genomeA
Journals.ASM.org

Draft Genome Sequences of Two Bulgarian *Bacillus anthracis* Strains

Dawn N. Birdsell,[a] Markus Antwerpen,[c] Paul Keim,[a,b] Matthias Hanczaruk,[c] Jeffrey T. Foster,[a] Jason W. Sahl,[b] David M. Wagner,[a] Gregor Grass[c]

Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, Arizona, USA[a]; Pathogen Genomics Division, Translational Genomics Research Institute, Flagstaff, Arizona, USA[b]; Bundeswehr Institute of Microbiology, Munich, Germany[c]

Bacillus anthracis strains previously isolated from Bulgaria form a unique subcluster within the A1.a cluster that is typical for isolates from southeastern Europe. Here, we report the draft genome sequences of two Bulgarian *B. anthracis* strains belonging to the A branch (A.Br.)008/009 canonical single nucleotide polymorphism (SNP) group of the major A branch.

Address correspondence to Gregor Grass, gregorgrass@bundeswehr.org.

Anthrax is a zoonotic disease that is endemic in many countries worldwide. The Gram-positive endospore-forming bacterium *Bacillus anthracis* produces several virulence factors required for establishing infections in animal and human hosts. Major virulence factors are encoded on two large plasmids, pX01 and pX02. Despite the importance of these plasmids for host infections, many strains of *B. anthracis*, especially those isolated from soil and/or after long-term culturing, lack one or both plasmids (1). Previously, we genetically analyzed 40 *B. anthracis* strains collected in Bulgaria, isolated mostly between 1960 and 1980, and found that over half of them lacked one or both plasmids (2). The three major phylogenetic lineages (branches A, B, and C) of *B. anthracis* are divided into 12 clonal groups by the analysis of 13 canonical single nucleotide polymorphisms (canSNPs) (3). We classified each strain as belonging to either the A branch (A.Br.)008/009 or A.Br.WNA genetic group (2) and placed all 40 strains onto the existing global phylogenetic tree (3, 4). As part of our continuing efforts aimed at determining the phylogeography of *B. anthracis* isolates in Europe, we sequenced the genomes of two Bulgarian *B. anthracis* strains representing the A.Br.008/009 genetic group. Whole genome sequences of these strains permit in-depth comparisons of Bulgarian *B. anthracis* A.Br.008/009 strains with other A.Br.008/009 strains from western and central Europe and areas south and east of the Black Sea (2).

Whole-genome shotgun (WGS) sequencing of the *B. anthracis* strains was performed using the Illumina GAIIx sequencing platform (Illumina, Inc.). For the WGS libraries, 4.6 to 5.79 million 100-bp reads were generated. Reads were assembled with AMOScmp (5), using the *B. anthracis* Ames ancestor strain (6) as the reference. Following assembly, contigs were processed with the PAGIT pipeline (7). A *de novo* assembly was also performed with Velvet (8), in conjunction with the VelvetOptimiser (http://bioinformatics.net.au/software.velvetoptimiser.shtml). The AMOS and Velvet assemblies were then aligned with Mugsy (9). Regions unique to the *de novo* assembly were then aligned back against the comparative assembly with BLASTn (10). Contigs that failed to align were concatenated with the comparative assembly into the

final assembly. The assembled contigs were submitted to the RAST annotation server for subsystem classification and functional annotation (11).

The total lengths of the draft genome shotgun sequences of *B. anthracis* strains 3154 and 3166 were 5.0 and 5.5 Mbp, respectively, and their mean G+C content was 35.35% ± 0.1%. The chromosomal sequences of these strains contained 3 to 5 contigs with 189- to 235-fold average coverage. The plasmids pX01 and pX02 in strain 3166 were assembled on 2 contigs with 434-fold coverage. The *B. anthracis* 3154 genome contains a ~200-kb deletion compared to the Ames ancestor (accession no. YP_019942 to YP_020195), comprising genes for a beta-lactam antibiotic acylase (accession no. YP_019964) and a sulfamethoxypyridazine (SMPR) multidrug efflux pump (accession no. YP_019986).

Nucleotide sequence accession numbers. The draft genome sequences of *B. anthracis* strains 3154 and 3166 have been included in the GenBank WGS database under accession no. ANFF00000000 and ANFG00000000, respectively. The versions described in this paper are the first versions, accession no. ANFF01000000 and ANFG01000000, respectively.

ACKNOWLEDGMENTS

REFERENCES

1. **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.
2. **Antwerpen M, Ilin D, Georgieva E, Meyer H, Savov E, Frangoulidis D.** 2011. MLVA and SNP analysis identified a unique genetic cluster in Bulgarian *Bacillus anthracis* strains. Eur. J. Clin. Microbiol. Infect. Dis. 30:923–930.
3. **Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J, Zanecki SR, Pearson T, Simonson TS, U'Ren JM, Kachur SM, Leadem-Dougherty RR, Rhoton SD, Zinser G, Farlow J, Coker PR,**

Smith KL, Wang B, Kenefic LJ, Fraser-Liggett CM, Wagner DM, Keim P. 2007. Global genetic population structure of *Bacillus anthracis*. PLoS One **2**:e461.

4. Kenefic LJ, Pearson T, Okinaka RT, Schupp JM, Wagner DM, Hoffmaster AR, Trim CB, Chung WK, Beaudry JA, Jiang L, Gaier P, Foster JT, Mead JI, Keim P, Ravel J, Keim P. 2009. Pre-Columbian origins for North American anthrax. PLoS One **4**:e4813.

5. Pop M, Phillippy A, Delcher AL, Salzberg SL. 2004. Comparative genome assembly. Brief. Bioinform. **5**:237–248.

6. 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.

7. Swain MT, Tsai IJ, Assefa SA, Newbold C, Berriman M, Otto TD. 2012. A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs. Nat. Protoc. **7**:1260–1284.

8. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.

9. Angiuoli SV, Salzberg SL. 2011. Mugsy: fast multiple alignment of closely related whole genomes. Bioinformatics **27**:334–42.

10. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

11. 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, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75.