AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of Seven *Neisseria gonorrhoeae* Clinical Isolates from Mucosal and Disseminated Gonococcal Infections

Freda E.-C. Jen,[a] John M. Atack,[a] Jennifer L. Edwards,[b,c] Michael P. Jennings[a]

[a]Institute for Glycomics, Griffith University, Brisbane, Queensland, Australia
[b]The Center for Microbial Pathogenesis, The Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, Ohio, USA
[c]The Department of Pediatrics, The Ohio State University, Columbus, Ohio, USA

ABSTRACT *Neisseria gonorrhoeae* is a Gram-negative bacterium that causes the sexually transmitted infection gonorrhea. *N. gonorrhoeae* has progressively developed resistance to all currently prescribed antibiotics, and no vaccine is available. Here, we report the closed, completed, annotated genome sequences for seven *N. gonorrhoeae* strains obtained by single-molecule real-time (SMRT) long-read genome sequencing.

Neisseria gonorrhoeae causes the sexually transmitted disease gonorrhea, which is currently one of the most common bacterial infectious diseases worldwide. Gonorrhea commonly presents as urethritis in men and cervicitis in women. Initially, gonorrhea could be easily treated with penicillin; however, it has since developed resistance to each successive recommended treatment, placing a burden on health care systems and threatening to breach last-line antibiotic treatment (1). No *N. gonorrhoeae* vaccines are available (2). Whole-genome sequencing of *N. gonorrhoeae* will provide a useful tool for unraveling the pathogenesis of this important bacterial species. There are many complete *N. gonorrhoeae* genome assemblies in the NCBI Reference Sequence Database; however, most of these were obtained using short-read Illumina sequencing. One limitation of short read lengths is that many repetitive features such as simple DNA sequence repeats (SSRs) and gene duplications may be lost during automated assembly (3). There are at least 36 translational, phase-variable genes in *N. gonorrhoeae* (4). In addition, *N. gonorrhoeae* contains 19 copies of silent, variable *pilS* genes, which can recombine with the pilin expression gene (5). Therefore, long-read sequencing (e.g., single-molecule real-time [SMRT]) is important for obtaining closed, complete genome sequences for *N. gonorrhoeae* pathogenesis research. Here, we used SMRT sequencing to sequence seven *N. gonorrhoeae* strains, 1291 (6), MS11 (7), O1G1370 (8, 9), 88G285 (8, 9), O2D156 (8), 98D159 (8, 9), and SK92-679 (10), isolated from mucosal and disseminated gonococcal infections, and report their closed, annotated whole-genome sequences. The improved synteny of these genome sequences will be useful for studying phase-variable and duplicate genes in *N. gonorrhoeae*.

N. gonorrhoeae strains were grown on GC agar supplemented with 1% IsoVitaleX (BD BBL) at 37°C and 5% CO_2 overnight and subcultured for 4 h. Cells were harvested from the plates, and genomic DNA was prepared using the GenElute kit (Sigma-Aldrich); PacBio long-read sequencing was carried out at SNPsaurus (Eugene, OR). SMRTbell libraries were prepared using the Express template prep kit 2.0 according to the manufacturer's protocol (Pacific Biosciences, CA). The samples were pooled into a single multiplexed library and size selected using Sage Sciences' BluePippin (BP) system according to the manufacturer's recommendations, with the 0.75% DF marker S1 high-pass 6 kb to 10 kb v3 run protocol and S1 marker. A size selection cutoff of 8,000 (BP start value) was used. The size-selected SMRTbell library for each strain was annealed and bound according to the SMRT Link setup, pooled, and sequenced using Sequel II chemistry v1.0 at SNPsaurus. The raw reads

Citation Jen FE-C, Atack JM, Edwards JL, Jennings MP. 2021. Complete genome sequences of seven *Neisseria gonorrhoeae* clinical isolates from mucosal and disseminated gonococcal infections. Microbiol Resour Announc 10:e00734-21. https://doi.org/10.1128/MRA.00734-21.

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

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

Address correspondence to Jennifer L. Edwards, jennifer.edwards@nationwidechildrens.org, or Michael P. Jennings, m.jennings@griffith.edu.au.

Received 19 July 2021
Accepted 9 October 2021
Published 28 October 2021

TABLE 1 Summary of information for the closed annotated genome and plasmid sequences for seven strains of *N. gonorrhoeae*

Strain	Type of infection[a]	Genome size (bp)	Genome coverage (×)	Total no. of reads (bases)	N_{50} (bp)	GC content (%)	No. of genes	No. of CDS[b]	GenBank accession no.	SRA accession no.	Plasmid found and sequenced	Plasmid homology:	
												%	With (plasmid name [GenBank accession no.])
1291	MI	2,177,032	275	635,585,034	20,140	52.6	2,276	2,208	CP078119	SRX11344351	Plasmid 1	100	WHO_O plasmid 4 (LT592149.1)
MS11	MI	2,234,079	289	669,009,143	20,168	52.4	2,350	2,279	CP078118	SRX11344352			
O1G1370	MI	2,215,052	167	424,231,185	19,999	52.4	2,327	2,256	CP078115	SRX11344355	Plasmid 1	99	WHO_W plasmid 2 (LT592164.1)
88G285	DGI	2,174,189	127	319,137,608	18,991	52.6	2,254	2,183	CP078116	SRX11344354	Plasmid 1	99	WHO_O plasmid 3 (LT592148.1)
O2D156	DGI	2,165,933	250	571,663,940	20,852	52.6	2,264	2,193	CP078113	SRX11344357			
98D159	DGI	2,172,572	168	403,272,244	19,828	52.6	2,271	2,200	CP078114	SRX11344356	Plasmid 1	99	WHO_M plasmid 4 (LT591907.1)
SK92-679	DGI	2,173,187	91	219,520,879	19,212	52.6	2,265	2,194	CP078117	SRX11344353	Plasmid 1	100	WHO_G plasmid 2 (LT591899.1)

[a]MI, mucosal infection; DGI, disseminated gonococcal infection.
[b]CDS, coding DNA sequences.

were converted to FASTA format using SAMtools (11). Flye v2.8 (12) was used to assemble and polish the sequenced genomes. The assembly quality was assessed using BUSCO v3 (13). Default parameters were used for all software unless otherwise specified. An average coverage of ~195-fold was obtained. The assembled sequences were annotated using the Prokaryotic Genome Annotation Pipeline (PGAP) during NCBI GenBank submission of the closed genome sequences (14). Information for each strain/genome/plasmid is summarized in Table 1.

Data availability. The genome sequences and whole-genome sequencing (WGS) reads have been deposited at NCBI. The accession numbers for the closed genome sequences and the raw data are provided in Table 1. The master record for the WGS reads and closed annotated genome sequences can be found at NCBI under BioProject accession number PRJNA743132.

ACKNOWLEDGMENTS

We thank SNPsaurus (Eugene, OR) for the PacBio SMRT genome sequencing and assembly. We also thank Hank Seifert and Joe Dillard for providing strain SK92-679, Michael Apicella for providing strains 1291 and MS11, and John Tapsall (deceased) for providing strains O1G137, 88G285, O2D156, and 98D159.

This work is supported by National Institutes of Health (NIH) grant number R01AI134848 (to J.L.E. and M.P.J.) and an NHMRC principal research fellowship (1138466 to M.P.J.) and Ideas grant (2001210 to F.E.-C.J.).

REFERENCES

1. Unemo M, Shafer WM. 2014. Antimicrobial resistance in Neisseria gonorrhoeae in the 21st century: past, evolution, and future. Clin Microbiol Rev 27:587–613. https://doi.org/10.1128/CMR.00010-14.
2. Edwards JL, Jennings MP, Seib KL. 2018. Neisseria gonorrhoeae vaccine development: hope on the horizon? Curr Opin Infect Dis 31:246–250. https://doi.org/10.1097/QCO.0000000000000450.
3. Alkan C, Sajjadian S, Eichler EE. 2011. Limitations of next-generation genome sequence assembly. Nat Methods 8:61–65. https://doi.org/10.1038/nmeth.1527.
4. Zelewska MA, Pulijala M, Spencer-Smith R, Mahmood HA, Norman B, Churchward CP, Calder A, Snyder LAS. 2016. Phase variable DNA repeats in Neisseria gonorrhoeae influence transcription, translation, and protein sequence variation. Microb Genom 2:e000078. https://doi.org/10.1099/mgen.0.000078.
5. Sechman EV, Rohrer MS, Seifert HS. 2005. A genetic screen identifies genes and sites involved in pilin antigenic variation in Neisseria gonorrhoeae. Mol Microbiol 57:468–483. https://doi.org/10.1111/j.1365-2958.2005.04657.x.
6. Apicella MA. 1974. Antigenically distinct populations of Neisseria gonorrhoeae: isolation and characterization of the responsible determinants. J Infect Dis 130:619–625. https://doi.org/10.1093/infdis/130.6.619.
7. Swanson J. 1972. Studies on gonococcus infection. II. Freeze-fracture, freeze-etch studies on gonocci. J Exp Med 136:1258–1271. https://doi.org/10.1084/jem.136.5.1258.
8. Power PM, Ku SC, Rutter K, Warren MJ, Limnios EA, Tapsall JW, Jennings MP. 2007. The phase-variable allele of the pilus glycosylation gene pglA is not strongly associated with strains of Neisseria gonorrhoeae isolated from patients with disseminated gonococcal infection. Infect Immun 75:3202–3204. https://doi.org/10.1128/IAI.01501-06.
9. Australian Gonococcal Surveillance Programme. 2005. Annual report of the Australian Gonococcal Surveillance Programme, 2004. Commun Dis Intell Q Rep 29:137–142.
10. Dillard JP, Seifert HS. 2001. A variable genetic island specific for Neisseria gonorrhoeae is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. Mol Microbiol 41:263–277. https://doi.org/10.1046/j.1365-2958.2001.02520.x.
11. Ramirez-Gonzalez RH, Bonnal R, Caccamo M, Maclean D. 2012. Bio-samtools: Ruby bindings for SAMtools, a library for accessing BAM files containing high-throughput sequence alignments. Source Code Biol Med 7:6. https://doi.org/10.1186/1751-0473-7-6.
12. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
13. Waterhouse RM, Seppey M, Simao FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol Biol Evol 35:543–548. https://doi.org/10.1093/molbev/msx319.
14. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105.