Draft Genome Sequence of *Talaromyces cellulolyticus* Strain Y-94, a Source of Lignocellulosic Biomass-Degrading Enzymes

Tatsuya Fujii,a Hideaki Koike,a Shigeki Sawayama,b Shinichi Yano,c Hiroyuki Inoueb

BiomaSS Refinery Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Higashi-Hiroshima, Hiroshima, Japan; Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan; Division of Applied Biosciences, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto, Japan

*Talaromyces cellulolyticus* (formerly *Acremonium cellulolyticus*) is a promising fungus for cellulase production. Here, we present the draft genome sequence of *T. cellulolyticus* strain Y-94. The genome is 36.4 Mbp long and contains genes for several enzymes involved in the degradation of lignocellulosic biomass, including cellulases, hemicellulases, pectinases, and amylases.

Received 7 January 2015 Accepted 13 January 2015 Published 26 February 2015


Copyright © 2015 Fujii et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported license.

Address correspondence to Hiroyuki Inoue, inoue-h@aist.go.jp.

*Talaromyces cellulolyticus* strain Y-94 (CBS 136886, FERM BP-5826), which was isolated in Japan (1), is one of several promising filamentous fungi for the industrial production of cellulase and hemicellulase to hydrolyze lignocellulosic biomass (2–5). The taxonomic classification of this organism has been revised from the genus *Acremonium* to *Talaromyces* (6). Recently, a transformation system with a uracil auxotrophic strain has been developed, and some glycoside hydrolase (GH) family enzymes and transcription factors have been analyzed (7–11).

We present here the draft genome sequence of *T. cellulolyticus* Y-94. 454/Roche (FLX Titanium) and Illumina Genome Analyzer II sequencers were used in this study. *T. cellulolyticus* Y-94 genomic DNA was sheared into 3-kb paired-end fragments to generate a library according to the Roche paired-end library preparation method manual. A 454 Titanium draft library with 699 Mb of total reads and an average read length of 391 bases, which provided 19-fold coverage of the genome, was generated. Reads were assembled using the gsAssembler (454 Life Sciences, Roche Applied Science, Branford, CT, USA), and 60 scaffolds were obtained. The genome size of *T. cellulolyticus* predicted by the total scaffold length was 36.4 Mbp. Illumina Genome Analyzer II sequencing was done to account for 454/Roche sequencing errors. In total, 10,980 open reading frames (ORFs) were predicted using AUGUSTUS (12) trained on *Aspergillus nidulans* FGSC A4 gene models.

We screened the draft sequence for GH family genes. At least 249 ORFs were annotated as GH family proteins, including 133 potentially secreted proteins (based on a SignalP version 4.1 analysis and our secretome data). The number of GH family genes of *T. cellulolyticus* was similar to that of *A. nidulans* (247 genes), less than that of *Aspergillus oryzae* (285 genes), and greater than that of *Trichoderma reesei* (200 genes) (13). Based on the CAZy database (14), carbohydrate-active enzymes related to the hydrolysis of lignocellulosic biomass were identified: 22 cellulases (12 GH5s [including hemicellulases such as mannanase], 1 GH6, 2 GH7s, 4 GH12s, 1 GH61, and 2 GH45s); 37 hemicellulases (22 GH43s, 1 GH10, 7 GH11s, 1 GH74, 1 GH62, 1 GH53, 1 GH54, 2 GH67s, and 1 GH26); 38 pectinases (16 GH28s, 12 GH78s, 4 PL1s, 2 PL4s, 2 CE8s, and 2 CE12s); and 8 amylases (5 GH13s and 3 GH15s). These results support the ability of *T. cellulolyticus* to degrade various types of biomass (1, 3, 10, 15). The genome sequence data will provide suggestions for improvements in cellulase and hemicellulase production in *T. cellulolyticus*.

**Nucleotide sequence accession numbers.** The nucleotide sequence of *T. cellulolyticus* strain Y-94 has been deposited at DDBJ/EMBL/GenBank as follows: 1,723 contigs under accession numbers BBPS0100001 to BBPS01001723 and 60 scaffolds under accession numbers DF933979 to DF933856. The version described in this study is the first version.

**ACKNOWLEDGMENTS**

This work was funded by a research grant from the National Institute of Advanced Industrial Science and Technology.

We thank Tamotsu Hoshino, Kazuhiko Ishikawa, Kenichi Nakayama, and Katsui Murakami (National Institute of Advanced Industrial Science and Technology) for helpful discussions.

**REFERENCES**


