



# Conservation and Recombination in the Genome Sequence of *Haemophilus influenzae* Type f WAPHL1

Allen C. Bateman,<sup>a</sup> Ailyn C. Perez-Osorio,<sup>b</sup> Zhen Li,<sup>b</sup> Michael Tran,<sup>b</sup> Alexander L. Greninger<sup>a</sup>

Department of Laboratory Medicine, University of Washington, Seattle, Washington, USA<sup>a</sup>; Washington State Public Health Laboratory, Seattle, Washington, USA<sup>b</sup>

**ABSTRACT** We report here the second draft genome sequence of a bloodstream isolate of *Haemophilus influenzae* serotype f. Three discrete 3.1- to 7.8-kb sites contained 80% of the variability in the genome, consistent with recombination in known virulence factors.

The incidence of invasive *Haemophilus influenzae* type b infection has decreased dramatically following vaccine introduction, but nontypeable and type f *H. influenzae* strains have increased in the United States and Sweden (1, 2). We report here a draft genome sequence of a bloodstream isolate of *H. influenzae* type f. To our knowledge, this is the second full genome of *H. influenzae* type f, which allows for a comparison to the previously published genome (3).

A man in his 50s presented with hypothermia, acidemia, and shock and was found to have hypoxemic respiratory failure. Aerobic and anaerobic bottles grew *H. influenzae* and a coagulase-negative *Staphylococcus* sp. The *H. influenzae* isolate tested beta-lactamase negative, and a workup at the Washington State Public Health laboratories showed that the *H. influenzae* isolate was type f, biotype I.

The DNeasy blood and tissue kit (Qiagen, Valencia, CA) was used to extract *H. influenzae* isolate DNA. The shotgun DNA library was prepared using a Nextera XT kit and sequenced via an Illumina MiSeq sequencer (4). A total of 2,379,348  $2 \times 250$ -bp paired-end reads were adapter and quality trimmed (Q30) using cutadapt, repaired using pairfq, *de novo* assembled using SPAdes, and annotated via Prokka (5–7). The genome assembly yielded 35 contigs measuring a total of 1,806,861 bp, with an  $N_{50}$  value of 137,750 bp and a GC content of 37.9%. A total of 1,709 coding sequences were annotated, and no plasmids were detected. The closest sequence in the NCBI NT database according to a BLASTN search and average nucleotide identity according to a BLAST search (ANIb) were *H. influenzae* KR494 (GenBank accession number CP005967) and >99.9% nucleotide identity. Multilocus sequence typing analysis of the draft genome showed the strain to be sequence type 124, representing *H. influenzae* serotype f (8). After mapping 1,983,017 of the 1,987,874 (99.8%) trimmed reads to the KR494 genome, a total of 666 variants were detected based on a minimum coverage of  $7\times$  and a minimum allele frequency of 75%.

Upon examination of these variants, we noticed three discrete regions with a high density of variants, consistent with recombination. These three regions measured 3.6 kb, 3.1 kb, and 7.8 kb in length and contained 135, 133, and 260 of the variants detected (79.3% of all variants). Genes in areas of potential recombination included those encoding IgA-specific serine endopeptidase, *S*-adenosylmethionine (SAM):tRNA ribosyltransferase-isomerase, and a hypothetical protein in the first region; the Yjll family glycine radical enzyme, radical SAM protein, and NupC/NupG family nucleoside CNT transporter in the second region; and the preprotein translocase subunit SecE, ATP-dependent protease

Received 26 July 2017 Accepted 7 August 2017 Published 21 September 2017

**Citation** Bateman AC, Perez-Osorio AC, Li Z, Tran M, Greninger AL. 2017. Conservation and recombination in the genome sequence of *Haemophilus influenzae* type f WAPHL1. Genome Announc 5:e00929-17. <https://doi.org/10.1128/genomeA.00929-17>.

**Copyright** © 2017 Bateman 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 Alexander L. Greninger, [agrening@uw.edu](mailto:agrening@uw.edu).

ATP-binding subunit ClpX, ATP-dependent Clp protease proteolytic subunit, trigger factor, TonB-dependent receptor, and RelE/StbE family type II toxin-antitoxin system mRNA interferase toxin in the third region. These regions aligned 94 to 97% by nucleotide identity to the *H. influenzae* CGSHiCZ412602 (accession number CP007805) and NCTC8143 (accession number LN831035) strains from the Czech Republic and United Kingdom, respectively.

*H. influenzae* type f has been primarily isolated from cases of pneumonia, sepsis, and meningitis (9). *H. influenzae* is well known to incorporate extracellular DNA; indeed, a laboratory study of *H. influenzae* competence demonstrated 3 to 6 recombination events per transformation of 4 to 12 kb in length, consistent with our data (10). While the spread of *H. influenzae* type f is likely clonal, the recombination ability of this organism provides it a mechanism for continued adaptation (11).

**Accession number(s).** The sequences described in this study are deposited in DDBJ/ENA/GenBank under the accession number [NIUV00000000](#).

## REFERENCES

- Rubach MP, Bender JM, Mottice S, Hanson K, Weng HYC, Korgenski K, Daly JA, Pavia AT. 2011. Increasing incidence of invasive *Haemophilus influenzae* disease in adults, Utah, USA. *Emerg Infect Dis* 17:1645–1650. <https://doi.org/10.3201/eid1709.101991>.
- Resman F, Ristovski M, Ahl J, Forsgren A, Gilsdorf JR, Jasir A, Kaijser B, Kronvall G, Riesbeck K. 2011. Invasive disease caused by *Haemophilus influenzae* in Sweden 1997–2009; evidence of increasing incidence and clinical burden of non-type b strains. *Clin Microbiol Infect* 17:1638–1645. <https://doi.org/10.1111/j.1469-0691.2010.03417.x>.
- Su YC, Resman F, Hörhold F, Riesbeck K. 2014. Comparative genomic analysis reveals distinct genotypic features of the emerging pathogen *Haemophilus influenzae* type f. *BMC Genomics* 15:38. <https://doi.org/10.1186/1471-2164-15-38>.
- Greninger AL, Chatterjee SS, Chan LC, Hamilton SM, Chambers HF, Chiu CY. 2016. Whole-genome sequencing of methicillin-resistant *Staphylococcus aureus* resistant to fifth-generation cephalosporins reveals potential non-*mecA* mechanisms of resistance. *PLoS One* 11:e0149541. <https://doi.org/10.1371/journal.pone.0149541>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Greninger AL, Langelier C, Cunningham G, Keh C, Melgar M, Chiu CY, Miller S. 2015. Two rapidly growing mycobacterial species isolated from a brain abscess: first whole-genome sequences of *Mycobacterium immunogenium* and *Mycobacterium llutzerense*. *J Clin Microbiol* 53:2374–2377. <https://doi.org/10.1128/JCM.00402-15>.
- Hoshino T, Hachisu Y, Kikuchi T, Tokutake S, Okui H, Kutsuna S, Fukasawa C, Murayama K, Oohara A, Shimizu H, Ito M, Takahashi Y, Ishiwada N. 2015. Analysis of *Haemophilus influenzae* serotype f isolated from three Japanese children with invasive *H. influenzae* infection. *J Med Microbiol* 64:355–358. <https://doi.org/10.1099/jmm.0.000031>.
- Campos J, Román F, Pérez-Vázquez M, Aracil B, Oteo J, Cercenado E, Spanish Study Group for *H. influenzae* type f. 2003. Antibiotic resistance and clinical significance of *Haemophilus influenzae* type f. *J Antimicrob Chemother* 52:961–966. <https://doi.org/10.1093/jac/dkh004>.
- Mell JC, Shumilina S, Hall IM, Redfield RJ. 2011. Transformation of natural genetic variation into *Haemophilus influenzae* genomes. *PLoS Pathog* 7:e1002151. <https://doi.org/10.1371/journal.ppat.1002151>.
- Bruun B, Gahrn-Hansen B, Westh H, Kilian M. 2004. Clonal relationship of recent invasive *Haemophilus influenzae* serotype f isolates from Denmark and the United States. *J Med Microbiol* 53:1161–1165. <https://doi.org/10.1099/jmm.0.45749-0>.