UNIVERSIDADE ESTADUAL DE CAMPINAS
SISTEMA DE BIBLIOTECAS DA UNICAMP
REPOSITÓRIO DA PRODUÇÃO CIENTIFICA E INTELECTUAL DA UNICAMP

Versão do arquivo anexado / Version of attached file:

Versão do Editor / Published Version

Mais informações no site da editora / Further information on publisher's website:

https://mra.asm.org.ez88.periodicos.capes.gov.br/content/8/30/e00642-19

DOI: 10.1128/MRA.00642-19

Direitos autorais / Publisher's copyright statement:

Draft Whole-Genome Sequences of *Haemophilus influenzae* Biogroup *aegyptius* Strains Isolated from Five Brazilian Purpuric Fever Cases and One Conjunctivitis Case

Rafaella F. C. Pereira,[a] Luciana S. Mofatto,[b] Ana C. A. Silva,[c] Danilo A. Alves,[c] Daisy Machado,[a] Thaís H. Theizen,[a] Gonçalo A. G. Pereira,[b] Carlos E. Levy,[d] Luciana M. de Hollanda,[a]* Marcelo F. Carazzolle,[b] Marcelo Lancellotti[a,c]

[a]Department of Biochemistry and Tissue Biology, Institute of Biology, State University of Campinas (UNICAMP), Campinas, São Paulo, Brazil
[b]Department of Genetics, Evolution and Bioagents, Institute of Biology, UNICAMP, Campinas, São Paulo, Brazil
[c]Faculty of Pharmaceutical Sciences, UNICAMP, Campinas, São Paulo, Brazil
[d]Department of Clinical Pathology, School of Medical Sciences, UNICAMP, Campinas, São Paulo, Brazil

ABSTRACT Brazilian purpuric fever is a febrile hemorrhagic pediatric disease caused by *Haemophilus influenzae* biogroup *aegyptius*, a bacterium which was formerly associated with only self-limited purulent conjunctivitis. Here, we present draft genomes of strains from five Brazilian purpuric fever cases and one conjunctivitis case.

For over a century, *Haemophilus influenzae* biogroup *aegyptius* was associated only with seasonal epidemics of self-limited purulent conjunctivitis ("pink eye"). In the 1980s, an emergent clone of *H. influenzae* biogroup *aegyptius* was identified as the etiological agent of Brazilian purpuric fever (BPF), a fulminant pediatric disease characterized by conjunctivitis, high fever, purpura, and sepsis with a fatality rate of 40 to 70% (1, 2). Major outbreaks of the disease occurred from 1984 to 1990 in the state of São Paulo, Brazil. Sporadic cases have been reported in Australia, the United States (2), and, more recently in 2007, in the Brazilian state of Pará (3).

The BPF clone refers to a group of closely similar, but not identical, strains of *H. influenzae* biogroup *aegyptius* that were associated with the Brazilian cases and have specific properties such as the presence of an approximately 32-kb plasmid referred to as 3031 and a characteristic multilocus enzyme electrophoresis (MLEE) profile (electrophoretic type 2) (2). To date, only one whole-genome sequence of a BPF clone strain is available in GenBank (strain F3031, GenBank accession no. FQ670178), in which were described 21 *H. influenzae* biogroup *aegyptius*-BPF-specific coding sequences (CDSs) (4). However, the origin and virulence mechanisms of *H. influenzae* biogroup *aegyptius* associated with BPF still remain a mystery. In this study, we sequenced five additional strains isolated from BPF cases in Brazil and one strain from a conjunctivitis case in the United States.

Strains stored at −80°C were grown on chocolate agar at 37°C with 5% CO_2. Genomic DNA was extracted using a cetyltrimethylammonium bromide (CTAB) extraction method (5). The DNA libraries were prepared with the Nextera XT DNA library preparation kit (Illumina, CA, USA) and sequenced using the Illumina HiSeq 2000 platform (100-bp paired-end reads). The number of sequenced reads ranged from 30,479,940 to 53,773,000, representing an extremely high sequencing coverage of 1,604× and 2,830×, respectively (see Table 1 for total reads and genome coverage per sample). To reduce coverage, the reads were filtered by quality (Phred quality score of >30), and only a total of 200× coverage for each sample was considered. The reads were *de novo* assembled using Velvet v. 1.2.03 (6), followed by gene annotation using RASTtk (7–9) for exploratory analysis and the NCBI Prokaryotic Genome Annotation

Address correspondence to Rafaella F. C. Pereira, rafaella.carneirop@gmail.com.

* Present address: Luciana M. de Hollanda, Institute of Technology and Research (ITP), Tiradentes University (UNIT), Aracaju, Sergipe, Brazil.

TABLE 1 Characteristics and genome sequencing and assembly statistics of *H. influenzae* biogroup *aegyptius* strains

Parameter	Data for strain:					
	F3030	F3039	F3042	F3283	F1946	KC1018
Isolation site	Blood	Blood	Oropharynx	Blood	Skin lesion	Conjunctiva
Genome size (bp)	1,962,639	1,961,979	1,969,657	1,962,133	1,957,726	1,909,282
No. of contigs	67	138	137	130	74	96
N_{50}	62,606	24,511	24,813	27,365	59,820	39,216
G+C content (%)	38.0	38.0	38.0	38.0	38.0	38.1
No. of genes	2,008	2,045	2,056	2,035	1,987	1,957
No. of CDSs	1,849	1,836	1,847	1,848	1,840	1,787
Total no. of reads	30,479,940	39,906,340	35,636,936	36,692,872	38,647,860	53,773,000
Genome coverage (×)	1,604	2,100	1,875	1,931	2,034	2,830
GenBank accession no.	LNKQ00000000	LNKN00000000	LNKO00000000	LNKP00000000	LNKS00000000	LNKR00000000
SRA accession no.	SRX5957578	SRX5975171	SRX5958526	SRX5957478	SRX5953066	SRX5954865

Pipeline (PGAP) v. 3.0 (10) for deposition to GenBank. The NCBI PGAP identified a total of 1,957 to 2,056 genes, 1,787 to 1,849 CDSs, 96 to 156 pseudogenes, and 51 to 54 RNA genes. The strain information and genome statistics are listed in Table 1.

A nucleotide BLAST search (11) of the assembled contigs showed the presence of the 21 specific *H. influenzae* biogroup *aegyptius*-BPF CDSs (4) only in the BPF strains. And, as previously described (1), the five BPF strains have the 3031 plasmid (strain F3031, GenBank accession no. AF447808) (12), which is absent in KC1018.

Draft genomes were also submitted to the *H. influenzae* multilocus sequence typing (MLST) website (https://pubmlst.org/hinfluenzae/) (13). The BPF-associated strains were found to be of sequence type 65 (ST65), while the conjunctivitis strain has the ST72 profile.

The data presented here will be useful for further studies on the genetic characterization of *H. influenzae* biogroup *aegyptius* associated with BPF.

Data availability. The GenBank and SRA accession numbers are given in Table 1.

REFERENCES

1. Brenner DJ, Mayer LW, Carlone GM, Harrison LH, Bibb WF, Brandileone MC, Sottnek FO, Irino K, Reeves MW, Swenson JM, Birkness KA, Weyant RS, Berkley SF, Woods TC, Steigerwalt AG, Grimont PA, McKinney RM, Fleming DW, Gheesling LL, Cooksey RC, Arko RJ, Broome CV, The Brazilian Purpuric Fever Study Group. 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.
2. 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.
3. Santana-Porto EA, Oliveira AA, da Costa MRM, Pinheiro AS, Oliveira C, Lopes ML, Pereira LE, Sacchi C, Aráujo WN, Sobel J. 2009. Suspected Brazilian purpuric fever, Brazilian Amazon region. Emerg Infect Dis 15:675–676. https://doi.org/10.3201/eid1504.090014.
4. 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.
5. Wilson K. 1987. Preparation of genomic DNA from bacteria, p 2.4.1–2.4.5. In Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (ed), Current protocols in molecular biology. Greene Publishing & Wiley Interscience, New York, NY.

6. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.
7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
8. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206–214. https://doi.org/10.1093/nar/gkt1226.
9. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, 3rd, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.
10. 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.

11. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

12. 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.

13. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1.