Draft Genome Sequence of the Plant Growth-Promoting Bacterium Bacillus siamensis KCTC 13613T

Haeyoung Jeong,a,b Da-Eun Jeong,a Sun Hong Kim,a Geun Cheol Song,a,b Soo-Young Park,a Choong-Min Ryu,a,b Seung-Hwan Park,a,b and Soo-Keun Choi*a,b

Bacillus siamensis KCTC 13613T, a novel halophilic Bacillus species isolated from a salted Thai food, produced antimicrobial compounds against plant pathogens and promoted plant growth by volatile emission. We determined the 3.8-Mb genome sequence of B. siamensis KCTC 13613T to reveal the plant-beneficial effect at the genomic level.

The genome of B. siamensis KCTC 13613T was determined using a whole-genome shotgun strategy and an Illumina HiSeq 2000 instrument. Paired-end reads of 4.96 Gb (average read length of 96.1 bp) were produced from a 490-bp genomic library. Quality trimming and de novo assembly (word size, 64) were performed using CLC Genomics Workbench, version 4.8. The assembly contained 3,779,696 bp in 51 large contigs over 200 bp (N50, 589,476 bp) with 46.3% G+C content. Maximum and average contig sizes were 965,626 bp and 74,112 bp, respectively. When Velvet, version 1.2.01, was used with various k-mers (55 to 75) (12), 38 scaffolds totaling 3,784,323 bp (N50, maximum scaffold size of 1,991,923 bp) were obtained. Not only were the two assembly results consistent with each other by direct comparison using MUMmer (6), but they were also the largest aligned blocks between two 100% identical results (593,857 bp), suggesting that the assembled sequences were accurate regardless of the assembly software. Automatic gene prediction and functional annotation of CLC assemblies were carried out using the RAST server (2). Among the 3,892 putative protein-coding genes, 47% were assigned to 446 subsystem categories. Average nucleotide identity analysis using completely sequenced Bacillus species showed that B. amyloliquefaciens was closely related to B. siamensis (>94% identity) (8). A polyketide synthase (PKS)/nonribosomal peptide synthetase (NRPS) gene analysis was performed using SBSPKS (1) and BLAST against known protein sequences from the NCBI nr database. We identified at least four gene clusters encoding bio-synthetic enzymes for one NRPS, two PKSs, and one PKS/NRPS hybrid. The sequence analysis revealed that they were all close to biosynthetic genes with known secondary metabolites such as fengycin, bacillaene, difficidin, and iturin, showing >90% identity at the amino acid level. In conclusion, this genome information may be useful for developing B. siamensis as a plant growth-promoting agent.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AJVF00000000. The version described in this paper is the first version, AJVF01000000.

ACKNOWLEDGMENTS

This study was supported by the 21C Frontier Microbial Genomics and Applications Center Program of the Ministry of Education, Science, and Technology, by the Technology Development Program for Agriculture and Forestry, Ministry for Agriculture, Forestry, and Fisheries, and by the KRIBB Research Initiative Program, Republic of Korea.

REFERENCES


Copyright © 2012, American Society for Microbiology. All Rights Reserved.

traditional food “kochujang” analyzed by a massive sequencing technique. J. Food Sci. 77:M250–M256.


