AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Closed Complete Genome Sequences of Two Nontypeable *Haemophilus influenzae* Strains Containing Novel *modA* Alleles from the Sputum of Patients with Chronic Obstructive Pulmonary Disease

John M. Atack,[a] Timothy F. Murphy,[b] Lauren O. Bakaletz,[c] Kate L. Seib,[a] Michael P. Jennings[a]

[a]Institute for Glycomics, Griffith University, Gold Coast, Queensland, Australia

[b]Clinical and Translational Research Center, University at Buffalo, State University of New York, Buffalo, New York, USA

[c]Center for Microbial Pathogenesis, The Research Institute at Nationwide Children's Hospital and The Ohio State University College of Medicine, Columbus, Ohio, USA

ABSTRACT Nontypeable *Haemophilus influenzae* (NTHi) is an important bacterial pathogen that causes otitis media and exacerbations of chronic obstructive pulmonary disease (COPD). Here, we report the complete genome sequences of NTHi strains 10P129H1 and 84P36H1, isolated from COPD patients, which contain the phase-variable epigenetic regulators ModA15 and ModA18, respectively.

Nontypeable *Haemophilus influenzae* (NTHi) is responsible for human respiratory tract infections (1, 2). Previous work characterizing NTHi showed that phase-variable N^6-adenine DNA methyltransferases (ModA) are involved in epigenetic regulation and virulence (3–6). Phase-variable methyltransferase expression leads to genome-wide methylation differences, epigenetically regulating multiple genes—a phase-variable regulon (phasevarion) (7, 8). *modA* alleles show high variability (<25% nucleotide identity) in their central target recognition domain (TRD), which dictates specificity (9). Different TRDs methylate different sequences and define a phasevarion (7). We have shown that ~65% of otitis media (OM) clinical isolates possessed one of five *modA* alleles, *modA2*, *-4*, *-5*, *-9*, or *-10* (4). Examination of *modA* alleles present in NTHi from a clinical collection of sputum samples from COPD patients (10) revealed two uncharacterized *modA* alleles, *modA15* and *modA18*.

We picked two strains, each containing a new *modA* allele (strain 10P129H1 contains *modA15*; strain 84P36H1 contains *modA18*), for genome sequencing and methylome analysis. DNA was sequenced at the Yale Center for Genome Analysis (YCGA) using a PacBio RS II platform with P6-C4 chemistry and a library size of 10 kb, with one strain per single-molecule real-time (SMRT) cell, and assembled *de novo* using the Hierarchical Genome Assembly Process (HGAP) (11). Preassembly was carried out using Celera Assembler v8.1 to the unitig step followed by a custom unitig consensus caller (YCGA). The first set of alignments was found by querying an index of the reference genome and then refining until only high-scoring alignments were retained (YCGA). Polishing for a pure PacBio assembly was carried out using the Quiver algorithm. Consensus sequences were submitted to NCBI for annotation with the Prokaryotic Genome Annotation Pipeline (PGAP), and annotated sequences were submitted to GenBank.

NTHi strain 10P129H1 resolved into a genome of 2,047,595 bp with a G+C content of 37.9% and containing 2,079 open reading frames (ORFs). Strain 10P129H1 encodes *modA15*, containing 5′-$AGCC_{(16)}$ repeats in its ORF. NTHi strain 84P36H1 resolved into a genome of 2,025,527 bp with a G+C content of 38.4% and containing 2,115 ORFs. Strain 84P36H1 encodes the *modA18* allele, containing 5′-$AGCC_{(19)}$ repeats in its ORF.

Received 5 June 2018 **Accepted** 19 June 2018 **Published** 19 July 2018

Citation Atack JM, Murphy TF, Bakaletz LO, Seib KL, Jennings MP. 2018. Closed complete genome sequences of two nontypeable *Haemophilus influenzae* strains containing novel *modA* alleles from the sputum of patients with chronic obstructive pulmonary disease. Microbiol Resour Announc 7:e00821-18. https://doi.org/10.1128/MRA.00821-18.

Editor Irene L. G. Newton, Indiana University Bloomington

Address correspondence to Kate L. Seib, k.seib@griffith.edu.au, or Michael P. Jennings, m.jennings@griffith.edu.au.

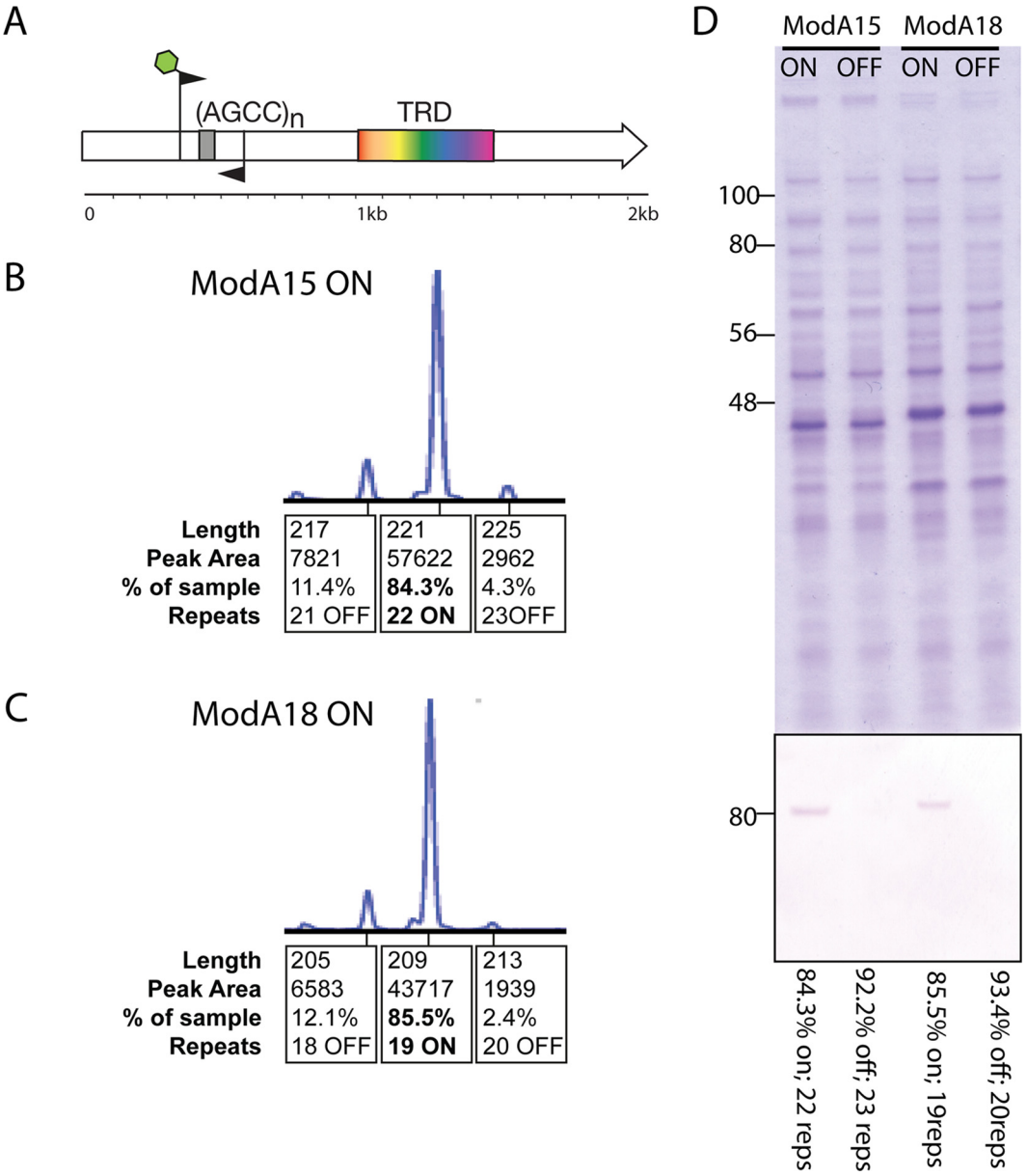

FIG 1 Confirmation of the on/off status of *modA15* and *modA18* in the studied strains. (A) *modA* gene, showing the location of the target recognition domain (TRD), which is highly variable between each *modA* allele and dictates the specificity of the ModA protein. The 5′ and 3′ regions are highly conserved (95% identity) between alleles. The location of the primers used for fragment analysis by performing PCR over the $AGCC_{(n)}$ repeat tracks is shown, with the forward primer containing a fluorescent FAM (6-fluorescein) label (green hexagon) so fragments can be sized by GenScanner. (B and C) Fragment analysis traces of *modA15* on and *modA18* on strains showing that the majority of the bacterial population contains 22 AGCC repeats (*modA15*) and 19 AGCC repeats (*modA18*) in their open reading frame, meaning the gene is in frame, i.e., is on and therefore expressed. (D) Western blot and accompanying Coomassie stain of the *modA15* and *modA18* on strains with paired isolates of the same strain where the *modA* gene is out of frame, i.e., off, and not expressed.

modA in both strains is expressed (on) with the strains' respective numbers of $AGCC_{(n)}$ repeats. The pregenome on/off status of *modA15* and *modA18* was verified using our fragment analysis approach as detailed previously (4) (Fig. 1). We also used Western blotting using an anti-ModA antibody to verify the presence of ModA in the culture used for DNA preparation for SMRT sequencing. Western blotting was carried out as described previously (4) (Fig. 1).

Strain 84P36H1 contains major NTHi virulence factors, such as lipooligosaccharide biosynthetic loci, and genes encoding the adhesins HMW1 and HMW2. Strain 10P129H1

contains a number of features associated with *Haemophilus influenzae* biogroup aegyptius (12), including a number of autotransporter adhesins and biogroup aegyptius-specific high-molecular-weight (HMW) proteins containing an octanucleotide 5′-GCATCATC$_{(n)}$ repeat in their upstream region (12).

These data provide insight into the pathobiology of NTHi and will aid in the development of novel vaccines and antibacterial strategies.

Data availability. The complete genome sequences of the *Haemophilus influenzae* strains described in this article have been deposited in NCBI GenBank under the accession numbers CP029620 (10P129H1) and CP029621 (84P36H1).

ACKNOWLEDGMENTS

We thank the Yale Center for Genome Analysis (YCGA).

Funding was provided by Australian National Health and Medical Research Council (NHMRC) project grant 1099279 to K.L.S. and J.M.A., career development fellowship 1045235 to K.L.S., and program grant 1071659 and principal research fellowship 1138466 to M.P.J.; a Garnett Passe and Rodney Williams grant-in-aid (supplementation) to K.L.S. and J.M.A.; NIH R01 grant DC015688 to L.O.B. and M.P.J. and grant AI19641 to T.F.M.; and Australian Research Council (ARC) discovery grant 180100976 to J.M.A.

REFERENCES

1. Pichichero ME, Casey JR, Hoberman A, Schwartz R. 2008. Pathogens causing recurrent and difficult-to-treat acute otitis media, 2003–2006. Clin Pediatr 47:901–906. https://doi.org/10.1177/0009922808319966.
2. Sethi S, Murphy TF. 2008. Infection in the pathogenesis and course of chronic obstructive pulmonary disease. N Engl J Med 359:2355–2365. https://doi.org/10.1056/NEJMra0800353.
3. Srikhanta YN, Maguire TL, Stacey KJ, Grimmond SM, Jennings MP. 2005. The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. Proc Natl Acad Sci U S A 102:5547–5551. https://doi.org/10.1073/pnas.0501169102.
4. Atack JM, Srikhanta YN, Fox KL, Jurcisek JA, Brockman KL, Clark TA, Boitano M, Power PM, Jen FE-C, McEwan AG, Grimmond SM, Smith AL, Barenkamp SJ, Korlach J, Bakaletz LO, Jennings MP. 2015. A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable *Haemophilus influenzae*. Nat Commun 6:7828. https://doi.org/10.1038/ncomms8828.
5. Brockman KL, Branstool MT, Atack JM, Robledo-Avila F, Partida-Sanchez S, Jennings MP, Bakaletz LO. 2017. The ModA2 phasevarion of nontypeable *Haemophilus influenzae* regulates resistance to oxidative stress and killing by human neutrophils. Sci Rep 7:3161. https://doi.org/10.1038/s41598-017-03552-9.
6. Brockman KL, Jurcisek JA, Atack JM, Srikhanta YN, Jennings MP, Bakaletz LO. 2016. ModA2 phasevarion switching in nontypeable *Haemophilus influenzae* increases the severity of experimental otitis media. J Infect Dis 214:817–824. https://doi.org/10.1093/infdis/jiw243.
7. Atack JM, Tan A, Bakaletz LO, Jennings MP, Seib KL. 2018. Phasevarions of bacterial pathogens: methylomics sheds new light on old enemies. Trends Microbiol. https://doi.org/10.1016/j.tim.2018.01.008.
8. Tan A, Atack JM, Jennings MP, Seib KL. 2016. The capricious nature of bacterial pathogens: phasevarions and vaccine development. Front Immunol 7:586. https://doi.org/10.3389/fimmu.2016.00586.
9. Gawthorne JA, Beatson SA, Srikhanta YN, Fox KL, Jennings MP. 2012. Origin of the diversity in DNA recognition domains in phasevarion associated *modA* genes of pathogenic *Neisseria* and *Haemophilus influenzae*. PLoS One 7:e32337. https://doi.org/10.1371/journal.pone.0032337.
10. Pettigrew MM, Ahearn CP, Gent JF, Kong Y, Gallo MC, Munro JB, D'Mello A, Sethi S, Tettelin H, Murphy TF. 2018. *Haemophilus influenzae* genome evolution during persistence in the human airways in chronic obstructive pulmonary disease. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.1719654115.
11. 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.
12. 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.