSU YG23 and *Allofrancisella frigidaquae* SYSU 10HL1970, Isolated from Water from Cooling Systems in China

 Caroline Öhrman,^a Ingrid Uneklint,^a Linda Karlsson,^a Daniel Svensson,^a Mats Forsman,^a  Andreas Sjödin^a

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

ABSTRACT Near neighbors to the causative agent of tularemia, *Francisella tularensis*, isolated from diverse sources, have been reported in recent years. In this announcement, we present the complete sequences of circular chromosomes of one of the closest neighboring genera of *Francisella* (i.e., the type strains of *Allofrancisella inopinata* and *Allofrancisella frigidaquae*).

Recently, a new genus within the family *Francisellaceae*, named *Allofrancisella*, was published, with three different species (1, 2). We present the complete sequences of the chromosome of the type strain of *Allofrancisella inopinata* SYSU YG23 (DSM 101834, FSC1302) and of the type strain of *Allofrancisella frigidaquae* SYSU 10HL1970 (DSM 101835, FSC1303), in addition to the already published complete genome sequence of the type strain of *Allofrancisella guangzhouensis* (3), previously named *Francisella guangzhouensis* (4). The strains of this genus are all isolated from water collected from different cooling systems in the city of Guangzhou, China (4, 5). Both strains were obtained from the DSMZ (German Collection of Microorganisms and Cell Cultures; DSM 101834 and DSM 101835). The strains were grown on McLeod agar supplemented with bovine hemoglobin (Becton, Dickinson, San Jose, CA) and IsoVitalax (provided by Umeå University Hospital) at room temperature and with 5% CO₂ for 24 to 48 h, and DNA was extracted with a magnetic bead-based protocol (MagAttract high-molecular-weight DNA kit; Qiagen, Hilden, Germany). Short-read sequencing libraries were prepared using a Nextera XT 600-cycle kit (Illumina, San Diego, CA) and sequenced on an Illumina MiSeq sequencer, followed by trimming using Trimmomatic (6) and quality checking with FastQC v0.11.9 (7) prior to assembly. The long-read sequencing library, with no shearing or size selection, was prepared using ligation sequencing kit 1D (SQK-LSK108), barcoded with native barcoding kit 1D (EXP-NBD103), and sequenced on a Nanopore MinION flow cell, FLO-MIN106 (R9.4). All steps were performed according to the manufacturer's instructions unless stated otherwise. Nanopore reads were basecalled using Albacore v1.1.1 (Oxford Nanopore Technologies, Oxford, UK) and trimmed with Porechop (8) prior to assembly. See Table 1 for the read statistics. Using both Illumina and Nanopore reads as the input, complete circular chromosomes were generated using Unicycler 0.4.7 (9), with the setting `-startgenes`, with complementary *dnaA* sequences for available *Francisella* species as the input. As reported by Unicycler, both assemblies were made with five rounds of pilon, and FSC1302 was rotated to the start of *dnaA*. No ambiguities were found when comparing the final hybrid assemblies using DNAdiff of the MUMmer package (10), with a short-read assembly, generated with an abyss-pe 2.2.2 k value of 51 (11). Nanopore reads were mapped to the hybrid assembly using minimap2 (12), and by viewing the bam files in Integrative Genomics Viewer (13), circularization could be confirmed by >10× coverage. Two structural variations were reported using Sniffles (14), a 339-bp inversion at nucleotide position 1520923 and an 81-bp insertion located in a repeat area at nucleotide position

Citation Öhrman C, Uneklint I, Karlsson L, Svensson D, Forsman M, Sjödin A. 2020. Complete genome sequences of *Allofrancisella inopinata* SYSU YG23 and *Allofrancisella frigidaquae* SYSU 10HL1970, isolated from water from cooling systems in China. *Microbiol Resour Announc* 9:e00554-20. <https://doi.org/10.1128/MRA.00554-20>.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Copyright © 2020 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 27 May 2020

Accepted 29 October 2020

Published 25 November 2020

TABLE 1 Read and genome statistics

Isolate	Bacterial species	Strain ID	DSM no.	No. of MiSeq reads	MiSeq mean read length (bp)	No. of MinION reads	MinION N_{50} (bp)	Genome size (bp)	GC%	No. of protein coding sequences	GenBank accession no.	SRA accession no.
SYSU YG23	<i>Allofrancisella inopinata</i>	FSC1302	101834	491,152	142.5	71,483	14,144	1,750,491	32.0	1,534	CP038241	SRR11853071, SRR11853072
SYSU 10HL1970	<i>Allofrancisella frigidaquae</i>	FSC1303	101835	852,168	143.7	149,804	9,351	1,674,180	32.1	1,480	CP038017	SRR11853073, SRR11853074

1646356. Both variations were corrected to generate the final assembly. The corrections were verified by more than 10× coverage by both Illumina and Nanopore reads using the same mapping procedure as that described above. The annotation of the final assembly was performed using NCBI Prokaryotic Genome Annotation Pipeline (15, 16). See Table 1 for genome statistics. All software programs were executed using default settings unless otherwise stated.

Data availability. These complete genome sequences are the first versions and have been deposited in GenBank/SRA. Genome and read sequences for SYSU YG23 (FSC1302) are available under the accession no. [CP038241](https://doi.org/10.1093/journal.pcbi.1005595), [SRR11853071](https://doi.org/10.1093/journal.pcbi.1005595), and [SRR11853072](https://doi.org/10.1093/journal.pcbi.1005595) and for SYSU 10HL1970 (FSC1303) under the accession no. [CP038017](https://doi.org/10.1093/journal.pcbi.1005595), [SRR11853073](https://doi.org/10.1093/journal.pcbi.1005595), and [SRR11853074](https://doi.org/10.1093/journal.pcbi.1005595).

REFERENCES

- Qu P-H, Chen S-Y, Scholz HC, Busse H-J, Gu Q, Kämpfer P, Foster JT, Glaeser SP, Chen C, Yang Z-C. 2013. *Francisella guangzhouensis* sp. nov., isolated from air-conditioning systems. *Int J Syst Evol Microbiol* 63:3628–3635. <https://doi.org/10.1099/ijs.0.049916-0>.
- Gu Q, Li X, Qu P, Hou S, Li J, Atwill ER, Chen S. 2015. Characterization of *Francisella* species isolated from the cooling water of an air conditioning system. *Braz J Microbiol* 46:921–927. <https://doi.org/10.1590/S1517-838246320140465>.
- Svensson D, Öhrman C, Bäckman S, Karlsson E, Nilsson E, Byström M, Lärkeryd A, Myrtennäs K, Stenberg P, Qu P, Trygg J, Scholz HC, Forsman M, Sjödin A. 2015. Complete genome sequence of *Francisella guangzhouensis* strain 08HL01032 T, isolated from air-conditioning systems in China. *Genome Announc* 3:e00024-15. <https://doi.org/10.1128/genomeA.00024-15>.
- Qu P-H, Li Y, Salam N, Chen S-Y, Liu L, Gu Q, Fang B-Z, Xiao M, Li M, Chen C, Li W-J. 2016. *Allofrancisella inopinata* gen. nov., sp. nov. and *Allofrancisella frigidaquae* sp. nov., isolated from water-cooling systems, and transfer of *Francisella guangzhouensis* Qu et al. 2013 to the new genus as *Allofrancisella guangzhouensis* comb. nov. *Int J Syst Evol Microbiol* 66:4832–4838. <https://doi.org/10.1099/ijs.0.001437>.
- Qu P, Deng X, Zhang J, Chen J, Zhang J, Zhang Q, Xiao YCS, Qu P, Deng X, Zhang JJ, Chen J, Zhang Q, Xiao Y, Chen S. 2009. Identification and characterization of the *Francisella* sp. strain 08HL01032 isolated in air conditioning systems. *Wei Sheng Wu Xue Bao* 49:1003–1010.
- 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>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequencing data. 0.11.9. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genomics* 3:1–7. <https://doi.org/10.1099/mgen.0.000132>.
- 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>.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol* 5:R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
- Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, Jahesh G, Khan H, Coombe L, Warren RL, Birol I. 2017. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res* 27:768–777. <https://doi.org/10.1101/gr.214346.116>.
- Li H. 2018. minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2012. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 14:178–192. <https://doi.org/10.1093/bib/bbs017>.
- 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>.
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetverin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwatz 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>.
- Tatusova T, DiCuccio M, Badretdin A, Chetverin 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>.