



# Closed Complete Annotated Genome Sequences of Five *Haemophilus influenzae* Biogroup *aegyptius* Strains

Zachary N. Phillips,<sup>a</sup> Greg Tram,<sup>a</sup> Michael P. Jennings,<sup>a</sup>  John M. Attack<sup>a</sup>

<sup>a</sup>Institute for Glycomics, Griffith University, Gold Coast, Queensland, Australia

**ABSTRACT** *Haemophilus influenzae* biogroup *aegyptius* is a cause of conjunctivitis in children. Biogroup *aegyptius* strains also caused fatal outbreaks of invasive disease, known as Brazilian purpuric fever (BPF), in the 1980s. BPF is fatal if untreated. Here, we report the complete genome sequences of five strains of *Haemophilus influenzae* biogroup *aegyptius*.

*Haemophilus influenzae* biogroup *aegyptius* (1) is a bacterial pathogen responsible for purulent conjunctivitis in young children. It was also the cause of the acute and deadly invasive infection Brazilian purpuric fever (BPF) in the 1980s (2). It was hypothesized that acquisition and/or expression of particular virulence factors led to the emergence of BPF strains (3, 4), but the exact mechanism(s) behind the transition from causing conjunctivitis to causing severe invasive disease was never elucidated. However, a number of *Haemophilus influenzae* biogroup *aegyptius*-specific virulence factors have been identified and studied (5), such as lipooligosaccharide structure (1), extracellular proteins (6), and outer membrane proteins/adhesins (4, 7). Nevertheless, no factors were conclusively shown to be completely necessary for virulence (1), in part due to a lack of high-quality genome sequences. Prior to this study, only four closed annotated genome sequences for *Haemophilus influenzae* biogroup *aegyptius*/*Haemophilus aegyptius* (NCBI taxonomy identification [ID] 725 and ID 197575) were available in the public domain (strains F3031 [GenBank accession number [FQ670178](https://doi.org/10.1093/nuclemta/15.12.178)], F3047 [[FQ670204](https://doi.org/10.1093/nuclemta/15.12.178)], NCTC8134 [[LR134395](https://doi.org/10.1093/nuclemta/15.12.178)], and NCTC8502 [[LS483429](https://doi.org/10.1093/nuclemta/15.12.178)]). Another 10 strains have been deposited as whole-genome shotgun (WGS) contigs. Herein, we report the closed annotated whole-genome sequences of five biogroup *aegyptius* strains, BPF isolates F1946, F3028, and F3037 and conjunctivitis (non-BPF) isolates F3043 and F3052. All five strains were originally isolated in Brazil in the 1980s (3).

Bacterial strains were grown on brain heart infusion (BHI) agar supplemented with NAD<sup>+</sup> (2 μg/ml) and hemin (1% [vol/vol]). Genomic DNA was prepared using the GenElute kit (Sigma-Aldrich) and sequenced at the Australian Genome Research Facility (AGRF) using the PacBio Sequel platform with P6-C4 chemistry and a library size of 10 kb, with barcoded libraries and a 10-hour read time using a single Sequel single-molecule real-time (SMRT) cell. Sequence reads were filtered and genomes were *de novo* assembled and polished using the hierarchical genome assembly process (HGAP) version 4 (8) using default settings, except that the genome size was set at 1,900,000 bp. The assembly quality was assessed using BUSCO analysis (9). Whole closed genome sequences for each strain were submitted to NCBI for annotation using the Prokaryotic Genome Annotation Pipeline (PGAP). The complete annotated closed genomes have been deposited in GenBank.

Information for each strain/genome is summarized in Table 1. All five strains contain a number of genes associated with *Haemophilus influenzae* biogroup *aegyptius* (4), including high-molecular-weight (HMW) adhesins containing an octanucleotide 5'-GCATCATC<sub>(n)</sub>-3' repeat in their promoter region (4) and a number of biogroup

**Citation** Phillips ZN, Tram G, Jennings MP, Attack JM. 2019. Closed complete annotated genome sequences of five *Haemophilus influenzae* biogroup *aegyptius* strains. *Microbiol Resour Announc* 8:e01198-19. <https://doi.org/10.1128/MRA.01198-19>.

**Editor** John J. Dennehy, Queens College

**Copyright** © 2019 Phillips 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 John M. Attack, [j.atack@griffith.edu.au](mailto:j.atack@griffith.edu.au).

**Received** 30 September 2019

**Accepted** 14 October 2019

**Published** 7 November 2019

**TABLE 1** Summary of information for the closed annotated genome sequences for five strains of *Haemophilus influenzae* biogroup *aegyptius*

Strain	Site of isolation	Disease	Genome size (bp)	Genome coverage (×)	Total no. of reads	Avg read length (bp)	GC content (%)	No. of genes	No. of CDS <sup>a</sup>	GenBank accession no.	SRA accession no.
F1946	Skin	BPF	1,985,844	128.4	43,712	5,833	38.2	2,028	1,947	CP043770	SRX6897973
F3028	CSF <sup>b</sup>	BPF	1,984,979	245.6	101,675	4,796	38.2	2,022	1,941	CP043771	SRX6897974
F3037	Blood	BPF	1,987,687	161.9	56,929	5,655	38.2	2,027	1,946	CP043772	SRX6897975
F3052	Conjunctiva	Conjunctivitis	1,877,864	138.2	34,648	6,488	40.7	1,898	1,822	CP043810	SRX6897976
F3043	Conjunctiva	Conjunctivitis	2,000,194	118.3	39,986	6,829	35.7	2,038	1,958	CP043811	SRX6897977

<sup>a</sup> CDS, coding DNA sequences.

<sup>b</sup> CSF, cerebrospinal fluid.

*aegyptius*-specific trimeric autotransporter proteins (4). Four of the five strains (F1946, F3028, F3043, and F3037) contain an ~32.5-kb plasmid (GenBank accession number AF447808) associated with *Haemophilus influenzae* biogroup *aegyptius* strains (10). Interestingly, our sequence of strain F1946 is ~28 kb larger than the WGS sequence deposited recently (11). The extra sequences, able to be resolved here due to long-read sequencing, consist of a number of duplicated regions, such as rRNA gene clusters (~5.5 kb each), the *licABC* operon involved in lipooligosaccharide (LOS) biosynthesis (~2.5 kb), and an ~6-kb region encoding pilin and the *haf* pilus export machinery.

These five closed and annotated genome sequences have more than doubled the number of whole closed annotated *Haemophilus influenzae* biogroup *aegyptius* genomes that are available in the public domain and will facilitate the elucidation of the exact factors responsible for the unusual virulence of this lineage of *Haemophilus influenzae*.

**Data availability.** The genomes have been deposited in GenBank. The accession numbers for the closed genomes and raw data (SRA) are provided in Table 1.

## ACKNOWLEDGMENTS

We thank the Australian Genome Research Facility (AGRF) for expert technical assistance in carrying out SMRT sequencing and methylome analysis.

This work was supported by Australian Research Council (ARC) discovery project grants 180100976 to J.M.A. and 170104691 to M.P.J., an Australian National Health and Medical Research Council (NHMRC) project grant (1099279) to J.M.A., and an NHMRC program grant (1071659) and principal research fellowship (1138466) to M.P.J.

## REFERENCES

- Harrison LH, Simonsen V, Waldman EA. 2008. Emergence and disappearance of a virulent clone of *Haemophilus influenzae* biogroup *aegyptius*, cause of Brazilian purpuric fever. *Clin Microbiol Rev* 21:594–605. <https://doi.org/10.1128/CMR.00020-08>.
- Harrison LH, da Silva GA, Pittman M, Fleming DW, Vranjac A, Broome CV. 1989. Epidemiology and clinical spectrum of Brazilian purpuric fever. Brazilian Purpuric Fever Study Group. *J Clin Microbiol* 27:599–604.
- Brenner DJ, Mayer LW, Carlone GM, Harrison LH, Bibb WF, Brandileone MC, Sottnek FO, Irino K, Reeves MW, Swenson JM. 1988. Biochemical, genetic, and epidemiologic characterization of *Haemophilus influenzae* biogroup *aegyptius* (*Haemophilus aegyptius*) strains associated with Brazilian purpuric fever. *J Clin Microbiol* 26:1524–1534.
- Strouts FR, Power P, Croucher NJ, Corton N, van Tonder A, Quail MA, Langford PR, Hudson MJ, Parkhill J, Kroll JS, Bentley SD. 2012. Lineage-specific virulence determinants of *Haemophilus influenzae* biogroup *aegyptius*. *Emerg Infect Dis* 18:449–457. <https://doi.org/10.3201/eid1803.110728>.
- Carlone GM, Gorelkin L, Gheesling LL, Hoiseth SK, Mulks MH, O'Connor SP, Weyant RS, Myrick JE, Mayer LW, Arko RJ. 1989. Potential virulence factors of *Haemophilus influenzae* biogroup *aegyptius* in Brazilian purpuric fever. The Brazilian Purpuric Fever Study Group. *Pediatr Infect Dis J* 8:245–247.
- Barbosa SF, Hoshino-Shimizu S, Alkmin M, Goto H. 2003. Implications of *Haemophilus influenzae* biogroup *aegyptius* hemagglutinins in the pathogenesis of Brazilian purpuric fever. *J Infect Dis* 188:74–80. <https://doi.org/10.1086/375739>.
- Cury GCG, Pereira RFC, de Hollanda LM, Lancellotti M. 2014. Inflammatory response of *Haemophilus influenzae* biotype *aegyptius* causing Brazilian purpuric fever. *Braz J Microbiol* 45:1449–1454. <https://doi.org/10.1590/s1517-83822014000400040>.
- 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>.
- Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
- Kroll JS, Farrant JL, Tyler S, Coulthart MB, Langford PR. 2002. Characterisation and genetic organisation of a 24-MDa plasmid from the Brazilian purpuric fever clone of *Haemophilus influenzae* biogroup *aegyptius*. *Plasmid* 48:38–48. [https://doi.org/10.1016/S0147-619X\(02\)00020-3](https://doi.org/10.1016/S0147-619X(02)00020-3).
- Pereira RFC, Mofatto LS, Silva ACA, Alves DA, Machado D, Theizen TH, Pereira GAG, Levy CE, de Hollanda LM, Carazzolle MF, Lancellotti M. 2019. Draft whole-genome sequences of *Haemophilus influenzae* biogroup *aegyptius* strains isolated from five Brazilian purpuric fever cases and one conjunctivitis case. *Microbiol Resour Announc* 8:e00642-19. <https://doi.org/10.1128/MRA.00642-19>.