

[SVIII-1]

Secrets of the Underground Lifestyle of the Plant-probiotic Bacterium *Paenibacillus polymyxa* E681 Revealed from Its Genomic Sequence

Jihyun F. Kim^{1*}, Changhoon Kim², Haeyoung Jeong⁴, Soo-Young Park¹,
Yon-Kyong Park¹, Soo Keun Choi¹, Cheol-Goo Hur², Tae Kwang Oh³,
Young Ho Moon⁴, Chang Seuk Park⁵, and Seung-Hwan Park¹

¹Laboratory of Microbial Genomics, ²National Center for Genome Information

³Microbial Genomics and Applications Center, Korea Research Institute of Bioscience and
Biotechnology (KRIBB), P.O. Box 115, Yusong, Taejeon 305-600;

⁴GenoTech Corporation, Daedeok BioCommunity, 461-6 Jeonmin-dong, Yusong, Taejeon 305-811;

⁵Gyeongsang National University, 900 Gazwa-dong, Chinju 660-701, Republic of Korea

The genus *Paenibacillus* belongs to endospore-forming low-G+C Gram-positive bacilli. Species in the genus can fix atmospheric nitrogen, produce antimicrobial compounds, and secrete diverse hydrolyzing enzymes. A strain of *Paenibacillus polymyxa* capable of suppressing plant diseases and promoting plant growth was isolated from the rhizosphere of winter barley grown in Chonnam province, Korea and designated E681. The sequence of the ca. ≥ 5 Mb genome was determined to understand the mechanisms of these beneficial effects. Genomic libraries of several different insert sizes ranging from two to ten kilobases were constructed, and both ends of each insert were sequenced. Total length of the reads after vector clipping and quality trimming amounted to 31 Mb. The G+C content of the sequences was 45.55%. The sequences were assembled, and 4,182 open reading frames were predicted. About 56.2% of them were then functionally assigned by matching the deduced proteins with entries in the COG database; about 23.3% of the genes were unique to E681. Interestingly, the E681 genome had high proportions of genes involved in transcription, signal transduction, carbohydrate metabolism, amino acid metabolism, and biosynthesis, transport and catabolism of secondary metabolites. This suggests that E681 is well adapted to changing environmental conditions, is capable of metabolizing a wide array of nutrients, and produces various secondary metabolites.

Introduction

Among the myriad bacteria thriving in the rhizosphere (the surface of the plant root and close vicinity), some are beneficial to plants. They are antagonistic to various plant pests and pathogens, induce plant defense systems against infection and infestation, and/or even promote plant growth. These so called plant-growth-promoting rhizobacteria can either fix atmospheric nitrogen, produce plant hormones or produce antibiotic compounds, and thus could be utilized for biofertilization, growth stimulation, and biological control of pests and pathogens (Bloemberg and Lugtenberg, 2001). Taxonomically diverse groups of “plant-probiotic” bacteria have been isolated and commercially developed. Some spore-forming bacteria, in particular Gram-positive

bacilli and streptomycetes, recently attracted special attention due to their advantages over non-spore formers in product formulation and stable maintenance in soil (Emmert and J. Handelsman, 1999).

The genus *Paenibacillus* (Ash *et al.*, 1993; previously *Bacillus*), comprised of more than thirty species, is facultatively anaerobic and belongs to endospore-forming low-G+C Gram-positive bacilli. Species in the genus can fix nitrogen, produce antimicrobial compounds, and secrete diverse hydrolyzing enzymes. *Paenibacillus polymyxa*, the type species of the genus, is known to produce two types of peptide antibiotics (Beatty and Jensen, 2002). One group is active against bacteria and includes the polymyxin-colistin-circulin family, polypeptins, jolipeptin, gavaserin, and saltavalin, which contain 2,4-diaminobutyric acid. The other is active against fungi and actinomycetes and includes gatavalin and fusaricidins. The species also synthesizes plant hormones auxin (Lebuhn *et al.*, 1997) and cytokinin (Timmusk *et al.*, 1999).

In view of the immense potential of *P. polymyxa* for use in plant health management and industrial applications, we determined the genome sequence of a “plant-probiotic” strain isolated in Korea by the whole-genome shotgun sequencing approach (Fraser *et al.*, 2000).

Results and Discussion

An endospore-forming strain, capable of inhibiting the growth of plant-pathogenic fungi and bacteria, and promoting plant growth, was isolated from the rhizosphere of winter barley grown in Jeollanam-do province, Korea, and was designated E681 (Ryu and Park, 1997). 16S rDNA sequencing and other characterization identified E681 as a strain of *Paenibacillus polymyxa*.

As a preliminary step towards the genome sequencing of E681, the genome size was measured by pulsed-field gel electrophoresis. Genomic DNA embedded in an agarose plug was digested with *NorI* or *PmeI*, and the resulting restriction fragments were electrophoresed. The sizes of each band were added up, and the genome size of E681 was estimated to be slightly larger than 5 Mb.

To construct sequencing libraries, genomic DNA of several different sizes ranging from two to ten kilobases was obtained by sonication or enzymatic treatment, and the DNA fragments were cloned into pUC, pBC, or pTB vector. Using ABI PRISM 3700, about 700 to 1,000 bases at both ends of each insert were determined. About six equivalents of the genome – total length of the reads after vector clipping and quality trimming amounting to 31 Mb – have been sequenced.

Using phrap (<http://www.phrap.org/>), the sequences were assembled into 1,084 contigs (494 contigs are ≥ 2 kb). The sum of the contigs reached 4.72 Mb, which corresponds to about 95% of the E681 genome. The average G+C content of the sequences was 45.55%, which is close to the experimentally determined value of the type strain. About ten 16S-23S-5S rRNA genes were found from the sequenced region. Also, seventy tRNA genes were identified by tRNAscan-SE (Lowe and Eddy, 1997).

From a preliminary annotation for the contigs by GLIMMER 2.02 (Delcher *et al.*, 1999), a total of 4,182 open reading frames (ORFs) were predicted. BLAST analysis (Altschul *et al.*, 1997) indicated that more than half of the E681 ORFs show highest similarities with genes in *Bacillus subtilis* and other low G+C Gram-positive bacteria (Fig. 1). About 56.2% of the ORFs could be functionally assigned by matching the deduced

proteins with entries in the COG database (<http://www.ncbi.nlm.nih.gov/COG/>), and further 20% were found to have homologs in the databases by BLAST searches. About 23.3% of the ORFs were unique to E681.

Interestingly, ORFs involved in biosynthesis, transport and catabolism of secondary metabolites were found in high frequencies in the E681 genome. This suggests that E681 produces various antimicrobial compounds and other secondary metabolites possibly to compete with other microbes in the rhizosphere and/or to communicate with each other. The E681 genome also had high proportions of ORFs involved in transcription, signal transduction, carbohydrate metabolism, and amino acid metabolism. It implies that E681 is well adapted to changing environmental conditions in the soil and the rhizosphere, and is capable of metabolizing a wide array of nutrients originating from plants.

Scaffolding of the contigs and gap closure by primer walking, end sequencing of bacterial artificial chromosome libraries, and other techniques are currently underway, and the progress will be presented.

References

1. Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
2. Ash, C., F. G. Priest, and M. D. Collins. 1993. Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Proposal for the creation of a new genus *Paenibacillus*. *Antonie Van Leeuwenhoek* 64, 253-260.
3. Beatty, P. H. and S. E. Jensen. 2002. *Paenibacillus polymyxa* produces fusaricidin-type antifungal antibiotics active against *Leptosphaeria maculans*, the causative agent of blackleg disease of canola. *Can. J. Microbiol.* 48, 159-169.
4. Bloemberg, G. V. and B. J. Lugtenberg. 2001. Molecular basis of plant growth promotion and biocontrol by rhizobacteria. *Curr. Opin. Plant Biol.* 4, 343-350.
5. Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 27, 4636-4641.
6. Emmert, E. A. and J. Handelsman. 1999. Biocontrol of plant disease: a (Gram-) positive perspective. *FEMS Microbiol. Lett.* 171, 1-9.
7. Fraser, C. M., J. A. Eisen, and S. L. Salzberg. 2000. Microbial genome sequencing. *Nature* 406, 799-803.
8. Lebuhn, M., T. Heulin, and A. Hartmann. 1997. Production of auxin and other indolic and phenolic compounds by *Paenibacillus polymyxa* strains isolated from different proximity to plant roots. *FEMS Microbiol. Ecol.* 22, 325-334.
9. Lowe, T. M. and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955-964.
10. Ryu, C.-M. and C. S. Park. 1997. Enhancement of plant growth induced by endospore forming PGPR strain, *Bacillus polymyxa* E681. In *Plant Growth-Promoting Rhizobacteria: Present Status and Future Prospects. Proceedings of the Fourth International Workshop on Plant Growth-Promoting Rhizobacteria, Japan-OECD Joint Workshop, Sapporo*, pp. 209-211.
11. Timmusk, S., B. Nicander, U. Granhall, and E. Tillberg. 1999. Cytokinin production by *Paenibacillus polymyxa*. *Soil Biol. Biochem.* 31, 1847-1852.

Fig. 1. Taxonomic distribution of highest BLAST hits. The number and the percentage of E681 ORFs showing the highest similarities with those in a given bacterial species or a taxonomic group are indicated along with the species or group name. Among low G+C Gram-positive bacteria, genome sequence information is available for *Bacillus subtilis*, *Bacillus halodurans*, *Clostridium acetobutylicum*, and several others.

