



# Draft Genome Sequence of Pediatric Otitis Media Isolate *Streptococcus pneumoniae* Strain EF3030, Which Forms *In Vitro* Biofilms That Closely Mimic *In Vivo* Biofilms

Edgar J. Scott II,<sup>a</sup> Nicole R. Luke-Marshall,<sup>b</sup> Anthony A. Campagnari,<sup>b</sup> David W. Dyer<sup>a</sup>

<sup>a</sup>Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma, USA

<sup>b</sup>Department of Microbiology and Immunology, State University of New York at Buffalo, Buffalo, New York, USA

**ABSTRACT** Here, we report the draft genome sequence of *Streptococcus pneumoniae* EF3030, a pediatric otitis media isolate active in biofilm assays of epithelial colonization. The final draft assembly included 2,209,198 bp; the annotation predicted 2,120 coding DNA sequences (CDSs), 4 complete rRNA operons, 58 tRNAs, 3 noncoding RNAs (ncRNAs), and 199 pseudogenes.

*Streptococcus pneumoniae* (the pneumococcus), an encapsulated Gram-positive diplococcus, is one of the most common colonizers and opportunistic pathogens of the human upper respiratory tract. It is estimated that 95% of children by the age of two are colonized on the nasopharyngeal mucosa by at least 1 of over 90 *S. pneumoniae* serotypes, which persist asymptotically in healthy individuals into adulthood (1–4). *S. pneumoniae* is the primary etiologic agent of otitis media and secondary bacterial pneumonia following viral infection and causes severe invasive disease, including acute pneumonia, meningitis, and sepsis (2, 5). Pneumococcal biofilm formation *in vivo* contributes to immune evasion, antibiotic resistance, and persistence and serves as a reservoir for initiating local and invasive disease (recently reviewed in reference 6). *S. pneumoniae* strain EF3030, a capsular serotype 19F, is a pediatric otitis media isolate that is an efficient colonizer of murine model systems (6–11). In addition, *in vitro* *S. pneumoniae* EF3030 biofilms on an epithelial substratum closely mimic *in vivo* biofilms that form during asymptomatic colonization (6, 8). On the human upper respiratory mucosa, polymicrobial interactions within the microbiome likely impact the mechanisms of disease induction by the pneumococcus and other cocolonizing microbes. Studies designed to delineate these complex interactions are warranted and the whole-genome sequence of *S. pneumoniae* EF3030 will contribute to the identification and characterization of bacterial factors critical for these processes.

Whole-genome sequencing was performed on an Illumina MiSeq instrument, which generated 1,646,744 paired-end reads, with an average read length of 151 bp (219× coverage). An initial reference assembly of the paired-end reads was first performed against *S. pneumoniae* R6, from which multilocus sequence type (MLST) (12) loci (*aroE*, *gdh*, *gki*, *recP*, *spi*, *xpt*, and *ddl*) were extracted. These loci were then used in an MLST comparison to the available completed *S. pneumoniae* genomes, from which we determined that *S. pneumoniae* CGSP14 (GenBank accession number CP001033) was the most closely related strain. The raw *S. pneumoniae* EF3030 reads were then reassembled to the *S. pneumoniae* CGSP14 genome, using Bowtie 2 with the very sensitive local preset (13), SAMtools (14), BCFtools (15), and vcfutils (16). The exact code for this assembly was “samtools mpileup -uf ./spn\_cgsp14.fna sorted.bam | bcfutils call -c | vcfutils.pl vcf2fq > cns.fq.” An additional *de novo* assembly employed SOAPdenovo2 (kmer size, 80) (17) and identified seven protein-coding contigs not in the reference assembly. These were all homologous to previously identified *S. pneumoniae*

**Citation** Scott EJ, II, 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:e01114-18. <https://doi.org/10.1128/MRA.01114-18>.

**Editor** Jason Stajich, University of California, Riverside

**Copyright** © 2019 Scott et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to David W. Dyer, David-dyer@ouhsc.edu.

**Received** 8 August 2018

**Accepted** 27 November 2018

**Published** 10 January 2019

genes and consequently were already contained in the *S. pneumoniae* pan-genome (18). These data were submitted to the NCBI Prokaryotic Genome Annotation Pipeline (19) for annotation. The annotation consisted of 2,193 total genes, including 2,120 coding DNA sequences (CDSs), 73 RNA-encoding genes (4 complete rRNA operons, 58 tRNAs, and 3 noncoding RNAs [ncRNAs]), and 199 pseudogenes.

**Data availability.** The draft genome sequence has been deposited in NCBI GenBank under the accession number [CP026549](https://doi.org/10.1016/S1473-3099(04)00938-7). The raw data were deposited in the Sequence Read Archive under BioProject accession number [PRJNA432428](https://doi.org/10.1097/MCP.0b013e328352103b).

## ACKNOWLEDGMENTS

We thank Shauna L. Sauberan for the technical support.

These studies were supported by grants from the National Institutes of Health, NIDCD (awards R01DC013554 and R01DC014576), and NIGMS (award P20GM103447). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## REFERENCES

- Bogaert D, De Groot R, Hermans PW. 2004. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *Lancet Infect Dis* 4:144–154. [https://doi.org/10.1016/S1473-3099\(04\)00938-7](https://doi.org/10.1016/S1473-3099(04)00938-7).
- Charalambous BM, Leung MH. 2012. Pneumococcal sepsis and nasopharyngeal carriage. *Curr Opin Pulm Med* 18:222–227. <https://doi.org/10.1097/MCP.0b013e328352103b>.
- Gray BM, Converse GM, III, Dillon HC, Jr. 1980. Epidemiologic studies of Streptococcus pneumoniae in infants: acquisition, carriage, and infection during the first 24 months of life. *J Infect Dis* 142:923–933. <https://doi.org/10.1093/infdis/142.6.923>.
- Huebner RE, Dagan R, Porath N, Wasas AD, T M, Klugman KP. 2000. Lack of utility of serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes. *Pediatr Infect Dis J* 19:1017–1020. <https://doi.org/10.1097/00006454-200010000-00019>.
- Vernatter J, Pirofski LA. 2013. Current concepts in host-microbe interaction leading to pneumococcal pneumonia. *Curr Opin Infect Dis* 26:277–283. <https://doi.org/10.1097/QCO.0b013e3283608419>.
- Chao Y, Marks LR, Pettigrew MM, Hakansson AP. 2014. Streptococcus pneumoniae biofilm formation and dispersion during colonization and disease. *Front Cell Infect Microbiol* 4:194. <https://doi.org/10.3389/fcimb.2014.00194>.
- 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>.
- 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>.
- 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>.
- 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>.
- Li H, Handsaker B, Dahan J, Freid 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>.
- Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11:595. <https://doi.org/10.1186/1471-2105-11-595>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* 27:2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1:18.
- Azarian T, Grant LR, Arnold BJ, Hammit LL, Reid R, Santosham M, Weatherholtz R, Goklish N, Thompson CM, Bentley SD, O'Brien KL, Hanage WP, Lipsitch M. 2018. The impact of serotype-specific vaccination on phylogenetic parameters of Streptococcus pneumoniae and the pneumococcal pan-genome. *PLoS Pathog* 14:e1006966. <https://doi.org/10.1371/journal.ppat.1006966>.
- 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>.