AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Sequence of *Bacillus clausii* AKU0647, a Strain That Produces Endo-β-*N*-Acetylglucosaminidase A

Yujiro Higuchi, Kazuki Mori, Akiko Suyama, Yibo Huang, Kosuke Tashiro, Satoru Kuhara, Kaoru Takegawa

Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, Fukuoka, Japan

To comprehensively identify glycosyl hydrolase genes in the genome of *Bacillus clausii* strain AKU0647, which produces endo-β-*N*-acetylglucosaminidase A (Endo-A), we conducted whole-genome shotgun sequencing. We identified several other putative glycosyl hydrolase genes apart from the Endo-A gene, and report these findings here.

Received 3 March 2016 **Accepted** 11 March 2016 **Published** 5 May 2016

Citation Higuchi Y, Mori K, Suyama A, Huang Y, Tashiro K, Kuhara S, Takegawa K. 2016. Draft genome sequence of *Bacillus clausii* AKU0647, a strain that produces endo-β-*N*-acetylglucosaminidase A. Genome Announc 4(3):e00310-16. doi:10.1128/genomeA.00310-16.

Address correspondence to Kaoru Takegawa, takegawa@agr.kyushu-u.ac.jp.

Endo-β-*N*-acetylglucosaminidase (ENGase) is an enzyme that hydrolyzes the N,N′-diacetylchitobiose moiety of the asparagine-linked oligosaccharides of glycoproteins. ENGases belong to two glycosyl hydrolase families, 18 (GH18) and 85 (GH85); distinctively, the former does not exhibit transglycosylation activity but the latter does. The GH85 ENGase A (Endo-A) was originally identified in *Arthrobacter protophormiae* AKU0647 when the strain was cultured in a medium containing ovalbumin, and the enzyme has been well characterized (1–5). However, how this microbe degrades such glycoproteins is not well understood. Therefore, to identify the total complement of glycosyl hydrolase genes in the genome of strain AKU0647, we generated a draft genome sequence.

A whole-genome shotgun sequencing strategy was carried out to acquire the data for the draft sequence. Sequencing was conducted using MiSeq (Illumina) to obtain 300 bp paired-end reads. A total of 5.18 Gbp of data were generated from 1.76×10^7 sequencing reads (1,112-fold-coverage). By assembling the sequence data using the program *Platanus* version 1.2.4, 16 contigs were obtained. The length of the longest contig is 1,277,614 bp and the N_{50} size is 713 kb. Genome annotation was performed with Glimmer version 3.02b, BLAST, and InterProScan to identify genes and predict their functions. The draft genome of strain AKU0647 is 4.66 Mbp with an average G+C content of 44.3% and contains 5,047 genes. The gene density is 923 bp/gene. The median and mean gene lengths are 744 bp and 808 bp, respectively.

Unexpectedly, the analysis of this draft sequence revealed that the genome of strain AKU0647 exhibits the highest sequence identity to that of *Bacillus clausii*, not that of *A. protophormiae*, with 100% identity of the 16S rRNA gene sequence. Thus, we renamed this sequenced strain as *B. clausii* AKU0647. According to our annotation of the whole-genome sequence, we confirmed the presence of the Endo-A gene and found two genes of ENGase belonging to GH18, which were designated ORF2621, ORF1208, and ORF2421, respectively. ORF1208 is located on contig0001 and both ORF2421 and ORF2621 are on contig0003 (accession numbers BCXJ01000001 and BCXJ01000003, respectively). Intriguingly, we also discovered four other genes encoding putative glycosyl hydrolases near the locus of the Endo-A gene, all of which might function together with Endo-A to hydrolyze oligosaccharides of glycoproteins. In total, we found 69 putative glycosyl hydrolase genes in the genome of *B. clausii* AKU0647. This draft genome sequence analysis will provide a solid basis for elucidating the enzymatic mechanisms of glycoprotein oligosaccharide degradation by *B. clausii* AKU0647.

Nucleotide sequence accession numbers. The contig sequences of *Bacillus clausii* strain AKU0647 have been deposited at DDBJ/EMBL/GenBank under the accession numbers BCXJ01000001 to BCXJ01000016.

ACKNOWLEDGMENTS

We thank Jun Ogawa and Hiroya Yurimoto at Kyoto University for providing samples. We also appreciate the support from the Center for Advanced Instrumental and Educational Supports, Faculty of Agriculture, Kyushu University with DNA sequencing using Illumina MiSeq.

This study was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (K. Takegawa).

FUNDING INFORMATION

This work, including the efforts of Kaoru Takegawa, was funded by JSPS KAKENHI.

REFERENCES

1. **Takegawa K, Nakoshi M, Iwahara S, Yamamoto K, Tochikura T.** 1989. Induction and purification of endo-β-*N*-acetylglucosaminidase from *Arthrobacter protophormiae* grown in ovalbumin. Appl Environ Microbiol **55**:3107–3112.
2. **Takegawa K, Tabuchi M, Yamaguchi S, Kondo A, Kato I, Iwahara S.** 1995. Synthesis of neoglycoproteins using oligosaccharide-transfer activity with endo-β-*N*-acetylglucosaminidase. J Biol Chem **270**:3094–3099. http://dx.doi.org/10.1074/jbc.270.7.3094.

3. **Takegawa K, Yamabe K, Fujita K, Tabuchi M, Mita M, Izu H, Watanabe A, Asada Y, Sano M, Kondo A, Kato I, Iwahara S.** 1997. Cloning, sequencing, and expression of *Arthrobacter protophormiae* endo-β-*N*-acetylglucosaminidase in *Escherichia coli*. Arch Biochem Biophys **338:** 22–28. http://dx.doi.org/10.1006/abbi.1996.9803.
4. **Fujita K, Takegawa K.** 2001. Chemoenzymatic synthesis of neoglycoproteins using transglycosylation with endo-β-*N*-acetylglucosaminidase A. Biochem Biophys Res Commun **282:**678–682. http://dx.doi.org/10.1006/bbrc.2001.4631.
5. **Huang W, Yang Q, Umekawa M, Yamamoto K, Wang LX.** 2010. Arthrobacter endo-β-*N*-acetylglucosaminidase shows transglycosylation activity on complex-type *N*-glycan oxazolines: one-pot conversion of ribonuclease B to sialylated ribonuclease C. Chembiochem **11:**1350–1355. http://dx.doi.org/10.1002/cbic.201000242.