genomeA
Journals.ASM.org

Draft Genome Sequences of Four Genetically Distinct Human Isolates of *Streptococcus dysgalactiae* subsp. *equisimilis*

Caitlin Evers,[a] Khushali Patel,[a] Varduhi Petrosyan,[a] Clay Morrison,[a,b] Viju Varghese,[c] Randy A. Chu,[a] Aymen Baig,[d] Erika J. Thompson,[a,c,d] Michael Chase,[b] Peter C. Hu,[a,d] Awdhesh Kalia[a]

Graduate Program in Diagnostic Genetics, School of Health Professions, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA[a]; Roche Diagnostics Corporation, Indianapolis, Indiana, USA[b]; Sequencing and Microarray Facility, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA[c]; Molecular Genetic Technology Program, School of Health Professions, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA[d]

C.E., K.P., and V.P. are joint first authors and are listed alphabetically.

***β*-Hemolytic group C and group G streptococci (GCS-GGS; *Streptococcus dysgalactiae* subsp. *equisimilis*) emerged as human pathogens in the late 1970s. We report here the draft genome sequences of four genetically distinct human strains of GCS-GGS isolated between the 1960s and 1980s. Comparative analysis of these genomes may provide a deeper understanding of GCS-GGS genome and virulence evolution.**

Received 17 August 2015 **Accepted** 19 August 2015 **Published** 1 October 2015

Citation Evers C, Patel K, Petrosyan V, Morrison C, Varghese V, Chu RA, Baig A, Thompson EJ, Chase M, Hu PC, Kalia A. 2015. Draft genome sequences of four genetically distinct human isolates of *Streptococcus dysgalactiae* subsp. *equisimilis*. Genome Announc 3(5):e01139-15. doi:10.1128/genomeA.01139-15.

Address correspondence to Awdhesh Kalia, akalia@mdanderson.org.

Large-colony-forming β-hemolytic isolates of Lancefield group C and group G Streptococci (GCS-GGS) identified as *Streptococcus dysgalactiae* subsp. *equisimilis* can infect humans and other mammals (1, 2). GCS and GGS emerged as human pathogens in the late 1970s and early 1980s and now approximate or surpass group A streptococci (GAS) as the predominant cause of invasive β-hemolytic streptococcal infection (3–5). The transfer of genes from GAS into GCS-GGS genomes via horizontal gene transfer (HGT) is common and ostensibly the most parsimonious explanation for the emergence of GCS-GGS as human pathogens (1, 6, 7). The dynamics of HGT between GAS and GCS-GGS are complicated. Some HGT events are asymmetric (aHGT), which can further be categorized as additive- or replacing-type aHGT (7); other HGT events are more typical and involve homologous recombination among orthologs, resulting in gene mosaics (8, 9). The mechanisms of replacing-type aHGT remain elusive. It is unclear (i) how replacing-type aHGT may shape diversity in the global GCS-GGS gene pool, (ii) whether aHGT dynamics differ spatially and temporally, and (iii) whether aHGT alone can account for the emergence of GCS-GGS as human pathogens.

Here, we report the draft genome sequences of four GCS-GGS isolates (Table 1). Lancefield groups of all isolates were determined by serotyping. High-quality genomic DNA was extracted using a previously described method (10) modified to include mutanolysin. Isolates were confirmed to be *S. dysgalactiae* subsp. *equisimilis* with 16S rRNA sequencing (GenBank accession numbers KP972460 to KP972463), and the genetic relatedness of the GCS-GGS isolates was determined via multilocus sequence typing (MLST) (GenBank accession numbers KT347549 to KT347565) (Table 1) (11, 12). Whole-genome shotgun sequencing was done with the Roche 454 GS Jr+ system. *De novo* assembly was performed with Newbler version 3.0 using default settings; contigs <200 bp were not included. The genome statistics are listed in Table 1.

(This work was presented in part at the 115th General Meeting of the American Society for Microbiology, New Orleans, LA, 30 May to 2 June 2015.)

Assembled genomes were annotated using NCBI PGAP version 2.1 (rev. 462191) (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/) and RASTtk pipelines (13) in conjunction with Blast2GO (14). The 16S rRNA sequences in the draft genome were identified using RNAmmer version 1.2 (15). Automated MLST using the draft genomes was performed with MLST version 1.8 (16). The pre- and postgenome sequencing of 16S rRNA and

TABLE 1 Strain descriptions and draft genome statistics

Strain	Serotype/yr	Disease	MLST genotype	Accession no.	No. of contigs >200 bp	Contig N50 (Kb)	Average coverage (×)	Estimated genome size (Mb)
UT-5345	C/1983	Bacteremia	ST-53	LAKV00000000	91	50.369	24	2.2
UT-SS1069	C/1974	Unknown	ST-3	LAKS00000000	86	52	16	2.01
UT-5354	G/1980s	Bacteremia	ND[a]	LAKU00000000	75	75.47	21	2.07
UT-SS957	C/1969	Unknown	ST-51	LAKT00000000	58	91.808	50	2.03

[a] ND, not determined; harbors a unique *xpt* allele not present in the *S. dysgalactiae* subsp. *equisimilis* MLST database and 99% identical to allele *xpt28* from GCS-GGS and 99% identical to allele *xpt29* from GAS.

MLST sequence data were 100% concordant and confirmed the lack of any contaminating DNA in the genomic DNA preparations or in sequencing libraries.

Nucleotide sequence accession numbers. The draft genome sequences have been deposited as whole-genome shotgun projects at DDBJ/EMBL/GenBank under the accession numbers listed in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health through Cancer Center Support grant P30CA016672 and used the Sequencing and Microarray Facility at the University of Texas MD Anderson Cancer Center.

We thank Dean Shirley Richmond for supporting the Project-Based Integrated Curriculum Development Initiative (PICDIn); laboratory exercises implemented in PICDIn formed the basis for this report. A.K. conceived and designed the project; C.E., K.P., V.P., R.A.C., A.B., V.V., and C.M. performed the experiments; C.E., K.P., V.P., and A.K. performed the analysis; C.M., E.J.T., M.C., P.C.H., and A.K. contributed reagents/materials/analysis tools; and A.K. wrote the paper. C.M. and M.C. are employees of Roche Diagnostics Corporation, USA.

REFERENCES

1. **Brandt CM, Spellerberg B**. 2009. Human infections due to *Streptococcus dysgalactiae* subspecies *equisimilis*. Clin Infect Dis **49:**766–772. http://dx.doi.org/10.1086/605085.
2. **Kasuya K, Yoshida E, Harada R, Hasegawa M, Osaka H, Kato M, Shibahara T**. 2014. Systemic *Streptococcus dysgalactiae* subspecies *equisimilis* infection in a Yorkshire pig with severe disseminated suppurative meningoencephalomyelitis. J Vet Med Sci **76:**715–718. http://dx.doi.org/10.1292/jvms.13-0526.
3. **Broyles LN, Van Beneden C, Beall B, Facklam R, Shewmaker PL, Malpiedi P, Daily P, Reingold A, Farley MM**. 2009. Population-based study of invasive disease due to β-hemolytic streptococci of groups other than A and B. Clin Infect Dis **48:**706–712. http://dx.doi.org/10.1086/597035.
4. **Schwartz IS, Keynan Y, Gilmour MW, Dufault B, Lagacé-Wiens P**. 2014. Changing trends in β-hemolytic streptococcal bacteremia in Manitoba, Canada: 2007–2012. Int J Infect Dis **28:**211–213. http://dx.doi.org/10.1016/j.ijid.2014.03.1376.
5. **Sylvetsky N, Raveh D, Schlesinger Y, Rudensky B, Yinnon AM**. 2002. Bacteremia due to beta-hemolytic streptococcus group G: increasing incidence and clinical characteristics of patients. Am J Med **112:**622–626. http://dx.doi.org/10.1016/S0002-9343(02)01117-8.
6. **Kalia A, Bessen DE**. 2003. Presence of streptococcal pyrogenic exotoxin A and C genes in human isolates of group G streptococci. FEMS Microbiol Lett **219:**291–295. http://dx.doi.org/10.1016/S0378-1097(03)00022-3.
7. **Choi SC, Rasmussen MD, Hubisz MJ, Gronau I, Stanhope MJ, Siepel A**. 2012. Replacing and additive horizontal gene transfer in streptococcus. Mol Biol Evol **29:**3309–3320. http://dx.doi.org/10.1093/molbev/mss138.
8. **Kalia A, Bessen DE**. 2004. Natural selection and evolution of streptococcal virulence genes involved in tissue-specific adaptations. J Bacteriol **186:**110–121. http://dx.doi.org/10.1128/JB.186.1.110-121.2004.
9. **Sriprakash KS, Hartas J**. 1996. Lateral genetic transfers between group A and G streptococci for M-like genes are ongoing. Microb Pathog **20:**275–285. http://dx.doi.org/10.1006/mpat.1996.0026.
10. **Kalia A, Rattan A, Chopra P**. 1999. A method for extraction of high-quality and high-quantity genomic DNA generally applicable to pathogenic bacteria. Anal Biochem **275:**1–5. http://dx.doi.org/10.1006/abio.1999.4259.
11. **Ahmad Y, Gertz RE, Li Z, Sakota V, Broyles LN, Van Beneden C, Facklam R, Shewmaker PL, Reingold A, Farley MM, Beall BW**. 2009. Genetic relationships deduced from *emm* and multilocus sequence typing of invasive *Streptococcus dysgalactiae* subsp. *equisimilis* and *S. canis* recovered from isolates collected in the United States. J Clin Microbiol **47:**2046–2054. http://dx.doi.org/10.1128/JCM.00246-09.
12. **McMillan DJ, Bessen DE, Pinho M, Ford C, Hall GS, Melo-Cristino J, Ramirez M**. 2010. Population genetics of *Streptococcus dysgalactiae* subspecies *equisimilis* reveals widely dispersed clones and extensive recombination. PLoS One **5:**e11741. http://dx.doi.org/10.1371/journal.pone.0011741.
13. **Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA III, Stevens R, Vonstein V, Wattam AR, Xia F**. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep **5:**8365. http://dx.doi.org/10.1038/srep08365.
14. **Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A**. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res **36:**3420–3435. http://dx.doi.org/10.1093/nar/gkn176.
15. **Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW**. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.
16. **Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, Lund O**. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol **50:**1355–1361. http://dx.doi.org/10.1128/JCM.06094-11.