



Complete Genome Sequence of Undomesticated *Bacillus subtilis* Strain NCIB 3610

Taylor M. Nye,^a Jeremy W. Schroeder,^a Daniel B. Kearns,^b Lyle A. Simmons^a

Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan, USA^a; Indiana University, Department of Biology, Bloomington, Indiana, USA^b

ABSTRACT *Bacillus subtilis* is a Gram-positive bacterium that serves as an important experimental system. *B. subtilis* NCIB 3610 is an undomesticated strain that exhibits phenotypes lost from the more common domesticated laboratory strains. Here, we announce the complete genome sequence of DK1042, a genetically competent derivative of NCIB 3610.

Bacterial strains passaged in the laboratory have lost several phenotypes that are readily observable with “wild” or undomesticated strains (1–3). *Bacillus subtilis* strain NCIB 3610 (abbreviated as 3610) is a derivative of Marburg with genomic similarity to *B. subtilis* 168 (4). Undomesticated 3610 forms multicellular structures with complex architecture and cell organization (1, 2). Pellicle biofilms, swarming motility, exopolysaccharide capsule, surfactant, and the production of antimicrobials are just a few examples of the many phenotypes that have been lost from laboratory strains (4–9). In addition to the differences in phenotypes, 3610 also harbors an 84 kb plasmid (pBS32) that has been cured from domesticated strains. Plasmid pBS32 encodes 102 genes including *comI*, the product of which is an inhibitor that prevents 3610 from developing genetic competence (10). The recent identification of ComI as a genetic competence inhibitor allowed for the generation of the naturally competent 3610 *comI*^{Q12L} (DK1042) strain, facilitating genetic study of the complex phenotypes associated with undomesticated strains (10). To further expedite genetic studies, we report the completed reference genome for DK1042, where the *comI*^{Q12L} point mutation is the only known mutation in an otherwise 3610 genetic background. Our results provide a publicly available complete reference genome for a competent 3610 derived strain, helping to forward the use of 3610 as a genetic platform for studying phenotypes and behaviors not present in domesticated laboratory strains.

Genomic DNA from DK1042 (NCIB 3610 *comI*^{Q12L}) (10) was purified via phenol-chloroform extraction (11). Sequencing libraries were prepared with 5-kb mean insert size for sequencing on the Pacific Biosciences (PacBio) RS II sequencer by the University of Michigan Sequencing Core. Two single-molecule real-time cells were used to sequence the libraries. The average subread length was 3.5 kb. *De novo* genome assembly was performed using RS_HGAP_Assembly.3 version 2.3.0 (12), resulting in two contigs representing the chromosome and plasmid (pBS32). A break was introduced at the chromosomal origin *in silico* and the genome was circularized using the minimus2 script from AMOS (13). The original PacBio sequencing reads were remapped to the circularized reference genome via RS_Resequencing.1, resulting in a consensus accuracy of 99.9997% and chromosomal coverage of 625. The previously sequenced pBS32 plasmid sequence [accession no. KF365913 (10)] was then added as a second contig to create the final NCIB 3610 *comI*^{Q12L} reference genome. The final reference genome consists of a 4,215,607 bp chromosome and 84,215 bp plasmid. The 2 contigs have a combined length of 4,299,831 bp. The previous reference genome for NCIB 3610

Received 24 March 2017 Accepted 27 March 2017 Published 18 May 2017

Citation Nye TM, Schroeder JW, Kearns DB, Simmons LA. 2017. Complete genome sequence of undomesticated *Bacillus subtilis* strain NCIB 3610. Genome Announc 5: e00364-17. <https://doi.org/10.1128/genomeA.00364-17>.

Copyright © 2017 Nye 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 Lyle A. Simmons, lasimm@umich.edu.

totaled 4,292,969 bp (chromosome and plasmid) and consisted of 84 contigs (14). Thus, the reference genome presented here has reduced the number of contigs from 84 to 2, with a 6.8 kb longer sequenced genome size.

Accession number(s). Gene annotation was performed using the Prokaryotic Genomes Annotation Pipeline (PGAP) through NCBI. The complete genome sequence of *B. subtilis* subsp. *subtilis* strain NCIB 3610 *comI*^{Q12L} was deposited in the DDBJ/EMBL/GenBank database with accession no. [CP020102](#) for the chromosome and [CP020103](#) for pBS32 (10).

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant R01 GM107312 to L.A.S., NIH Genetics training grant T32GM007544 to J.W.S., and National Science Foundation predoctoral fellowship to T.M.N. (DGE 1256260). D.B.K. was supported by National Institutes of Health R01 grant GM093030.

Author contributions: conceptualization, T.M.N., J.W.S., L.A.S., and D.B.K.; methodology, T.M.N., J.W.S., D.B.K., and L.A.S.; formal analysis, T.M.N. and J.W.S.; investigation, T.M.N. and J.W.S.; original draft, T.M.N. and L.A.S.; review and editing, T.M.N., J.W.S., D.B.K., and L.A.S.; supervision, D.B.K. and L.A.S.; funding acquisition, T.M.N., J.W.S., D.B.K., and L.A.S.

Conflict of interest statement: L.A.S. owns common stock in Illumina, Inc. and Pacific Biosciences.

REFERENCES

1. Branda SS, González-Pastor JE, Ben-Yehuda S, Losick R, Kolter R. 2001. Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 98:11621–11626. <https://doi.org/10.1073/pnas.191384198>.
2. Branda SS, Chu F, Kearns DB, Losick R, Kolter R. 2006. A major protein component of the *Bacillus subtilis* biofilm matrix. *Mol Microbiol* 59:1229–1238. <https://doi.org/10.1111/j.1365-2958.2005.05020.x>.
3. Kearns DB, Chu F, Branda SS, Kolter R, Losick R. 2005. A master regulator for biofilm formation by *Bacillus subtilis*. *Mol Microbiol* 55:739–749. <https://doi.org/10.1111/j.1365-2958.2004.04440.x>.
4. Earl AM, Losick R, Kolter R. 2007. *Bacillus subtilis* genome diversity. *J Bacteriol* 189:1163–1170. <https://doi.org/10.1128/JB.01343-06>.
5. McLoon AL, Guttenplan SB, Kearns DB, Kolter R, Losick R. 2011. Tracing the domestication of a biofilm-forming bacterium. *J Bacteriol* 193:2027–2034. <https://doi.org/10.1128/JB.01542-10>.
6. Stanley NR, Lazazzera BA. 2005. Defining the genetic differences between wild and domestic strains of *Bacillus subtilis* that affect poly-gamma-dl-glutamic acid production and biofilm formation. *Mol Microbiol* 57:1143–1158. <https://doi.org/10.1111/j.1365-2958.2005.04746.x>.
7. Butcher RA, Schroeder FC, Fischbach MA, Straight PD, Kolter R, Walsh CT, Clardy J. 2007. The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 104:1506–1509. <https://doi.org/10.1073/pnas.0610503104>.
8. Kinsinger RF, Shirk MC, Fall R. 2003. Rapid surface motility in *Bacillus subtilis* is dependent on extracellular surfactin and potassium ion. *J Bacteriol* 185:5627–5631. <https://doi.org/10.1128/JB.185.18.5627-5631.2003>.
9. Kearns DB, Chu F, Rudner R, Losick R. 2004. Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. *Mol Microbiol* 52:357–369. <https://doi.org/10.1111/j.1365-2958.2004.03996.x>.
10. Konkol MA, Blair KM, Kearns DB. 2013. Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J Bacteriol* 195:4085–4093. <https://doi.org/10.1128/JB.00696-13>.
11. Hardwood CR, Cutting SM. 1990. *Molecular biological methods for Bacillus*. John Wiley & Sons, Chichester.
12. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
13. Treangen TJ, Sommer DD, Angly FE, Koren S, Pop M. 2011. Next generation sequence assembly with AMOS. *Curr Protoc Bioinformatics Chapter 11:Unit 11.8*. <https://doi.org/10.1002/0471250953.bi1108s33>.
14. Srivatsan A, Han Y, Peng J, Tehranchi AK, Gibbs R, Wang JD, Chen R. 2008. High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. *PLoS Genet* 4:e1000139. <https://doi.org/10.1371/journal.pgen.1000139>.