Complete Genome Sequence of *Paenibacillus mucilaginosus* 3016, a Bacterium Functional as Microbial Fertilizer

Mingchao Ma,[a] Zhenya Wang,[b] Li Li,[a] Xin Jiang,[a] Dawei Guan,[a] Fengming Cao,[a] Huijun Chen,[a] Xuan Wang,[b] Delong Shen,[a] Binghai Du,[b] and Jun Li[a]

Institute of Agricultural Resources and Regional Planning Chinese Academy of Agricultural Sciences, Beijing, China,[a] and College of Life Sciences, Shandong Agricultural University, Taian, Shandong, China[b]

Paenibacillus mucilaginosus is a ubiquitous functional bacterium in microbial fertilizer. Here we report the complete sequence of *P. mucilaginosus* 3016. Multiple sets of functional genes have been found in the genome. To the best of our knowledge, this is the first announcement about the complete genome sequence of a *P. mucilaginosus* strain.

*P*aenibacillus mucilaginosus (formerly *Bacillus mucilaginosus*) (6), the type species of *Paenibacillus*, is consequently widely used as a microbial fertilizer in agricultural applications due to its growth-promoting properties (7, 12). *P. mucilaginosus* 3016 was isolated from rhizosphere soil and has been widely used in microbial fertilizer by most manufacturers in China (20).

The *P. mucilaginosus* genome was determined by Roche 454 pyrosequencing (15), and Illumina sequencing by synthesis. The Illumina adapters were ligated onto fragmented *P. mucilaginosus* genomic DNA, and then gel electrophoresis was used to select DNA fragments about 3 kbp in size. Libraries were PCR amplified using Phusion polymerase. Sequencing libraries were denatured with sodium hydroxide and diluted in hybridization buffer for loading onto a single lane of an Illumina genome analyzer (GA) flow cell. Cluster formation, primer hybridization, and sequencing were performed using proprietary reagents according to manufacturer-recommended protocols. The paired-end DNA library (8-kbp span) for the 454 platform was prepared according to the manual from Roche. All the low-quality sequences were trimmed before assembly. The mate-paired reads (718 Mbp; 82.53× coverage) were generated by a Solexa sequencer and assembled by the SOAPdenovo method (11). Then, the 454 reads (190 Mbp; 21.84× coverage) and the split fragments of contigs generated by SOAPdenovo were used for a hybrid assembly with the Newbler sequence assembler (version 2.6). To finish the genome, conventional Sanger sequencing technologies were used to fill the gaps. Coding sequences were predicted by Glimmer3 (4). Functional assignment and classification were determined by performing sequence similarity searches with BLAST (1) (E-value cutoff, 1e−5) against the eggNOG database (16), the KEGG reference database (8), and the nonredundant GenBank CDS database. The tRNA and the rRNA genes were predicted by tRNASCAN-SE (13) and RNAmmer (9), respectively. A set of in-house Perl scripts and EMBOSS (17) were used for sequence manipulation.

The complete genome sequence of strain 3016 is a circular 8,739,048-bp chromosome with mean GC contents of 58.3%. There are 7,285 coding genes, 42 rRNA operons, and 170 tRNAs in the chromosome. Many essential genes were detected in the chromosome, such as the genes involving the metabolism of nitrogen, phosphorus, and potassium, as well as nitrogen-fixing NifU domain-containing protein, potassium channel protein, and potassium-transporting ATPase subunit A, indicating its importance to the strain's life.

In comparison to *Paenibacillus polymyxa* SC2, a strain of plant growth-promoting rhizobacteria in the same genus, there are 1,641 shared genes and 5,416 unique genes in the chromosome of strain 3016. The unique genes, such as the *ywlF* gene (NCBI M3016-39), the *ppdK* gene (M3016-1649), the *ytiB* gene (M3016-3117), the *yfkO* gene (M3016-4045), the *yvdA* gene (M3016-4842), and the *yurN* gene (M3016-5050), are most involved in carbohydrate transport and metabolism (10.73% of the unique genes), transcription (9.31% of the unique genes), signal transduction mechanisms (6.96% of the unique genes), and amino acid transport and metabolism (5.89% of the unique genes), based on their similarities to previously published genes (2, 3, 5, 10, 14, 18, 19, 21).

Nucleotide sequence accession number. The *Paenibacillus mucilaginosus* 3016 chromosome sequence has been deposited in GenBank under accession number CP003235.

REFERENCES

1. **Altschul SF, et al.** 1997. Gapped blast and psi-blast: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.
2. **Chen XH, et al.** 2007. Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium Bacillus amyloliquefaciens FZB42. Nat. Biotechnol. **25**:1007–1014.
3. **Chivian D, et al.** 2008. Environmental genomics reveals a single-species ecosystem deep within Earth. Science **332**:275–278.
4. **Delcher AL, Harmon D, Kasif S, White O, Salzberg SL.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27**:4636–4641.

5. **Gibbons HS, et al.** 2011. Genomic signatures of strain selection and enhancement in Bacillus atrophaeus var. globigii, a historical biowarfare stimulant. PLoS One **6**:e17836.
6. **Hu XF, et al.** 2010. Transfer of *Bacillus mucilaginosus* and *Bacillus edaphicus* to the genus *Paenibacillus* as *Paenibacillus mucilaginosus* comb. nov. and *Paenibacillus edaphicus* comb. nov. Int. J. Syst. Evol. Microbiol. **60**:8–14.
7. **Hu X-F, Gao Y-Y, Fang Q-L, Wu J-G, Chen J-S.** 2008. Effect of ion implantation on *B. mucilaginosus* KNP414 and screening for mutants with higher release of phosphate and potassium. J. Nucl. Agric. Sci. **22**(4):420–425.
8. **Kanehisa M, Goto S.** 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. **28**:27.
9. **Lagesen K, et al.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35**:3100.
10. **Lapidus A, Galleron N, Sorokin A, Ehrlich SD.** 1997. Sequencing and functional annotation of the Bacillus subtilis genes in the 200 kb rrnB-dnaB region. Microbiology **143**:3431–3441.
11. **Li R, et al.** 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. **20**:265.
12. **Liu WX, et al.** 2006. Decomposition of silicate minerals by *Bacillus mucilaginosus* in liquid culture. Environ. Geochem. Health **28**:133–140.
13. **Lowe TM, Eddy SR.** 1997. tRNAscan-se: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25**:955–964.
14. **Lu YK, et al.** 2010. Metabolic flexibility revealed in the genome of the cyst-forming alpha-1 proteobacterium Rhodospirillum centenum. BMC Genomics **11**:325–337.
15. **Margulies M, et al.** 2005. Genome sequencing in open microfabricated high density picoliter reactors. Nature **437**:376–380.
16. **Muller J, et al.** 2010. eggNOG v2.0: extending the evolutionary genealogy of genes with enhanced non-supervised orthologous groups, species and functional annotations. Nucleic Acids Res. **38**(database issue):D190–D195.
17. **Rice P, et al.** 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. **16**(6):276–277.
18. **Ueda K, et al.** 2004. Genome sequence of Symbiobacterium thermophilum, an uncultivable bacterium that depends on microbial commensalism. Nucleic Acids Res. **32**:4937–4944.
19. **Veith B, et al.** 2004. The complete genome sequence of Bacillus licheniformis DSM13, an organism with great industrial potential. J. Mol. Microbiol. Biotechnol. **7**:204–211.
20. **Wang X, et al.** 2011. Rapid identification for Paenibacillus mucilaginosus by PCR. Wei Sheng Wu Xue Bao **51**:1485–1493. (In Chinese.)
21. **Zhang G, et al.** 2011. Complete genome sequence of *Bacillus amyloliquefaciens* TA208, a strain for industrial production of guanosine and ribavirin. J. Bacteriol. **193**:3142–3143.