JB Journals.ASM.org

GENOME ANNOUNCEMENT

Draft Genome Sequence of *Bacillus pumilus* BA06, a Producer of Alkaline Serine Protease with Leather-Dehairing Function

Chuan-Wu Zhao, Hai-Yan Wang, Yi-Zheng Zhang, and Hong Feng

Sichuan Key Laboratory of Molecular Biology and Biotechnology, Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, Sichuan, People's Republic of China

***Bacillus pumilus* BA06 was isolated from the proteinaceous soil and produced an extracellular alkaline protease with leather-dehairing function. The genome of BA06 was sequenced. The comparative genome analysis indicated that strain BA06 is different in genome from the other *B. pumilus* strains, with limited insertions, deletions, and rearrangements.**

Alkaline serine protease is widely used in various industries (15). Representative subtilisins from *Bacillus subtilis* and the related species have been extensively studied for many years (3, 13). Recently, alkaline serine proteases from various strains of *Bacillus pumilus* were characterized with unique biochemical characteristics, such as higher thermostability and catalytic efficiency (6, 9, 10, 12, 16). *B. pumilus* BA06, isolated from the proteinaceous soil, is able secrete a major alkaline serine protease that shows good potential in leather dehairing (8, 17). The gene encoding this protease was cloned and expressed in *B. subtilis* (WB600) (11). In order to get deep insights into transcriptional regulation and for comparative genome analysis, we sequenced the genome of *B. pumilus* BA06 using Illumina HiSeq 2000 at Shenzhen Huada Genomics Institute (China).

The sequencing generated 5,555,556 paired-end reads that were 90 bp in length, with genome coverage of about 250×. The clean reads were assembled primarily using the software ABySS (14) and Edena (7). The resulting two sets of contigs (>200 bp, 23 each) were aligned using BLAST (1) to improve the assembly results from ABySS; for that, several in-house Python (www.python.org) scripts were used to merge contigs and replace the Ns and other ambiguous bases. Finally, 15 contigs, with an N_{50} of 630,692 bp and a largest contig of 835,516 bp, were obtained. In total, the draft genome of BA06 contains 3,747,698 bp with a G+C content of 41.3%, about 0.43 Mb more than the genome of *B. pumilus* SARF-32. The genome phylogenetic trees inferred from the whole orthologs and constructed using PhyML (5) showed that BA06 was close to (in order from near to far) *B. pumilus* SAFR-32, *B. pumilus* ATCC 7061, and *Bacillus licheniformis* ATCC 14580.

The draft genome was annotated by submitting it to the RAST system (2). Consequently, 3,890 protein-coding sequences (CDSs) were identified, among which 2,773 CDSs were able to be assigned to one of the 461 RAST subsystems. Numbers of predicted genes belonging to metabolisms of DNA, nitrogen, protein, carbohydrates, phosphorous, and potassium did not show obvious differences from those for *B. pumilus* SARF-32 and even *B. subtilis* 168. In contrast, the numbers of CDSs in the BA06 genome belonging to the subsystems of the membrane transporter, sulfur metabolism, mobility and chemotaxis, regulation, and cell signaling were much more than those for SARF-32. Comparative genome analysis was also carried out with the closest strains of *Bacillus* using the software Mauve (4). Overall, the sequences of BA06 and SARF-32 were colinear, with only limited insertions, deletions, and rearrangements. Compared to the BA06 genome, a prophage, type I restriction modification system, and some hypothetical CDSs did not present in SARF-32. Several operons or their members that are involved in assimilation of sulfate and thiosulfate, alkanesulfonates, nitrate, and arsenate were different in or omitted from the strain SARF-32. All the differences between those strains may result from the fitness to various niches.

Nucleotide sequence accession number. The sequence of *B. pumilus* BA06 from this whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AMDH00000000.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31171204) and the National High Technology Research and Development Program of China (2006AA02Z221).

REFERENCES

1. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.
2. **Aziz RK, et al.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75.
3. **Bryan PN.** 2000. Protein engineering of subtilisin. Biochim. Biophys. Acta **1543**:203–222.
4. **Darling AC, Mau B, Blattner FR, Perna NT.** 2010. ProgressiveMauve: multiple alignment with gene gain, loss and rearrangements. PLoS One **5**:e11147. doi:10.1371/journal.pone.0011147.
5. **Guindon S, Gascuel O.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**:696–704.
6. **Han X-Q, Damodaran S.** 1997. Isolation, identification, and fermentation of a *Bacillus* species producing a detergent-stable endopeptidase. J. Agric. Food Chem. **45**:4191–4195.
7. **Hernandez D, François P, Farinelli L, Osteras M, Schrenzel J.** 2008. *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. Genome Res. **18**:802–809.
8. **Huang Q, Peng Y, Li X, Wang H-Y, Zhang Y-Z.** 2003. Purification and characterization of an extracellular alkaline serine protease with dehairing function from *Bacillus pumilus*. Curr. Microbiol. **46**:169–173.
9. **Jaouadi B, Ellouz-Chaaboum S, Rhimi M, Bejar S.** 2008. Biochemical and molecular characterization of a detergent-stable serine alkaline pro-

Received 10 September 2012 **Accepted** 28 September 2012

Address correspondence to Hong Feng, hfeng@scu.edu.cn.

doi:10.1128/JB.01694-12

tease from *Bacillus pumilus* CBS with high catalytic efficiency. Biochimie **90**:1291–1305.
10. **Kumar CG.** 2002. Purification and characterization of a thermostable alkaline protease from alkalophilic *Bacillus pumilus*. Lett. Appl. Microbiol. **34**:13–17.
11. **Pan J, Huang Q, Zhang Y-Z.** 2004. Gene cloning and expression of an alkaline serine protease with dehairing function from *Bacillus pumilus*. Curr. Microbiol. **49**:165–169.
12. **Rahman RNZRA, Mahamad S, Salleh AB, Basri M.** 2007. A new organic solvent tolerant protease from *Bacillus pumilus*. J. Ind. Microbiol. Biotechnol. **34**:509–517.
13. **Rao MB, Tanksale AM, Ghatge MS, Deshpande VV.** 1998. Molecular and biotechnological aspects of microbial proteases. Microbiol. Mol. Biol. Rev. **62**:597–635.
14. **Simpson JT, et al.** 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. **19**:1117–1123.
15. **Verma J, Modi DR, Sharma R, Saxena S.** 2011. Vital role of alkaline protease in bio-industries: a review. Plant Arch. **11**:1083–1092.
16. **Wan M-Y, Wang H-Y, Zhang Y-Z, Feng H.** 2009. Substrate specificity and thermostability of the dehairing alkaline protease from *Bacillus pumilus*. Appl. Biochem. Biotechnol. **159**:394–403.
17. **Wang H-Y, et al.** 2007. Screening and mutagenesis of a novel *Bacillus pumilus* strain producing alkaline protease for dehairing. Lett. Appl. Microbiol. **44**:1–6.