



First Complete Genome Sequence of *Haemophilus influenzae* Serotype a

✉ Mariam Iskander,^{a,b} Kristy Hayden,^b Gary Van Domselaar,^{a,b} Raymond Tsang^b

University of Manitoba, Winnipeg, Manitoba, Canada^a; National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada^b

ABSTRACT *Haemophilus influenzae* is an important human pathogen that primarily infects small children. In recent years, *H. influenzae* serotype a has emerged as a significant cause of invasive disease among indigenous populations. Here, we present the first complete whole-genome sequence of *H. influenzae* serotype a.

Haemophilus influenzae is a Gram-negative coccobacillus that can cause disease such as otitis media, pneumonia, and meningitis (1). *H. influenzae* can be classified into six serotypes (a to f) based on the characteristics of the surface polysaccharides; nonencapsulated strains also exist and are referred to as nontypeable *H. influenzae* (2). Historically, *H. influenzae* serotype b (Hib) had been the major cause of invasive disease and outbreaks worldwide until the development of a Hib conjugate vaccine in the 1980s (3). The introduction of the vaccine drastically decreased the incidence rate of invasive Hib disease; however, the vaccine does not protect against non-serotype b strains. Since the introduction of the vaccine, serotype f has replaced Hib as the most dominant invasive *H. influenzae* strain, while serotype a (Hia) has emerged as a significant cause of invasive disease among the indigenous populations of North America (4–7). Little else is known about Hia, and no publicly available Hia genomes have been reported to date. Here, we report the first whole-genome sequence of a *H. influenzae* serotype a strain (NML-Hia-1, isolated from a blood culture specimen).

Genomic DNA was extracted using the Qiagen DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA). Whole-genome sequencing was performed with the PacBio RSII platform (Pacific Biosciences, Menlo Park, CA, USA) using a single-molecule real-time cell. The run generated 111,380 reads, with an average length of 9,889 bp and 576× coverage. The genome assembly was done using the Hierarchical Genome Assembly Process (HGAP) workflow, and the sequences were polished using Quiver (8) to produce a single 1,829,217 bp contig representing the closed, finished Hia chromosomal DNA with an average G+C content of 38.02% (Hia is not known to contain plasmids). The genome sequence was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP, https://www.ncbi.nlm.nih.gov/genome/annotation_prok). The annotation predicted 1,643 coding sequences, 19 rRNA operons, and 59 tRNA genes.

The Hia genome size is relatively small in comparison to other publicly available *H. influenzae* genomes. The *H. influenzae* serotype d genome (Rd KW20; GenBank accession no. NC_000907) is only 0.92 kb larger than Hia, while serotypes f (KR494) and b (10810) (GenBank accession nos. NC_022356 and NC_016809) are 26.96 kb and 152.32 kb larger than Hia, respectively. The serotype f genome has 1,646 coding sequences, only three more than Hia, while serotype d has 1,610 coding sequences and serotype b has 1,853 coding sequences.

Accession number(s). The whole-genome sequence of *H. influenzae* NML-Hia-1 has been deposited in GenBank under the accession number [CP017811](https://www.ncbi.nlm.nih.gov/nuccore/CP017811).

Received 10 November 2016 Accepted 14 November 2016 Published 19 January 2017

Citation Iskander M, Hayden K, Van Domselaar G, Tsang R. 2017. First complete genome sequence of *Haemophilus influenzae* serotype a. *Genome Announc* 5:e01506-16. <https://doi.org/10.1128/genomeA.01506-16>.

© Crown copyright 2017. 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 Gary Van Domselaar, gary.van.domselaar@phac-aspc.gc.ca, or Raymond Tsang, raymond.tsang@phac-aspc.gc.ca.

ACKNOWLEDGMENTS

We acknowledge the contribution of the scientists and staff of the McGill University and Génome Québec Innovation Centre for sequencing the *H. influenzae* genome on the PacBio RSII platform. This work was supported by operational funding from the National Microbiology Laboratory.

REFERENCES

1. Turk DC. 1984. The pathogenicity of *Haemophilus influenzae*. *J Med Microbiol* 18:1–16. <https://doi.org/10.1099/00222615-18-1-1>.
2. Pittman M. 1931. Variation and type specificity in the bacterial species *Haemophilus influenzae*. *J Exp Med* 53:471–492. <https://doi.org/10.1084/jem.53.4.471>.
3. Adams WG, Deaver KA, Cochi SL, Plikaytis BD, Zell ER, Broome CV, Wenger JD. 1993. Decline of childhood *Haemophilus influenzae* type b (Hib) disease in the Hib vaccine era. *JAMA* 269:221–226. <https://doi.org/10.1001/jama.1993.03500020055031>.
4. Ulanova M, Tsang RSW. 2014. *Haemophilus influenzae* serotype a as a cause of serious invasive infections. *Lancet Infect Dis* 14:70–82. [https://doi.org/10.1016/S1473-3099\(13\)70170-1](https://doi.org/10.1016/S1473-3099(13)70170-1).
5. Kelly L, Tsang RSW, Morgan A, Jamieson FB, Ulanova M. 2011. Invasive disease caused by *Haemophilus influenzae* type a in Northern Ontario First Nations communities. *J Med Microbiol* 60:384–390. <https://doi.org/10.1099/jmm.0.026914-0>.
6. Tsang RSW, Li YA, Mullen A, Baikie M, Whyte K, Shuel M, Tyrrell G, Rotondo JAL, Desai S, Spika J. 2016. Laboratory characterization of invasive *Haemophilus influenzae* isolates from Nunavut, Canada, 2000–2012. *Int J Circumpolar Health* 75:29798. <https://doi.org/10.3402/ijch.v75.29798>.
7. Tsang RSW, Bruce MG, Lem M, Barreto L, Ulanova M. 2014. A review of invasive *Haemophilus influenzae* disease in the indigenous populations of North America. *Epidemiol Infect* 142:1344–1354. <https://doi.org/10.1017/S0950268814000405>.
8. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.