AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of Neonatal Clinical Group B Streptococcal Isolate CJB111

Brady L. Spencer,[a] Anushila Chatterjee,[a] Breck A. Duerkop,[a] Carol J. Baker,[b,c] Kelly S. Doran[a]

[a]Department of Immunology and Microbiology, School of Medicine, University of Colorado-Anschutz Medical Center, Aurora, Colorado, USA
[b]Department of Pediatrics, McGovern Medical School, University of Texas Health Science Center, Houston, Texas, USA
[c]Children's Memorial Hermann Hospital, Houston, Texas, USA

ABSTRACT Group B *Streptococcus* (GBS) is an asymptomatic colonizer of the female reproductive tract but can cause maternal and neonatal infections and adverse pregnancy outcomes. Here, we closed the genome sequence of strain CJB111, a neonatal GBS clinical isolate from a case of late-onset bacteremia without focus (Houston, TX; 1990).

The Gram-positive beta-hemolytic bacterium *Streptococcus agalactiae* (group B *Streptococcus* [GBS]) asymptomatically colonizes the gastrointestinal and female genital tracts of healthy adults but can cause neonatal infections (pneumonia, bacteremia, meningitis [1]) and adverse pregnancy outcomes (2). Serotype V GBS isolates are emerging among adults and infants (3–6), and serotype V isolate CJB111 exhibits hypervirulence and vaginal persistence in murine models of GBS infection and colonization (7, 8).

CJB111 (ATCC BAA23) was isolated by Carol J. Baker from the blood of a female infant with late-onset sepsis on 18 July 1990, grown in Todd Hewitt broth (THB), and stored in glycerol at −90°C. The patient received intravenous ampicillin and gentamicin 1 day post onset of illness. Upon GBS isolation, therapy was switched on day 3 to intravenous penicillin (10-day treatment total). Following clean cerebrospinal fluid tests, she was diagnosed with bacteremia without focus at age 55 days and discharged without apparent sequelae. While CJB111's sequence is currently available in 155 contigs (GenBank accession no. AAJQ01000000), a closed genome sequence may ease future genomic analyses.

CJB111 was grown statically overnight at 37°C in THB, genomic DNA was purified (Gentra PureGene Yeast/Bact kit), and concentration and quality were confirmed by NanoDrop spectroscopy. The Microbial Genome Sequencing Center (MiGS; Pittsburgh, PA) performed short- and long-read sequencing (Illumina and Oxford Nanopore technologies [ONT], respectively) and *de novo* assembly. Default parameters were used except where otherwise noted. Short reads were obtained using the Illumina Nextera kit and NextSeq 550 platform (9). For ONT sequencing, libraries were prepared using kit SQK-LSK109 to the manufacturer's specifications (no DNA size selection/shearing), sequencing was performed on a MinION R9 flow cell, and base calling was performed using Guppy v4.2.2 (GPU mode) (10). Illumina paired-end reads (2 × 150 bp) and ONT long reads were provided as fastq files (Illumina: 7,410,044 reads, 989,364,400 bases, 472× coverage; ONT: 175,394 reads, 650,701,562 bases, 310× coverage, N_{50} value of 4,577 bp). bcl2fastq v2.20.0.422 was used for demultiplexing, quality control, and trimming of the Illumina reads (11) and Porechop v0.2.4 for quality trimming and removing adapters for ONT sequencing (12). Hybrid assembly via Unicycler v0.4.8 with a verbosity value of 2 (13) yielded six contigs, which were further assembled into three nonoverlapping contigs upon mapping to CJB111 contigs (AAJQ01000000) in Geneious v11.1.5 (14). The genome

Citation Spencer BL, Chatterjee A, Duerkop BA, Baker CJ, Doran KS. 2021. Complete genome sequence of neonatal clinical group B streptococcal isolate CJB111. Microbiol Resour Announc 10:e01268-20. https://doi.org/10.1128/MRA.01268-20.

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

Address correspondence to Kelly S. Doran, kelly.doran@cuanschutz.edu.

Received 3 November 2020
Accepted 21 December 2020
Published 14 January 2021

TABLE 1 Primers used in this study

PCRs flanking nonoverlapping contigs	Primer name[a]	Sequence (5′–3′)	PCR T_m (°C)	PCR product length (bp)	PCR extension time (min:s)
PCR 1	5121	TATCAATAACGATAGTATGCCCAGCG	65	4,712	2:21
	3121	TTCCAATAGGTCTTGATAGTGAGGTG			
PCR 2	5122	GTTTGTTGCAGTCGTCGTTATCTC	63	3,392	1:42
	3122	CGTCGGAATTAAATCTTGGAATACC			
PCR 3	5123	GGCATCAGGAATGATCTGATTTACAC	65	2,064	1:02
	3123	TGCCTCCCATTGGATTACTGTATAC			
PCR 4	5124	GACTCGATAGGGTATATGGTGCC	65	5,126	2:34
	3124	GGTTCGATTGCGTTACTGCG			
Sequencing primers					
PCR 1	5131	GTGACATAGATTGGAATAGGGTTAGC			
	5132	TATTCTCAGTGTCTGTGTACTACTGC			
	5133	AAATCTTGGCAGACAGTGGTTATC			
	5134	CAACAGGAGGAACCTGTAGAAGTTC			
	5135	TACAATCCATCTCTGGAATTCAC			
	3131	GTGACATAGATTGGAATAGGGTTAGC			
	3132	ATAATAAGGTGTCAGACAAACTCGC			
	3133	GGTTCGTCATTTATGAATGGTGTAC			
	3134	TTGACTATGGTTATGCTTTCAGG			
	3135	TTCTCAACCTTGATTCTCTCTTTGG			
	3136	GTGCCGTTTCAAAGGTCGCT			
	3137	CCGGGCTCGCTCCATATAGATAAG			
PCR 2	5136	TATGCTCTCATAGGTAACACCACC			
	5137	AACGATCACCTAAATTAGTACCTGC			
	5138	TCTATCTTGTTCCTGTTTCCTTG			
	5139	TTTAGGTTAGAAAGGAGATACTGCC			
	5140	TACTTCAAATGGTATGCAAGCTATGG			
	3138	GCTGAACAAGCTGCTGTTATTGC			
	3139	TTTAGTTGAGGATGCTTATCGAG			
	3140	AGTTATCTGTCTATAAGGAATGTCG			
	3141	AAGCTATGGTTGAAGCTGTTG			
PCR 3	5141	TTAAATTAACTCCTGAAGTACTCCG			
	3142	AGGTAATTTCCATTTCTCACCTGAAG			
	3143	TTTCGGCGACAATTCATTGAACTGAG			
PCR 4	5125	GTAACTAGTTATCTCTAGCCTGTAGC			
	5126	CACGAAAGCAACTTAATCCGTCG			
	5127	CCCTTGACTACATAAGTACTAACCC			
	5128	CTGTTAATAAATCAGCTCCATGAGC			
	5129	TTTCCCTTGCATTTCCCATAGACC			
	5130	GCCTATCCAATTATTCGTTTGGAG			
	3125	TTTACCTCTGTTGCATCCACAATC			
	3126	GCAAAGCAATTGTATTCCGTCTT			
	3127	AAAGTGTCGTTACCAACTCTGAAG			
	3128	AAATTATGAATCAGGCATGCTCCTGG			
	3129	AATAAAGCCTGAAACCAGTTCAGAG			
	3130	CATCACTCTGGCCTCTATTATTT			

[a] Primer names beginning with 5 indicate forward primers. Primer names beginning with 3 indicate reverse primers.

sequence was closed via PCR using primers flanking the nonoverlapping contigs (Table 1) and Phusion high-fidelity polymerase/buffer (New England Biolabs) under the following cycling conditions on a Bio-Rad T100 thermal cycler: 98°C, 2-min hot start; 34 cycles (98°C, 10^{seconds}; T_m °C, 20^{seconds}; 72°C, 30- second extension/kb); and 72°C, 10-min extension. Purified PCR amplicons (Qiagen) were Sanger sequenced using Applied Biosystems 3730/3500xl genetic analyzers, yielding 2× sequencing in both directions. Contigs were assembled *de novo*, overlapping ends were trimmed, and the genome sequence was rotated manually in Geneious v11.1.5 to start with *dnaA*.

The CJB111 sequence was deposited at GenBank as one circular contig (2,093,987 bp;

GC content, 35.52%). BUSCO_v1 and CheckM v1.0.18 confirmed the genome completeness (15, 16). GenBank annotated the CJB111 genome sequence using PGAP v4.13 (17).

Data availability. The CJB111 sequence is available in GenBank under accession no. CP063198. The raw sequence reads are accessible under Sequence Read Archive accession no. SRX9273111, SRX9273112, and SRX9273113; BioProject accession no. PRJNA663970; and BioSample accession no. SAMN16191206.

ACKNOWLEDGMENTS

This work was funded by NIH/NIAID F32 AI143203 to B.L.S., NIH/NINDS R01 NS116716 to K.S.D., and NIH/NIAID R01 AI141479 to B.A.D.

We acknowledge Dan Snyder and the Microbial Genome Sequencing Center (Pittsburgh, PA) for Illumina and Nanopore sequencing and Mark Farrell and the Barbara Davis Center Molecular Biology Core at University of Colorado-Anschutz Medical Campus (Aurora, CO) for Sanger sequencing.

REFERENCES

1. Verani JR, McGee L, Schrag SJ. 2010. Prevention of perinatal group B streptococcal disease: revised guidelines from CDC, 2010. MMWR Recomm Rep 59:1–32.
2. Seale AC, Bianchi-Jassir F, Russell NJ, Kohli-Lynch M, Tann CJ, Hall J, Madrid L, Blencowe H, Cousens S, Baker CJ, Bartlett L, Cutland C, Gravett MG, Heath PT, Ip M, Le Doare K, Madhi SA, Rubens CE, Saha SK, Schrag SJ, Sobanjo-ter Meulen A, Vekemans J, Lawn JE. 2017. Estimates of the burden of group B streptococcal disease worldwide for pregnant women, stillbirths, and children. Clin Infect Dis 65:S200–S219. https://doi.org/10.1093/cid/cix664.
3. A'Hearn-Thomas B, Khatami A, Randis TM, Vurayai M, Mokomane M, Arscott-Mills T, Banda FM, Mazhani T, Lepere T, Gaolebale P, Nchingane S, Chamby A, Gegick M, Suzman E, Steenhoff AP, Ratner AJ. 2019. High rate of serotype V *Streptococcus agalactiae* carriage in pregnant women in Botswana. Am J Trop Med Hyg 100:1115–1117. https://doi.org/10.4269/ajtmh.18-0847.
4. Burcham LR, Spencer BL, Keeler LR, Runft DL, Patras KA, Neely MN, Doran KS. 2019. Determinants of group B streptococcal virulence potential amongst vaginal clinical isolates from pregnant women. PLoS One 14:e0226699. https://doi.org/10.1371/journal.pone.0226699.
5. Yoon IA, Jo DS, Cho EY, Choi EH, Lee HJ, Lee H. 2015. Clinical significance of serotype V among infants with invasive group B streptococcal infections in South Korea. Int J Infect Dis 38:136–140. https://doi.org/10.1016/j.ijid.2015.05.017.
6. Flores AR, Galloway-Pena J, Sahasrabhojane P, Saldana M, Yao H, Su X, Ajami NJ, Holder ME, Petrosino JF, Thompson E, Margarit YRI, Rosini R, Grandi G, Horstmann N, Teatero S, McGeer A, Fittipaldi N, Rappuoli R, Baker CJ, Shelburne SA. 2015. Sequence type 1 group B Streptococcus, an emerging cause of invasive disease in adults, evolves by small genetic changes. Proc Natl Acad Sci U S A 112:6431–6436. https://doi.org/10.1073/pnas.1504725112.
7. Patras KA, Doran KS. 2016. A murine model of group B *Streptococcus* vaginal colonization. J Vis Exp e54708. https://doi.org/10.3791/54708.
8. Faralla C, Metruccio MM, De Chiara M, Mu R, Patras KA, Muzzi A, Grandi G, Margarit I, Doran KS, Janulczyk R. 2014. Analysis of two-component systems in group B *Streptococcus* shows that RgfAC and the novel FspSR modulate virulence and bacterial fitness. mBio 5:e00870-14. https://doi.org/10.1128/mBio.00870-14.
9. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS One 10:e0128036. https://doi.org/10.1371/journal.pone.0128036.
10. Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol 20:129. https://doi.org/10.1186/s13059-019-1727-y.
11. Illumina. bcl2fastq: a proprietary Illumina software for the conversion of bcl files to basecalls. https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html.
12. Wick RR. 2017. Porechop. http://github.com/rrwick/Porechop.
13. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
14. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
15. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
16. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
17. 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.