



Draft Genome Sequences of Six Skin Isolates of *Streptococcus pyogenes*

Zhong Liang,^a Melissa Stephens,^b Victoria A. Ploplis,^{a,c} Shaun W. Lee,^d Francis J. Castellino^{a,c}

^aW.M. Keck Center for Transgene Research, University of Notre Dame, Notre Dame, Indiana, USA

^bGenomics and Bioinformatics Core Facility, University of Notre Dame, Notre Dame, Indiana, USA

^cDepartment of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana, USA

^dDepartment of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, USA

ABSTRACT Whole-genome shotgun sequences and bottom-up assembly of contigs of six skin isolates of *Streptococcus pyogenes*, viz., NS88.3 (*emm98.1*), NS223 (*emm91*), NS455 (*emm52*), SS1448 (*emm86.2*), SS1572 (*emm223*), and SS1574 (*emm224*), are presented here. All contigs were annotated, and the gene arrangements and the inferred proteins were consistent with a pattern D classification.

Group A *Streptococcus pyogenes* (GAS) is a beta-hemolytic human-pathogenic bacterium. It causes a spectrum of diseases ranging from mild infections to life-threatening tissue destruction and organ failure. This microbe is able to infect multiple human niches, from epithelial surfaces to deeper tissues. Superficial GAS infections at times become highly invasive and result in lethal outcomes if medical intervention does not occur (1).

Herein, we report the newly sequenced bottom-up-assembled contigs of six strains of GAS isolated from skin for future investigations of the pathogenesis associated with skin infection. Isolates NS88.3, NS223, and NS455 were obtained from M. J. Walker (Queensland, Australia), and isolates SS1448, SS1572, and SS1574 were obtained from the Centers for Disease Control and Prevention.

Sequencing libraries were prepared with the NEBNext Ultra II DNA library prep kit (New England BioLabs) and subsequently sequenced across two lanes of an Illumina MiSeq platform using the MiSeq reagent Nano kit version 2 (500 cycles) with 251-bp paired-end reads (Illumina). An average for each genome of 530,000 paired reads passing filtering were produced, yielding ~135 million bp of data for each of the six GAS strains. Quality control on the library pool was obtained using the Kapa Biosystems quantitative PCR (qPCR) and Agilent Bioanalyzer DNA 7500 assays. Base calling was done by Illumina Real-Time Analysis version 2. The sequencing generated an average of 50-fold coverage for each library, and the sequences were assembled *de novo* using Velvet version 1.2.10 (2). Gene predictions and annotations were obtained using Glimmer 3.0 (3) and RAST (4).

The assembled genomes have an average length of ~1.76 million bp in ~34 contigs (range, 26 to 45 contigs), with an average N_{50} value of $201,592 \pm 41,543$ bp. Gene annotations identified an average of 1,730 (range, 1,674 to 1,788) open reading frames. Since gaps between contigs have not as yet been filled, the genes at the termini of the contigs may be fragmented, and small fragments may be absent. Nonetheless, we find that genomic features related to pattern type and virulence exist in all six genomes, including the unique GAS pattern D multiple-gene activator (Mga) arrangement and the type 3 arrangement of the fibronectin-collagen-T-antigen (FCT) locus (5). All six strains also encode the major virulence determinant, viz., the direct plasminogen-binding type M protein (PAM), which is found only in pattern D GAS strains. In the case of strain NS223, PAM is partially present in two contigs, with a missing fragment in the

Received 30 May 2018 Accepted 30 May 2018 Published 28 June 2018

Citation Liang Z, Stephens M, Ploplis VA, Lee SW, Castellino FJ. 2018. Draft genome sequences of six skin isolates of *Streptococcus pyogenes*. *Genome Announc* 6:e00592-18. <https://doi.org/10.1128/genomeA.00592-18>.

Copyright © 2018 Liang 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 Francis J. Castellino, fcastell@nd.edu.

center of the gene. Also of importance, all six isolates contain the streptokinase subtype 2b gene that activates plasminogen bound to PAM-type M proteins (6). Thus, all genomic data obtained to date are consistent with the pattern D skin tropicity of these strains (7–9), similar to the well-characterized genomes of AP53 (9) and Alab49 (10), GAS skin isolates. Identifying the specific virulomes and resistance genes of GAS serotypes by whole-genome sequencing is of great importance in vaccine development and public health approaches to outbreak control.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession numbers [QFXO0000000](https://doi.org/10.1016/j.bbr.2014.01.109) (NS88.3), [QFXP0000000](https://doi.org/10.1016/j.bbr.2014.01.109) (SS1574), [QFXQ0000000](https://doi.org/10.1016/j.bbr.2014.01.109) (SS1572), [QFXR0000000](https://doi.org/10.1016/j.bbr.2014.01.109) (SS1448), [QFXS0000000](https://doi.org/10.1016/j.bbr.2014.01.109) (NS455), and [QFXT0000000](https://doi.org/10.1016/j.bbr.2014.01.109) (NS223).

ACKNOWLEDGMENTS

These studies were supported by grant HL013423 from the NIH.

We acknowledge the assistance of Yunjuan Bao with analysis of the contigs.

REFERENCES

- Schwartz B, Facklam RR, Breiman RF. 1990. Changing epidemiology of group A streptococcal infection in the USA. *Lancet* 336:1167–1171. [https://doi.org/10.1016/0140-6736\(90\)92777-F](https://doi.org/10.1016/0140-6736(90)92777-F).
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23: 673–679. <https://doi.org/10.1093/bioinformatics/btm009>.
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards R, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
- Bessen DE, Sotir CM, Readdy TL, Hollingshead SK. 1996. Genetic correlates of throat and skin isolates of group A streptococci. *J Infect Dis* 173:896–900. <https://doi.org/10.1093/infdis/173.4.896>.
- Zhang Y, Mayfield J, Ploplis VA, Castellino FJ. 2014. The β -domain of cluster 2b streptokinase is a major determinant for the regulation of its plasminogen activation activity by cellular plasminogen receptors. *Biochem Biophys Res Commun* 444:595–598. <https://doi.org/10.1016/j.bbrc.2014.01.109>.
- McGregor KF, Spratt BG, Kalia A, Bennett A, Bilek N, Beall B, Bessen DE. 2004. Multilocus sequence typing of *Streptococcus pyogenes* representing most known *emm* types and distinctions among subpopulation genetic structures. *J Bacteriol* 186:4285–4294. <https://doi.org/10.1128/JB.186.13.4285-4294.2004>.
- Bessen DE, Lizano S. 2010. Tissue tropisms in group A streptococcal infections. *Future Microbiol* 5:623–638. <https://doi.org/10.2217/fmb.10.28>.
- Bao Y-J, Liang Z, Mayfield JA, Donahue DL, Carothers KE, Lee SW, Ploplis VA, Castellino FJ. 2016. Genomic characterization of a pattern D *Streptococcus pyogenes emm53* isolate reveals a genetic rationale for invasive skin tropicity. *J Bacteriol* 198:1712–1724. <https://doi.org/10.1128/JB.01019-15>.
- Bessen DE, Kumar N, Hall GS, Riley DR, Luo F, Lizano S, Ford CN, McShan WM, Nguyen SV, Dunning Hotopp JC, Tettelin H. 2011. Whole-genome association study on tissue tropism phenotypes in group A *Streptococcus*. *J Bacteriol* 193:6551–6563. <https://doi.org/10.1128/JB.05263-11>.