Draft Genome Sequence of *Paenibacillus popilliae* ATCC 14706<sup>T</sup>

Kazuhiro Iiyama<sup>1</sup>*, Kazuki Mori<sup>2</sup>, Hiroaki Mon<sup>3</sup>, Yuuka Chieda<sup>1</sup>, Jae Man Lee<sup>3</sup>, Takahiro Kusakabe<sup>3</sup>, Kousuke Tashiro<sup>2</sup>, Shin-ichiro Asano<sup>4</sup>, Chisa Yasunaga-Aoki<sup>1</sup> and Susumu Shimizu<sup>1</sup>

<sup>1</sup>Laboratory of Insect Pathology and Microbial Control, Institute of Biological Control, Faculty of Agriculture, Graduate School, Kyushu University, Japan
<sup>2</sup>Laboratory of Molecular Gene Technology, Faculty of Agriculture, Graduate School, Kyushu University, Japan
<sup>3</sup>Laboratory of Silkworm Science, Faculty of Agriculture, Graduate School, Kyushu University, Japan
<sup>4</sup>Laboratory of Applied Molecular Entomology, Faculty of Agriculture, Hokkaido University, Japan

(Received March 29, 2013; Accepted June 10, 2013)

Key words: *Paenibacillus popilliae*, draft genome sequence, entomopathogenic bacteria

INTRODUCTION

*Paenibacillus popilliae* (formerly *Bacillus popilliae*) is a rod-shaped, endospore-forming bacterium, and its Gram stain reaction is variable (Pettersson et al., 1999; Tanada and Kaya, 1992). The bacterium infects the Japanese beetle *Popillia japonica* and related species, and then causes milky disease. *P. popilliae* was the first insect pathogen to be registered in the United States as a microbial control agent targeted at the Japanese beetle (Tanada and Kaya, 1992). In this study, we report the draft genome sequence of *P. popilliae* ATCC 14706<sup>T</sup>.

MATERIALS AND METHODS

Bacterial culture, genome extraction, sequencing and annotation

*P. popilliae* ATCC 14706<sup>T</sup> was purchased from the American Type Culture Collection (ATCC). *P. popilliae* ATCC 14706<sup>T</sup> was cultured in MYPGP broth (Costilow and Coulter, 1971) at 30°C for 24 h. Genomic DNA was extracted from the cells by lysozyme-CTAB method (Moore et al., 2004).

A paired-end (fragment size 300-400 bp) library of the *P. popilliae* ATCC 14706<sup>T</sup> genome was generated. Sequencing was performed using HiSeq 2000 (Illumina, San Diego, CA, USA), yielding 46,295,320 reads and a total of 4,676 Mb. Assembly was performed using Velvet v.1.2.01 (Zerbino and Birney, 2008). Sequences were annotated with Microbial Genome Annotation Pipeline, MiGAP (Sugawara et al., 2009).

Phylogenetic analysis

In our previous report, phylogenetic analysis based on housekeeping genes of *Paenibacillus* species was demonstrated (Iiyama et al., 2013). Among these *Paenibacillus* species, draft or complete genome analyses of *P. dendritiformis* C454 (Sirota-Madi et al., 2012), *P. alvei* JCM 20131<sup>T</sup> (Djukic et al., 2012), *P. polymyxa* SC2 (Ma et al., 2011), and *P. larvae* subsp. *larvae* BRL2300-10 (Chan et al., 2011) were carried out and the data were made publicly available. The phylogenetic relationship of these *Paenibacillus* species was reanalyzed using the data set in the previous work by the methods described before (Iiyama et al., 2013). *B. subtilis* subsp. *subtilis* 168 (Belda et al., 2013) was used as the outgroup.

RESULTS AND DISCUSSION

Sequencing, assembly and annotation

The assembled sequences contained 583 contigs. Frequency of contig length was shown in Fig. 1. The longest contig was 160,279 bp. The length of all contigs combined was 3,833,720 bp with a G+C ratio of 51.0%. The N50 and N90 sizes of the contigs were 43,349 bp and 5,673 bp, respectively (Fig. 2). The sequence coverage was approximately 1,200× when genome size was assumed to be 3,833,720 bp. Since the genome of *Paenibacillus popilliae* ATCC 14706<sup>T</sup> was estimated to be 3,395 kb by pulsed-field gel electrophoresis (Macdonald and Kalmakoff, 1995), this draft genome was estimated to almost cover the full genome sequence. Since the number of contigs was high (583 contigs), improvement of draft assemblies will require a more intensive analysis with multiple assembly algorithms, additional libraries, and a second next generation sequencing technology.

A total of 3,855 protein-coding sequences (CDS), four rRNAs, and 75 tRNAs are encoded in the *P. popilliae* ATCC 14706<sup>T</sup> draft genome. Among the CDS, 2,802 (65.9%) were assigned to clusters of orthologous groups (COG) categories (Tatusov et al., 2000) (Fig. 3).

Comparison with other *Paenibacillus* species

*P. popilliae* and *P. larvae* subsp. *larvae* are entomo-
pathogenic bacteria and they have many common characteristics in bacterial and pathological properties (Iiyama et al., 2013). However, the phylogenetic analysis that was based on sequences of 16S ribosomal RNA (rRNA) genes showed that the relationship between these bacterial species was not close (Pettersson et al., 1999). This overall topology was supported by alternative phylogenetic markers such as glyceraldehyde 3-phosphate dehydrogenase A, GroEL protein, DNA gyrase subunit A and glucose-6-phosphate isomerase (Iiyama et al., 2013). In this study, phylogenetic analysis based on DNA gyrase subunit A was reanalyzed (Fig. 4). Both the estimated genome sizes and the number of CDS in \textit{P. popilliae} and \textit{P. larvae} subs. \textit{larvae} were smaller than those of other \textit{Paenibacillus} species used in this analysis. Elimination of particular genes may correlate with acquisition of pathogenicity toward insects. However, the result shown in Fig. 4 is completely insufficient for proving this hypothesis, although it is a basis for proposal of this hypothesis. To evaluate this hypothesis, a detailed study of genome comparison among \textit{Paenibacillus} species is needed.

This report is the first of the draft genome sequence of \textit{P. popilliae}, and a further analysis of this genome and a comparative analysis with related bacteria will contribute to understanding the pathogenicity of \textit{P. popilliae} in insects.
The 583 contigs contained in the genome have been deposited at DDBJ/EMBL/GenBank under the accession numbers BALG01000001 to BALG01000583.

ACKNOWLEDGEMENTS

We would like to thank Dr. Itsuro Sugimura (Hokkaido System Science Co., Ltd.) for his technical support. This work was supported in part by JSPS KAKENHI Grant Number 21380037.

REFERENCES


