


Draft Genome Sequences of Two South African *Bacillus anthracis* Strains

Kgaugelo E. Lekota,^{a,b,c,d} Joseph Mafofo,^a Evelyn Madoroba,^{c,d} Jasper Rees,^a  Henriette van Heerden,^b Farai C. Muchadeyi^a

Biotechnology Platform, Agricultural Research Council, Onderstepoort, South Africa^a; Department of Veterinary Tropical Diseases, University of Pretoria, Onderstepoort, South Africa^b; Agricultural Research Council, Onderstepoort Veterinary Institute, Onderstepoort, South Africa^c; University of South Africa College of Agriculture and Environmental Sciences, Florida Campus, Florida, South Africa^d

***Bacillus anthracis* is a Gram-positive bacterium that causes anthrax, mainly in herbivores through exotoxins and capsule produced on plasmids, pXO1 and pXO2. This paper compares the whole-genome sequences of two *B. anthracis* strains from an endemic region and a sporadic outbreak in South Africa. Sequencing was done using next-generation sequencing technologies.**

Received 29 September 2015 Accepted 1 October 2015 Published 19 November 2015

Citation Lekota KE, Mafofo J, Madoroba E, Rees J, van Heerden H, Muchadeyi FC. 2015. Draft genome sequences of two South African *Bacillus anthracis* strains. *Genome Announc* 3(6):e01313-15. doi:10.1128/genomeA.01313-15.

Copyright © 2015 Lekota et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Farai C. Muchadeyi, muchadeyif@arc.agric.za.

Anthrax is a zoonotic bacterial disease caused by the spore-forming soil-borne bacterium *Bacillus anthracis*. Fully virulent wild-type strains of *B. anthracis* carry two large plasmids, pXO1 and pXO2 (182 and 96 kb, respectively), which harbor the virulence factors (1). Anthrax occurs endemically in the Northern Cape region (NCP) and Kruger National Park (KNP), but sporadic outbreaks occur in livestock and wildlife throughout the country.

In this study, two strains were sequenced, 3631_1C, isolated from a kudu (*Tragelaphus strepsiceros*) in the NCP during the 2009 outbreak in Klipfontein, and 20SD, isolated from a sporadic outbreak in sheep (*Ovis aries*) in the Standerton area, Mpumalanga Province, South Africa. The genotyping method, multilocus variable-number tandem-repeat analysis (MLVA) (2, 3), clustered both *B. anthracis* strains in the A clade.

The sequencing libraries were generated using the Nextera XT DNA sample preparation kit (Illumina) and sequenced on the Illumina HiScan SQ (Illumina) instrument. A total of approximately 2.0 and 3.0 million 100-bp paired-end reads were generated for 20SD and 3631_1C, respectively, and analyzed using CLC Genomics Workbench version 7. The genome assemblies of the strains were aligned to *B. anthracis* Ames Ancestor (GenBank accession numbers AE017334, AE017336, and AE017335) (4) using the Burrows-Wheeler Aligner (BWA) (5). Approximately 99% of the reads generated for both the 20SD and 3631_1C strains mapped to the reference chromosome and the plasmids. The sequence data for 20SD had 30-fold coverage of the chromosome, 98-fold coverage of pXO1, and 48-fold coverage of the pXO2. The sequence data for the 3631_1C strain gave 53-fold coverage of the chromosome, 127-fold coverage of plasmid pXO1, and lacked pXO2.

The sequence data were also used for *de novo* assembly of the two strains using CLC Genomics Workbench version 7. The assembled contigs were aligned using BLASTn (6) using *B. anthracis* Ames ancestor as a reference. The assembled contigs were ordered using Mauve (7) and ABACAS (8) to check for consistency in

overlapping contigs. The assembled genomes comprise 52 and 45 contigs for strains 20SD and 3631_1C, respectively, with 50 and 44 contigs mapping to the chromosomes for strains 20SD and 3631_1C, respectively, while the plasmids, pXO1 and pXO2, assembled into 1 contig each.

The draft genomes were annotated using the RAST (9). Strain 3631_1C gave a total of 5,680 coding sequences (CDSs) on the chromosome, 204 CDSs on plasmid pXO1, and a total of 64 tRNAs, whereas pXO1 and pXO2 of strain 20SD produced 201 and 117 CDSs, respectively, 5,645 CDSs on the chromosome and a total of 49 tRNAs. The draft genome sequences for *B. anthracis* 20SD and 3631_1C were approximately 5.44 and 5.35 Mb, respectively. Both strains yielded an average G+C content of 35%.

Previous studies have shown that *B. anthracis* strains, especially those isolated from the soil and/or after long-term culturing, usually lack one or both plasmids (10, 11). This was also observed in this study in strain 3631_1C, whereas strain 20SD was a fully virulent *B. anthracis* strain.

Nucleotide sequence accession numbers. The draft genome sequences of the *B. anthracis* 20SD and 3631_1C strains are available in GenBank under the accession numbers [LGCC00000000](https://ncbi.nlm.nih.gov/nucl/LGCC00000000) and [LGCD00000000](https://ncbi.nlm.nih.gov/nucl/LGCD00000000), respectively.

ACKNOWLEDGMENTS

This project was supported by the Agricultural Research Council (ARC)-Biotechnology Platform and National Research Foundation.

We thank the ARC-Biotechnology Platform sequencing facility for sequencing technical support.

REFERENCES

- Mock M, Fouet A. 2001. Anthrax. *Annu Rev Microbiol* 55:647–671. <http://dx.doi.org/10.1146/annurev.micro.55.1.647>.
- Van Ert MN, Easterday WR, Simonson TS, U'Ren JM, Pearson T, Kenefic LJ, Busch JD, Huynh LY, Dukerich M, Trim CB, Beaudry J, Welty-Bernard A, Read T, Fraser CM, Ravel J, Keim P. 2007. Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* Ames strain. *J Clin Microbiol* 45:47–53. <http://dx.doi.org/10.1128/JCM.01233-06>.

3. Beyer W, Bellan S, Eberle G, Ganz HH, Getz WM, Haumacher R, Hilss KA, Kilian W, Lazak J, Turner WC, Turnbull PCB. 2012. Distribution and molecular evolution of *Bacillus anthracis* genotypes in Namibia. *PLoS Negl Trop Dis* 6:e1534.
4. Ravel J, Jiang L, Stanley ST, Wilson MR, Decker RS, Read TD, Worsham P, Keim PS, Salzberg SL, Fraser-Liggett CM, Rasko DA. 2009. The complete genome sequence of *Bacillus anthracis* Ames "Ancestor." *J Bacteriol* 191:445–446. <http://dx.doi.org/10.1128/JB.01347-08>.
5. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <http://dx.doi.org/10.1093/bioinformatics/btp324>.
6. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic Local Alignment Search Tool. *J Mol Biol* 215:403–410. [http://dx.doi.org/10.1016/S0022-2836\(05\)80360-2](http://dx.doi.org/10.1016/S0022-2836(05)80360-2).
7. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <http://dx.doi.org/10.1101/gr.2289704>.
8. Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics* 25:1968–1969. <http://dx.doi.org/10.1093/bioinformatics/btp347>.
9. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
10. Marston CK, Hoffmaster AR, Wilson KE, Bragg SL, Plikaytis B, Brachman P, Johnson S, Kaufmann AF, Popovic T. 2005. Effects of long-term storage on plasmid stability in *Bacillus anthracis*. *Appl Environ Microbiol* 71:7778–7780. <http://dx.doi.org/10.1128/AEM.71.12.7778-7780.2005>.
11. Birdsell DN, Antwerpen M, Keim P, Hanczaruk M, Foster JT, Sahl JW, Wagner DM, Grass G. 2013. Draft genome sequences of two Bulgarian *Bacillus anthracis* strains. *Genome Announc* 1(2):e00152-13. <http://dx.doi.org/10.1128/genomeA.00152-13>.