

GENOME ANNOUNCEMENTS

Draft Genome Sequence of *Paenibacillus elgii* B69, a Strain with Broad Antimicrobial Activity[∇]

Rui Ding,^{1†} Yudong Li,^{2†} Chaodong Qian,¹ and Xuechang Wu^{1*}

Institute of Microbiology, College of Life Sciences, Zhejiang University, Hangzhou, China,¹ and Department of Biological Engineering, Zhejiang Gongshang University, Hangzhou, China²

Received 25 March 2011/Accepted 10 June 2011

Here, we report the draft genome sequence of *Paenibacillus elgii* B69, which was isolated from soil and has broad-spectrum antimicrobial activity. As far as we know, the *P. elgii* genome is the largest of the *Paenibacillus* genus for which genome sequences are available. Multiple sets of genes related to antibiotic biosynthetic pathways have been found in the genome.

Recently, interest in the *Paenibacillus* genus has been rapidly growing, for many *Paenibacillus* species have been shown to produce various enzymes and antimicrobial substances, such as xylanase and polymyxins (3, 7). There are more than 110 species that have been discovered in this genus (8). The *Paenibacillus elgii* strain B69 was isolated from soil samples collected from Tianmu Mountain in Hangzhou, China, and displays significant activity against all tested microorganisms (9).

Whole-genome shotgun (WGS) sequencing of *Paenibacillus elgii* B69 was performed with Solexa paired-end sequencing technology (1). A total of 2,564.46 million base pairs of reads were generated to reach a depth of ~322-fold genome coverage and were assembled into 278 contigs using SOAPdenovo version 1.04 (4). Furthermore, the paired-end information was used to join contigs into 184 scaffolds (>1 kb in size). The open reading frames were annotated by using Glimmer version 3.0 (2) and by searching against the COG, KEGG, UniProt, and nr databases.

The unclosed draft genome sequence of *Paenibacillus elgii* B69 showed a genome size of 7,981,270 bp (scaffold length) with a mean GC content of 52.57%. There are 7,907 coding sequences (CDSs), and 51 tRNAs were identified by tRNAscan-SE (5). In addition, 330 insertion sequence elements, 330 transposases, and 227 tandem repeats were found.

Gene family analysis of *Paenibacillus elgii* B69 and *Paenibacillus* sp. strain JDR-2 showed that 123 gene families were constructed (more than 10 genes for each of them), and the largest gene family included 414 genes (186 from JDR-2). From the 7,907 identified genes, 5,704 CDSs were classified into 22 functional COG groups, and the unannotated genes may be assigned upon closure of the genome. One of the notable features of B69 is its potential to produce secondary metabolites, such as pelgipeptins (9) and paenibactin, which is a novel catecholic siderophore iden-

tified recently in our lab (unpublished data). No polymyxin synthetase gene was found, but several unknown giant gene clusters were found in the B69 genome, including the genes for one polyketide synthetase (PKS), two nonribosomal peptide synthetases (NRPS), three hybrid NRPS-PKS synthetases, and a lantibiotic-synthetic gene cluster. Surprisingly, the genome size of *P. elgii* is larger than that of other *Paenibacillus* species with genome sequences available (3, 6), such as *P. polymyxa* (~5.7 Mb), suggesting that the size difference could have arisen through acquisition of more genes by gene duplication or horizontal transfer during evolution. Comparative genome analysis of *Paenibacillus* strains will help us identify diverse antibiotic biosynthetic genes and elucidate the evolution of biosynthetic pathways among the members of this genus.

Nucleotide sequence accession number. The sequence from this whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under accession number AFHW00000000. The version described in this paper is the first version, AFHW01000000.

This work was supported in part by grant Y201016960 from the Education Bureau of Zhejiang Province.

We thank BGI (the Beijing Genomics Institute) for their work in Solexa sequencing, assembly, and annotation of the genome.

REFERENCES

1. Bentley, D. R., et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**:53–59.
2. Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**:673–679.
3. Kim, J. F., et al. 2010. Genome sequence of the polymyxin-producing plant-probiotic rhizobacterium *Paenibacillus polymyxa* E681. *J. Bacteriol.* **192**:6103–6104.
4. Li, R., et al. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20**:265–272.
5. Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of tRNA genes in genomic sequence. *Nucleic Acids Res.* **25**:955–964.
6. Ma, M., et al. 2011. Complete genome sequence of *Paenibacillus polymyxa* SC2, a strain of plant growth-promoting rhizobacterium with broad-spectrum antimicrobial activity. *J. Bacteriol.* **193**:311–312.
7. St. John, F. J., J. D. Rice, and J. F. Preston. 2006. *Paenibacillus* sp. strain JDR-2 and XynA1: a novel system for methylglucuronoxylan utilization. *Appl. Environ. Microbiol.* **72**:1496–1506.
8. Wu, X., et al. 2011. *Paenibacillus tianmuensis* sp. nov., isolated from soil. *Int. J. Syst. Evol. Microbiol.* **61**:1133–1137.
9. Wu, X., et al. 2010. Isolation and partial characterization of antibiotics produced by *Paenibacillus elgii* B69. *FEMS Microbiol. Lett.* **310**:32–38.

* Corresponding author. Mailing address: Institute of Microbiology, College of Life Sciences, Zhejiang University, Hangzhou 310058, People's Republic of China. Phone: 86-571-88206637. Fax: 86-571-88206627. E-mail: mblab@163.com.

† The first two authors contributed equally to this work.

∇ Published ahead of print on 24 June 2011.