AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Neisseria musculi* Using Illumina and PacBio Sequencing

Eliza Thapa,[a] Jain Aluvathingal,[b] Suvarna Nadendla,[b] Aditya Mehta,[b] Hervé Tettelin,[c] Nathan J. Weyand[a,d,e]

[a]Department of Biological Sciences, Ohio University, Athens, Ohio, USA
[b]Institute for Genome Sciences (IGS), University of Maryland School of Medicine, Baltimore, Maryland, USA
[c]Department of Microbiology and Immunology, IGS, University of Maryland School of Medicine, Baltimore, Maryland, USA
[d]The Infectious and Tropical Disease Institute, Ohio University, Athens, Ohio, USA
[e]Molecular and Cellular Biology Program, Ohio University, Athens, Ohio, USA

ABSTRACT *Neisseria musculi* is an oral commensal of wild-caught mice. Here, we report the complete genome sequence of *N. musculi* strain NW831, generated using a combination of the Illumina and PacBio platforms.

Animal models that mimic human *Neisseria* infections and persistence are often compromised by strict tropism of pathogenic species for the human host (1). Using *Neisseria* species indigenous to an animal host can avoid host restriction barriers. *Neisseria musculi* is an oral commensal of healthy wild-caught mice (2). *N. musculi* can colonize the oral cavity and gastrointestinal tract of at least three lines of laboratory inbred mice for extended periods (3). The *N. musculi*-mouse model will allow studies of host association factors and their role in asymptomatic colonization of a natural host.

N. musculi was isolated from the oral cavity of wild-caught mice and stored as previously described (2). Strain NW831 was subcultured directly from the original stock of AP2031 (2). A cryopreserved stock was struck on GC medium base (GCB) agar with Kellogg's supplements I and II and incubated at 37°C with 5% CO_2 for 48 h (4). Isolated colonies were lawned on GCB agar. After 18 h, DNA was purified using an organic phenol-chloroform extraction method (5). DNA was sequenced using both Illumina and PacBio sequencing platforms. For Illumina sequencing, short-read libraries were generated with a KAPA HyperPrep kit and sequenced using 150-bp paired-end reads on the NovaSeq 6000 system. For PacBio sequencing, DNA was sheared using g-Tubes at 6,000 rpm, and libraries were constructed using the SMRTbell Express template prep kit 2.0. Libraries were size selected with Blue Pippin with a cutoff size of 7 kb and sequenced on the Sequel I system using Chemistry 3.0 (movie length, 1,200 min; read N_{50}, 9,706 bp). PacBio reads were assembled into the main chromosome using SMRTLink 8.0/HGAP 4.0 with the microbial assembly module (the length_cutoff parameter was set to −1). Polishing was accomplished with the smrtlink-release_8.0 resequencing pipeline using the Arrow algorithm and then with Pilon 1.21 (6) using Illumina reads. The single contig obtained was trimmed of overlaps using Minimus 2 (7) and manually rotated to place the *dnaA* gene at nucleotide 4. Illumina reads were assembled separately using SPAdes 3.11.1 (8) to obtain the plasmid assembly. Default parameters were used for all software unless otherwise noted. Other features of the two sequencing methods are provided in Table 1.

The complete circular genome and plasmid were each assembled into single contigs with lengths of 2,928,421 bp (53% G+C content) and 4,650 bp (48% G+C content), respectively. DNA uptake sequences are hallmark repetitive elements in *Neisseria* species that facilitate neisserial transformation (9). There were 3,996 10-mer (GCCGTCTGAA) and 2,661 12-mer (AGGCCGTCTGAA) DNA uptake sequences in the genome and 0 in the

Citation Thapa E, Aluvathingal J, Nadendla S, Mehta A, Tettelin H, Weyand NJ. 2021. Complete genome sequence of *Neisseria musculi* using Illumina and PacBio sequencing. Microbiol Resour Announc 10:e00452-21. https://doi.org/10.1128/MRA.00452-21.

Editor Julia A. Maresca, University of Delaware

Copyright © 2021 Thapa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Nathan J. Weyand, weyand@ohio.edu.

Received 5 May 2021
Accepted 18 May 2021
Published 10 June 2021

TABLE 1 Further sequencing information for the *N. musculi* genome

Feature	Datum
Illumina sequencing	
SRA accession no.	SRX9090710
No. of reads	8,512,816
Coverage (×)	439
PacBio sequencing	
SRA accession no.	SRX9090709
No. of reads	209,605
Coverage (×)	603

plasmid. The genome and plasmid sequences were annotated using the Institute for Genome Sciences (IGS) annotation pipeline and by NCBI using PGAP 4.13 (10, 11). The IGS pipeline identified 2,787 genes (2,778 genes in the chromosome and 9 genes in the plasmid), 2,721 protein-coding sequences, 4 rRNA operons, and 54 tRNAs in the genome.

Data availability. The complete genome sequence of *N. musculi* has been deposited in GenBank under the accession numbers CP060414.1 (for the genome) and CP060415.1 (plasmid), BioProject number PRJNA656535, and BioSample number SAMN15791842.

ACKNOWLEDGMENTS

We are grateful for financial support from the Infectious and Tropical Disease Institute, the Molecular and Cellular Biology Graduate Program, and the Department of Biological Sciences, Ohio University.

REFERENCES

1. Weyand NJ, Wertheimer AM, Hobbs TR, Sisko JL, Taku NA, Gregston LD, Clary S, Higashi DL, Biais N, Brown LM, Planer SL, Legasse AW, Axthelm MK, Wong SW, So M. 2013. *Neisseria* infection of rhesus macaques as a model to study colonization, transmission, persistence, and horizontal gene transfer. Proc Natl Acad Sci U S A 110:3059–3064. https://doi.org/10.1073/pnas.1217420110.
2. Weyand NJ, Ma M, Phifer-Rixey M, Taku NA, Rendón MA, Hockenberry AM, Kim WJ, Agellon AB, Biais N, Suzuki TA, Goodyer-Sait L, Harrison OB, Bratcher HB, Nachman MW, Maiden MCJ, So M. 2016. Isolation and characterization of *Neisseria musculi* sp. nov., from the wild house mouse. Int J Syst Evol Microbiol 66:3585–3593. https://doi.org/10.1099/ijsem.0.001237.
3. Ma M, Powell DA, Weyand NJ, Rhodes KA, Rendón MA, Frelinger JA, So M. 2018. A natural mouse model for *Neisseria* colonization. Infect Immun 86:e00839-17. https://doi.org/10.1128/IAI.00839-17.
4. Kellogg DS, Jr, Peacock WL, Jr, Deacon WE, Brown L, Pirkle DI. 1963. *Neisseria gonorrhoeae*. I. Virulence genetically linked to clonal variation. J Bacteriol 85:1274–1279. https://doi.org/10.1128/JB.85.6.1274-1279.1963.
5. Sambrook J, Russell DW. 2006. Purification of nucleic acids by extraction with phenol:chloroform. CSH Protoc 2006:pdb.prot4455. https://doi.org/10.1101/pdb.prot4455.
6. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
7. Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. BMC Bioinformatics 8:64. https://doi.org/10.1186/1471-2105-8-64.
8. 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.
9. Graves JF, Biswas GD, Sparling PF. 1982. Sequence-specific DNA uptake in transformation of *Neisseria gonorrhoeae*. J Bacteriol 152:1071–1077.
10. Galens K, Orvis J, Daugherty S, Creasy HH, Angiuoli S, White O, Wortman J, Mahurkar A, Giglio MG. 2011. The IGS standard operating procedure for automated prokaryotic annotation. Stand Genomic Sci 4:244–251. https://doi.org/10.4056/sigs.1223234.
11. 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.