

## GENOME ANNOUNCEMENT

### Complete Genome Sequence of *Aggregatibacter (Haemophilus) aphrophilus* NJ8700<sup>∇</sup>

Maria Pia Di Bonaventura,<sup>1</sup> Rob DeSalle,<sup>1\*</sup> Mihai Pop,<sup>2</sup> Niranjana Nagarajan,<sup>2</sup> David H. Figurski,<sup>3</sup>  
Daniel H. Fine,<sup>4</sup> Jeffrey B. Kaplan,<sup>4</sup> and Paul J. Planet<sup>1,3,5</sup>

Sackler Institute for Comparative Genomics, American Museum of Natural History, 79th Street at Central Park West, New York, New York 10024<sup>1</sup>; Center for Bioinformatics and Computational Biology, University of Maryland, College Park, Maryland 20742<sup>2</sup>; Department of Microbiology, Columbia University, College of Physicians & Surgeons, New York, New York 10032<sup>3</sup>; Department of Oral Biology, New Jersey Dental School, Newark, New Jersey 07103<sup>4</sup>; and Department of Pediatrics, Children's Hospital of New York, Columbia University, College of Physicians and Surgeons, New York, New York 10032<sup>5</sup>

Received 1 April 2009/Accepted 5 May 2009

**We report the finished and annotated genome sequence of *Aggregatibacter aphrophilus* strain NJ8700, a strain isolated from the oral flora of a healthy individual, and discuss characteristics that may affect its dual roles in human health and disease. This strain has a rough appearance, and its genome contains genes encoding a type VI secretion system and several factors that may participate in host colonization.**

*Aggregatibacter aphrophilus* (formerly *Haemophilus aphrophilus*) (11) is well known as an etiologic agent in infectious endocarditis caused by gram-negative bacteria (7). Most often, however, it is found as a nonpathogenic, commensal resident of dental plaque and the oropharyngeal flora. The complete genome sequence of *A. aphrophilus* NJ8700 was achieved using a hybrid approach of a shotgun sequencing strategy combined with 454 pyrosequencing (two runs). The 454 sequences were assembled with the Newbler assembler (454 Life Sciences), and the Sanger reads were added to the resulting contigs using MUMmer (8) and custom scripts. The contigs were linked together into scaffolds using Bambus (15), and gaps between contigs were closed by direct sequencing using a technique described by N. Nagarajan et al. (submitted for publication), achieving a 25-fold coverage. Automated annotation was done at the Institute for Genomic Research/J. Craig Venter Institute through the Annotation Engine Service.

The *A. aphrophilus* NJ8700 genome is 2,313,035 bp in length, with a GC content of 42.23% and 2,320 predicted coding sequences. Approximately 88.4% of nucleotides are predicted to encode proteins. The genome contains 57 tRNAs, including one gene for tRNA<sup>Sec</sup> (AAP\_1961), and five rRNA nontandem cistrons. Like other *Pasteurellaceae* (3, 4), the genome has four RNA subunit genes (*rpoA*, *rpoB*, *rpoC*, and *rpoZ*; AAP\_2188, AAP\_1813, AAP\_1812, and AAP\_1427), and five sigma factor genes (AAP\_1594, AAP\_1967, AAP\_2019, AAP\_2021, and AAP\_2324).

The *A. aphrophilus* NJ8700 genome contains genes encoding

a type VI secretion system (T6SS) (AAP\_1851 to -1862, AAP\_2123), which is the first instance of its presence in a member of the *Pasteurellaceae* (1, 2, 9, 10, 16, 17, 20). There are several open reading frames (ORFs) similar to *vgrG* (AAP\_0259, AAP\_0279 to -0281, AAP\_0288, AAP\_292, AAP\_1540, AAP\_1541, AAP\_2121) that encode other possible substrates. The *flp-tad* cluster (AAP\_0177 to -0190) is similar to the *tad* locus involved in the rough colony phenotype in *Aggregatibacter actinomycetemcomitans* (12, 13, 14, 21). Also present is a locus required for the assembly of type IVa pili, including *pilF*, *pilA*, *pilB*, *pilC*, and *pilD* (AAP\_0008, AAP\_1464 to -1467). The *A. aphrophilus* genome contains genes encoding several adhesins that may participate in host colonization (EmaA, AAP\_0065; Aae, AAP\_0152; YadA and Hia, AAP\_0523 and AAP\_0527).

Genes for the production of PGA (poly-*N*-acetylglucosamine), i.e., *hmsD*, *pgaC*, *pgaB*, and *pgaA* (AAP\_1678 to -1681); *N*-acetylneuraminase lyase (*nanA*, AAP\_A0548); and the *dspB* enzyme that degrades PGA (AAP\_0383 and AAP\_0384), all implicated in biofilm formation, are present (6).

*A. aphrophilus* NJ8700 has several loci implicated in iron utilization, including one for a predicted hemoglobin binding protein, *hgpA* (AAP\_1269), and a hemoglobin/transferrin binding receptor (AAP\_2099). Genes for the hemophore receptor HasR (AAP\_1311) and a heme utilization protein (AAP\_2308) are present. A gene for the TbpA (transferrin binding protein; AAP\_1194, AAP\_1226) may signal an ability to use transferrin. *A. aphrophilus* carries genes encoding potential siderophore receptors (AAP\_0347, AAP\_0905), a TonB-dependent hemoglobin/transferrin/lactoferrin family receptor (AAP\_1145), and a receptor for ferrienterochelin/colicins (AAP\_1677). Two ORFs may encode chelatin transporters (AAP\_1146 to -1149; AAP\_783 to -785), along with the ferric-dicitrate transport system (*fecBCDE*, AAP\_1294 to -1297). The genome harbors the genes coding for the Fur

\* Corresponding author. Mailing address: Sackler Institute for Comparative Genetics, American Museum of Natural History, 79th Street at Central Park West, New York, NY 10024. Phone: (212) 769-5670. Fax: 212-769-5277. E-mail: desalle@amnh.org.

<sup>∇</sup> Published ahead of print on 15 May 2009.

regulator (AAP\_0360) and periplasmic-binding transport systems: the *afe* locus (AAP\_0393, AAP\_0395 to -0397), the *hit* locus (AAP\_1640, AAP\_1644 to -1654), and *afu* loci (AAP\_0695 to -0697, AAP\_1193 to -1196) (5, 18, 19, 22).

There are three regions of the bacterial chromosome where phage/prophage gene clusters were identified, including the *accA*-GMP gene intergenic region (acetyl coenzyme A carboxylase, AAP\_0460; GMP synthase, AAP\_0517), which harbors a prophage (M. Di Bonaventura et al., submitted for publication).

**Nucleotide sequence accession number.** The complete genome sequence of *Aggregatibacter (Haemophilus) aphrophilus* strain NJ8700 has been assigned GenBank accession number CP001607.

We thank the Sackler Institute for Comparative Genomics at the American Museum of Natural History for continued support during this project. This genome sequencing project was supported by funding from the Department of Energy.

#### REFERENCES

- Bingle, L. E., C. M. Bailey, and M. J. Pallen. 2008. Type VI secretion: a beginner's guide. *Curr. Opin. Microbiol.* **11**:3–8.
- Filloux, A., A. Hachani, and S. Bleves. 2008. The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiology* **154**:1570–1583.
- Foote, S. J., J. T. Bosse, A. E. Bouevitch, P. R. Langford, N. M. Young, and J. H. E. Nash. 2007. The complete genome sequence of *Actinobacillus pleuropneumoniae* L20 (serotype 5b). *J. Bacteriol.* **190**:1495–1496.
- Gioia, J., X. Qin, H. Jiang, K. Clinkenbeard, R. Lo, Y. Liu, G. E. Fox, S. Yerrapragada, M. P. McLeod, T. Z. McNeill, L. Hemphill, E. Sodergren, Q. Wang, D. M. Muzny, F. J. Homsy, G. M. Weinstock, and S. K. Highlander. 2006. The genome sequence of *Mannheimia haemolytica* A1: insights into virulence, natural competence, and *Pasteurellaceae* phylogeny. *J. Bacteriol.* **188**:7257–7266.
- Graber, K. R., L. M. Smoot, and L. A. Actis. 1998. Expression of iron binding proteins and hemin binding activity in the dental pathogen *Actinobacillus actinomycetemcomitans*. *FEMS Microbiol. Lett.* **163**:135–142.
- Kaplan, J. B., M. F. Meyenhofer, and D. H. Fine. 2003. Biofilm growth and detachment of *Actinobacillus actinomycetemcomitans*. *J. Bacteriol.* **185**:1399–1404.
- Khairat, O. 1940. Endocarditis due to a new species of *Haemophilus*. *J. Pathol. Bacteriol.* **50**:497–505.
- Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, and S. L. Salzberg. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* **5**:R12.
- Ludu, J. S., O. M. de Bruin, B. N. Duplantis, C. L. Scherck, A. Y. Chou, K. L. Elkins, and F. E. Nano. 2008. The *Francisella* pathogenicity island protein PdpD is required for full virulence and associates with homologues of the type VI secretion system. *J. Bacteriol.* **190**:4584–4595.
- Mougous, J. D., M. E. Cuff, S. Raunser, A. Shen, M. Zhou, C. A. Gifford, A. L. Goodman, G. Joachimiak, C. L. Ordonez, S. Lory, T. Walz, A. Joachimiak, and J. J. Mekalanos. 2006. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* **312**:1526–1530.
- Norskov-Lauritsen, N., and M. Kilian. 2006. Reclassification of *Actinobacillus actinomycetemcomitans*, *Haemophilus aphrophilus*, *Haemophilus paraphrophilus* and *Haemophilus segnis* as *Aggregatibacter actinomycetemcomitans* gen. nov., comb. nov., *Aggregatibacter aphrophilus* comb. nov. and *Aggregatibacter segnis* comb. nov., and emended description of *Aggregatibacter aphrophilus* to include V factor-dependent and V factor-independent isolates. *Int. J. Syst. Evol. Microbiol.* **56**:2135–2146.
- Perez, B. A., P. J. Planet, S. C. Kachlany, M. Tomich, D. H. Fine, and D. H. Figurski. 2006. Genetic analysis of the requirement for *flp-2*, *tadV*, and *rcpB* in *Actinobacillus actinomycetemcomitans* biofilm formation. *J. Bacteriol.* **188**:6361–6375.
- Planet, P. J. 2006. Tree disagreement: measuring and testing incongruence in phylogenies. *J. Biomed. Inform.* **39**:86–102.
- Planet, P. J., S. C. Kachlany, D. H. Fine, R. DeSalle, and D. H. Figurski. 2003. The widespread colonization island of *Actinobacillus actinomycetemcomitans*. *Nat. Genet.* **34**:193–198.
- Pop, M., D. S. Kosack, and S. L. Salzberg. 2004. Hierarchical scaffolding with Bambus. *Genome Res.* **14**:149–159.
- Pukatzki, S., A. T. Ma, A. T. Revel, D. Sturtevant, and J. J. Mekalanos. 2007. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci. USA* **104**:15508–15513.
- Pukatzki, S., A. T. Ma, D. Sturtevant, B. Krastins, D. Sarracino, W. C. Nelson, J. F. Heidelberg, and J. J. Mekalanos. 2006. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. USA* **103**:1528–1533.
- Rhodes, E. R., S. Menke, C. Shoemaker, A. P. Tomaras, G. McGillivray, and L. A. Actis. 2007. Iron acquisition in the dental pathogen *Actinobacillus actinomycetemcomitans*: what does it use as a source and how does it get this essential metal? *Biometals* **20**:365–377.
- Rhodes, E. R., A. P. Tomaras, G. McGillivray, P. L. Connerly, and L. A. Actis. 2005. Genetic and functional analyses of the *Actinobacillus actinomycetemcomitans* AfeABCD siderophore-independent iron acquisition system. *Infect. Immun.* **73**:3758–3763.
- Schell, M. A., R. L. Ulrich, W. J. Ribot, E. E. Brueggemann, H. B. Hines, D. Chen, L. Lipscomb, H. S. Kim, J. Mrazek, W. C. Nierman, and D. Deshazer. 2007. Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol. Microbiol.* **64**:1466–1485.
- Tomich, M., P. J. Planet, and D. H. Figurski. 2007. The *tad* locus: postcards from the widespread colonization island. *Nat. Rev. Microbiol.* **5**:363–375.
- Willemssen, P. T., I. Vulto, M. Boxem, and J. de Graaff. 1997. Characterization of a periplasmic protein involved in iron utilization of *Actinobacillus actinomycetemcomitans*. *J. Bacteriol.* **179**:4949–4952.