Complete Genome of *Bacillus subtilis* Myophage Grass

Stanton Y. Miller,[a] Jennifer M. Colquhoun,[b] Abbey L. Perl,[b] Karthik R. Chamakura,[a] Gabriel F. Kuty Everett[a]

Center for Phage Technology, Texas A&M University, College Station, Texas, USA[a]; Department of Biological Sciences, Lehigh University, Bethlehem, Pennsylvania, USA[b]

***Bacillus subtilis* is a ubiquitous Gram-positive model organism. Here, we describe the complete genome of *B. subtilis* myophage Grass. Aside from genes encoding core proteins pertinent to the life cycle of the phage, Grass has several interesting features, including an FtsK/SpoIIIE protein.**

Address correspondence to Gabriel F. Kuty Everett, gabbyeverett@tamu.edu.

*B*acillus subtilis is a soil-dwelling, endospore-forming, Gram-positive bacterium (1). *B. subtilis* is well studied and has been used in many commercial applications, such as the protection of crop growth and biofilm analysis (2, 3). It is also a well-understood model for determining the properties of spores. The study of *B. subtilis* phages has contributed greatly to our understanding of the genetics and physiology of this model organism. Here, we present the complete annotated genome of the *B. subtilis* phage Grass.

Grass was isolated from a soil sample collected in Bethlehem, PA. Phage DNA was sequenced using 454 pyrosequencing at the Emory GRA Genome Center (Emory University, Atlanta, GA). The trimmed FLX Titanium reads were assembled to a single contig at 128.6-fold coverage using the Newbler assembler version 2.5.3 (454 Life Sciences) with the default settings. The contigs were confirmed to be complete by PCR. Genes were predicted using GeneMarkS (4) and corrected using software tools available on the Center for Phage Technology (CPT) portal (https://cpt.tamu.edu/cpt-software/portal/).

Grass contains a 152,440-bp unique genome with a coding density of 89.1% and a G+C content of 42.3%. In the unit genome, 242 coding sequences were identified, of which 92 were novel hypothetical genes and 95 were conserved hypothetical genes. The genome encodes 3 tRNAs and contains 38 rho-independent terminators. Electron microscopy imaging shows that Grass is a myophage. The TerL showed homology to the TerLs of phages with long terminal repeats. The raw sequencing data were processed using the Pause (https://cpt.tamu.edu/cpt-software/releases/pause/) method to show that the terminal repeat is 4,208 bp in length.

Genes encoding proteins pertaining to replication and recombination, DNA biosynthesis, packaging, morphogenesis, and lysis were identified. Genes for replication and modification proteins include UvsW helicase, DnaB-like helicase, primase, polymerase, RecA, and several DNA-binding proteins and nucleases. Genes encoding DNA biosynthesis proteins include dUTPase, dihydrofolate reductase, thymidylate synthase, nucleotidyltransferase, ribonucleotide diphosphate reductase subunits alpha and beta, and a ribonucleotide reductase stimulatory protein. The large terminase and portal proteins were also annotated. Genes encoding phage morphogenesis proteins were identified (major and minor capsid proteins, tail sheath proteins, tail lysin, base-plate assembly proteins [P2 W {InterPro database no. IPR007048} and P2 J {pfam04865}], and a tailspike protein with a pectin lyase domain). A holin (class II with two transmembrane domains in an N-in C-in topology) and endolysin (*N*-acetylmuramoyl-L-alanine amidase) were also found.

Interesting features in the Grass genome are genes encoding an FtsK/SpoIIIE protein, an RtcB family protein, and a ParM protein. SpoIIIE is an ATPase common in sporulating bacteria that pumps DNA into the forespore (5). RtcB is an RNA ligase involved in tRNA splicing and repair (6). ParM is an actin homolog involved in plasmid segregation that is also found in the lytic *Bacillus* phage SPO1. How these proteins are involved in phage replication and assembly is unknown.

Nucleotide sequence accession number. The genome sequence of phage Grass was contributed as accession no. KF669652 to GenBank.

ACKNOWLEDGMENTS

REFERENCES

1. **Tam NK, Uyen NQ, Hong HA, Duc LH, Hoa TT, Serra CR, Henriques AO, Cutting SM.** 2006. The intestinal life cycle of *Bacillus subtilis* and close relatives. J. Bacteriol. **188:**2692–2700.
2. **Islam MR, Jeong YT, Lee YS, Song CH.** 2012. Isolation and identification of antifungal compounds from *Bacillus subtilis* C9 inhibiting the growth of plant pathogenic fungi. Mycobiology **40:**59–66.
3. **Lemon KP, Earl AM, Vlamakis HC, Aguilar C, Kolter R.** 2008. Biofilm development with an emphasis on *Bacillus subtilis*. Curr. Top. Microbiol. Immunol. **322:**1–16.
4. **Besemer J, Lomsadze A, Borodovsky M.** 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. **29:**2607–2618.
5. **Kaimer C, Graumann PL.** 2011. Players between the worlds: multifunctional DNA translocases. Curr. Opin. Microbiol. **14:**719–725.
6. **Tanaka N, Meineke B, Shuman S.** 2011. RtcB, a novel RNA ligase, can catalyze tRNA splicing and HAC1 mRNA splicing *in vivo*. J. Biol. Chem. **286:**30253–30257.