AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Streptococcus pneumoniae* Serotype 19F Strain EF3030

R. Junges,[a] M. Maienschein-Cline,[b] D. A. Morrison,[c] F. C. Petersen[a]

[a]Institute of Oral Biology, Faculty of Dentistry, University of Oslo, Oslo, Norway
[b]Research Informatics Core, Research Resources Center, University of Illinois at Chicago, Chicago, Illinois, USA
[c]Department of Biological Sciences, College of Liberal Arts and Sciences, University of Illinois at Chicago, Chicago, Illinois, USA

ABSTRACT We report the complete genome sequence of *Streptococcus pneumoniae* EF3030, a serotype 19F isolate that colonizes the nasopharynx of mice while being mostly noninvasive. Such attributes make this strain highly attractive in pneumococcal carriage studies. The availability of its complete genomic sequence is likely to advance studies in the field.

Streptococcus pneumoniae is the causative agent of important invasive and noninvasive human diseases (1). These morbidities cause a significant financial burden and many mortalities every year, especially in children under 5 years of age (2, 3). The pneumococcal strain EF3030 of the serotype 19F was isolated from a patient with otitis media (4, 5). One of the most interesting features of this strain is its suitability for animal models of pneumococcal disease, as it can colonize the upper respiratory tract for weeks while rarely causing bacteremia (6–10). Such lack of virulence in mice improves the study of host-*S. pneumoniae* interactions, which are particularly relevant in the assessment of immunization strategies against pneumococcal carriage and disease (11–15). In addition, EF3030 forms dense biofilms and is compatible in models of mixed infection, facilitating an investigation of pneumococcal strain behavior, including the study of natural transformation and other forms of horizontal gene transfer (16).

The *S. pneumoniae* strain EF3030 was grown in Todd-Hewitt broth supplemented with 5% yeast extract (THY) until late-log phase, and DNA was extracted using the Quick-DNA fungal/bacterial microprep kit (Zymo Research), according to the manufacturer's protocol. Library preparation utilized the Illumina Nextera XT kit. Sequencing was performed on an Illumina NextSeq platform with 2 × 150-bp reads and an Oxford Nanopore MinION device with a MasterPure complete DNA purification kit (Epicentre, Biosearch Technologies) and the Ligation sequencing kit 1D. Nanopore reads were processed for base calling using Albacore v2.1.10 (17). Reads were assembled using Canu v1.5 (18), which yielded one contig with a total sequence length of 2,142,815 bases. Initial error correction of the Nanopore data assembly was performed using Minimap2 v2.6 (19) and Racon (20), with additional polishing using Illumina data, mapping reads to the contig using BWA-MEM v0.7.15, and fixing single-nucleotide polymorphism (SNP) and indel errors (21). Further analysis of the contig revealed a 41,194-base segment that was duplicated on either end of the contig, which prevented circularization of the genome. To confirm circularization, we trimmed one copy of this duplicated sequence to create a trimmed genome contig of 2,101,618 bases in length. We then remapped reads with BWA-MEM v0.7.15 (21), extracted read pairs where either end mapped within 500 bp of either end of the contig, and performed an assembly on those reads using SPAdes (22). This yielded a single circularization contig of 947 bases in length. Alignment to the trimmed genome contig with nucmer (23) revealed an overlap of the circularization contig with either end of the genome contig, confirming

Citation Junges R, Maienschein-Cline M, Morrison DA, Petersen FC. 2019. Complete genome sequence of *Streptococcus pneumoniae* serotype 19F strain EF3030. Microbiol Resour Announc 8:e00198-19. https://doi.org/10.1128/MRA.00198-19.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Address correspondence to D. A. Morrison, damorris@uic.edu.

Received 3 March 2019
Accepted 16 April 2019
Published 9 May 2019

that the trimmed genome contig was circular. We replaced the aligned sequences from the trimmed genome contig with the bases that overlapped with the circularization contig, which added 3 bases to the total genome size for a final circular genome of 2,101,621 bases in length. Finally, we rotated the genome to match the circularized contig to *Streptococcus pneumoniae* R6 (NCBI RefSeq accession no. NC_003098) with nucmer (23), and the start position of the genome was set as the origin of replication locus upstream of the *dnaA* gene (24). Gene annotation was performed following the NCBI Prokaryotic Genome Annotation Pipeline revision 4.7 (25).

The complete genome of *S. pneumoniae* EF3030 presents a GC content of 39.8% and 2,222 genes, including 2,149 coding sequences (CDS), 73 RNA-coding genes (4 complete rRNA operons, 58 tRNAs, and 3 noncoding RNAs [ncRNAs]), and 210 pseudogenes. The hybrid assembly filled the 85 gaps in the recent EF3030 draft genome (26). Newly identified genes included the capsule locus that contains the genes specific to the strain EF3030 serotype 19F (capsule locus starts at position ~306000). Among these are the glycosyltransferase genes *wchO* (EF3030_01700), *wchP* (EF3030_01705), and *wchQ* (EF3030_01710), the polymerase gene *wzy* (EF3030_01715), the flippase gen *wzx* (EF3030_01720), and the *mnaA* and rhamnose pathway genes *rmlABCD* (EF3030_01725 to EF3030_01745). Altogether, we expect that the availability of the complete genome sequence for *S. pneumoniae* EF3030 will facilitate the genetic manipulation of this strain and the further study of pneumococcal colonization and disease.

Data availability. The nucleotide sequence of the *S. pneumoniae* EF3030 genome is deposited in NCBI GenBank under accession no. CP035897, and the raw reads are available in the Sequence Read Archive with BioProject no. PRJNA521678.

ACKNOWLEDGMENTS

We thank Jeremy S. Brown for sending the *S. pneumoniae* strain EF3030 and David W. Dyer for thoughtfully revising the manuscript and genome sequence. Library preparation and sequencing were performed at the University of Illinois at Chicago Sequencing Core (UICSQC).

This study was partially funded by the International Partnership for Outstanding Education, Research, and Innovation (INTPART) grant no. 274867.

REFERENCES

1. Henriques-Normark B, Tuomanen EI. 2013. The pneumococcus: epidemiology, microbiology, and pathogenesis. Cold Spring Harb Perspect Med 3:a010215. https://doi.org/10.1101/cshperspect.a010215.
2. Balsells E, Dagan R, Yildirim I, Gounder PP, Steens A, Muñoz-Almagro C, Mameli C, Kandasamy R, Givon Lavi N, Daprai L, van der Ende A, Trzciński K, Nzenze SA, Meiring S, Foster D, Bulkow LR, Rudolph K, Valero-Rello A, Ducker S, Vestrheim DF, von Gottberg A, Pelton SI, Zuccotti G, Pollard AJ, Sanders EAM, Campbell H, Madhi SA, Nair H, Kyaw MH. 2018. The relative invasive disease potential of Streptococcus pneumoniae among children after PCV introduction: a systematic review and meta-analysis. J Infect 77:368–378. https://doi.org/10.1016/j.jinf.2018.06.004.
3. O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, Lee E, Mulholland K, Levine OS, Cherian T, Hib and Pneumococcal Global Burden of Disease Study Team. 2009. Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. Lancet 374:893–902. https://doi.org/10.1016/S0140-6736(09)61204-6.
4. Andersson B, Dahmen J, Frejd T, Leffler H, Magnusson G, Noori G, Eden CS. 1983. Identification of an active disaccharide unit of a glycoconjugate receptor for pneumococci attaching to human pharyngeal epithelial cells. J Exp Med 158:559–570. https://doi.org/10.1084/jem.158.2.559.
5. Andersson B, Eriksson B, Falsen E, Fogh A, Hanson LA, Nylen O, Peterson H, Svanborg Eden C. 1981. Adhesion of Streptococcus pneumoniae to human pharyngeal epithelial cells in vitro: differences in adhesive capacity among strains isolated from subjects with otitis media, septicemia, or meningitis or from healthy carriers. Infect Immun 32:311–317.
6. Briles DE, Crain MJ, Gray BM, Forman C, Yother J. 1992. Strong association between capsular type and virulence for mice among human isolates of *Streptococcus pneumoniae*. Infect Immun 60:111–116.
7. Balachandran P, Brooks-Walter A, Virolainen-Julkunen A, Hollingshead SK, Briles DE. 2002. Role of pneumococcal surface protein C in nasopharyngeal carriage and pneumonia and its ability to elicit protection against carriage of *Streptococcus pneumoniae*. Infect Immun 70:2526–2534. https://doi.org/10.1128/IAI.70.5.2526-2534.2002.
8. Marks LR, Parameswaran GI, Hakansson AP. 2012. Pneumococcal interactions with epithelial cells are crucial for optimal biofilm formation and colonization *in vitro* and *in vivo*. Infect Immun 80:2744–2760. https://doi.org/10.1128/IAI.00488-12.
9. Shah P, Briles DE, King J, Hale Y, Swiatlo E. 2009. Mucosal immunization with polyamine transport protein D (PotD) protects mice against nasopharyngeal colonization with *Streptococcus pneumoniae*. Exp Biol Med (Maywood) 234:403–409. https://doi.org/10.3181/0809-RM-269.
10. Blevins LK, Wren JT, Holbrook BC, Hayward SL, Swords WE, Parks GD, Alexander-Miller MA. 2014. Coinfection with *Streptococcus pneumoniae* negatively modulates the size and composition of the ongoing influenza-specific $CD8^+$ T cell response. J Immunol 193:5076–5087. https://doi.org/10.4049/jimmunol.1400529.
11. Wilson R, Cohen JM, Jose RJ, de Vogel C, Baxendale H, Brown JS. 2015. Protection against *Streptococcus pneumoniae* lung infection after nasopharyngeal colonization requires both humoral and cellular immune responses. Mucosal Immunol 8:627–639. https://doi.org/10.1038/mi.2014.95.
12. Palaniappan R, Singh S, Singh UP, Sakthivel SK, Ades EW, Briles DE, Hollingshead SK, Paton JC, Sampson JS, Lillard JW, Jr. 2005. Differential PsaA-, PspA-, PspC-, and PdB-specific immune responses in a mouse model of pneumococcal carriage. Infect Immun 73:1006–1013. https://doi.org/10.1128/IAI.73.2.1006-1013.2005.
13. Briles DE, Hollingshead SK, Paton JC, Ades EW, Novak L, van Ginkel FW,

Benjamin WH, Jr. 2003. Immunizations with pneumococcal surface protein A and pneumolysin are protective against pneumonia in a murine model of pulmonary infection with *Streptococcus pneumoniae*. J Infect Dis 188:339–348. https://doi.org/10.1086/376571.
14. Palaniappan R, Singh S, Singh UP, Singh R, Ades EW, Briles DE, Hollingshead SK, Royal W, III, Sampson JS, Stiles JK, Taub DD, Lillard JW, Jr. 2006. CCL5 modulates pneumococcal immunity and carriage. J Immunol 176:2346–2356. https://doi.org/10.4049/jimmunol.176.4.2346.
15. Keller LE, Luo X, Thornton JA, Seo KS, Moon BY, Robinson DA, McDaniel LS. 2015. Immunization with pneumococcal surface protein K of nonencapsulated Streptococcus pneumoniae provides protection in a mouse model of colonization. Clin Vaccine Immunol 22:1146–1153. https://doi.org/10.1128/CVI.00456-15.
16. Marks LR, Reddinger RM, Hakansson AP. 2012. High levels of genetic recombination during nasopharyngeal carriage and biofilm formation in Streptococcus pneumoniae. mBio 3:e00200-12. https://doi.org/10.1128/mBio.00200-12.
17. Oxford Nanopore Technologies. 2016. Albacore, v2.1.10. https://mirror.oxfordnanoportal.com/software/analysis/ont_albacore-2.3.1-cp36-cp36m-manylinux1_x86_64.whl.
18. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
19. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
20. Vaser R, Sovic I, Nagarajan N, Sikic M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.
21. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN]. https://arxiv.org/abs/1303.3997.
22. 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.
23. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. Genome Biol 5:R12. https://doi.org/10.1186/gb-2004-5-2-r12.
24. Gasc AM, Giammarinaro P, Richter S, Sicard M. 1998. Organization around the *dnaA* gene of *Streptococcus pneumoniae*. Microbiology 144:433–439. https://doi.org/10.1099/00221287-144-2-433.
25. 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.
26. Scott EJ, Jr, Luke-Marshall NR, Campagnari AA, Dyer DW. 2019. Draft genome sequence of pediatric otitis media isolate Streptococcus pneumoniae strain EF3030, which forms in vitro biofilms that closely mimic in vivo biofilms. Microbiol Resour Announc 8:e00114-18. https://doi.org/10.1128/MRA.01114-18.