

# Draft Genome Sequence of *Aspergillus oryzae* Strain 3.042

Guozhong Zhao, Yunping Yao, Wei Qi, Chunling Wang, Lihua Hou, Bin Zeng, and Xiaohong Cao

Key Laboratory of Food Nutrition and Safety, Tianjin University of Science and Technology, Ministry of Education, Tianjin, China

***Aspergillus oryzae* is the most important fungus for the traditional fermentation in China and is particularly important in soy sauce fermentation. We report the 36,547,279-bp draft genome sequence of *A. oryzae* 3.042 and compared it to the published genome sequence of *A. oryzae* RIB40.**

The filamentous fungus *Aspergillus oryzae* is well known for its capacity to secrete large amounts of hydrolytic enzymes, such as protease, cellulase, and amylase (2, 6). *A. oryzae* strain 3.042, originally isolated from soil, is extensively used for the production of both soy sauce and other fermented foods in China (7).

The genome sequence of *A. oryzae* strain 3.042 was determined with a combined strategy using Roche 454 and Solexa paired-end, and Solexa mate-paired sequencing technology. The paired-end reads (2,560 Mbp; 70× coverage) and mate-paired reads (2,867 Mbp; 77× coverage) generated by Solexa sequencer were assembled by SOAPdenovo (4), and then the 454 reads (900 Mbp; 24× coverage) and the split fragments of contigs generated by SOAPdenovo were used for hybrid assembly with the Newbler sequence assembler (version 2.3).

The draft genome of *A. oryzae* 3.042 consists of 225 sequence contigs with a total length of 36,547,279 bp and a G+C content of 48%. The numbers of protein-coding genes and tRNAs predicted were 11,397 and 243. The average gene density is one gene per 1.911 kb. The protein-coding sequence occupies 44% of the sequenced portion of the genome of strain 3.042. An estimated total of 23,470 introns, ranging from 16 to 2,753 nucleotides long, with a mean length of 97 nucleotides, are distributed among 98% of strain 3.042 genes. Comparative genomic analysis was performed with the published genome of *A. oryzae* RIB40 (5).

With the availability of the draft genome sequence, the complete mitochondrial genome sequence of *A. oryzae* 3.042 was determined, representing a circular DNA molecule of 29,192 bp with a G+C content of 26%. Compared to strain RIB40, most of the functional genes of strain 3.042 are conserved. We found hundreds of unique genes in strains 3.042 and RIB40. Some unique genes in the transport systems of the two strains, like the major facilitator superfamily (MFS) transporters, amino acid transporters, and predicted flavoprotein involved in K<sup>+</sup> transport, were found, indicating that some different substrates can be transported into the cells and that the two strains may have different environmental adaptations (1). Strain 3.042 possesses one CipC protein that contributed to hyperbranching, whereas strain RIB40 has SNARE protein SED5, chitin synthase, etc., which may explain the phenotype differences between the two strains under the same culture conditions (3). We also found some unique genes related to ester formation and amino acid metabolism which make contributions to flavor formation.

In conclusion, the genome sequence of *A. oryzae* 3.042 provides new data with which to further explore gene-based functional mechanisms of *A. oryzae*. Moreover, comparative genomics analysis and functional genomics analysis would clarify the reasons for different transport systems and growth rates of the two strains and help trace flavor production by *A. oryzae*.

**Nucleotide sequence accession number.** The draft genome sequence and annotation information have been deposited in DDBJ/EMBL/GenBank under accession no. [AKHY00000000](https://doi.org/10.1093/nucleic-acids/gkz000). Contig 85 is labeled as a mitochondrion and is a complete circular mitochondrial sequence with no gaps under accession no. [JX129489](https://doi.org/10.1093/nucleic-acids/gkz000).

## ACKNOWLEDGMENTS

This work was supported by project grants 2012BAD33B04, 10ZCZDSY07000, SS2012AA023408, 31171731, 210-bk130006, and IRT1166.

## REFERENCES

1. Hamari Z, et al. 2009. Convergent evolution and orphan genes in the Fur4p-like family and characterization of a general nucleoside transporter in *Aspergillus nidulans*. *Mol. Microbiol.* 73:43–57.
2. Hara S, Jin FJ, Takahashi T, Koyama Y. 2012. A further study on chromosome minimization by protoplast fusion in *Aspergillus oryzae*. *Mol. Genet. Genomics* 287:177–187.
3. Jin FJ, et al. 2011. SclR, a basic helix-loop-helix transcription factor, regulates hyphal morphology and promotes sclerotial formation in *Aspergillus oryzae*. *Eukaryot. Cell* 10:945–955.
4. Li R, et al. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20:265–272.
5. Machida M, et al. 2005. Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438:1157–1161.
6. McKelvey SM, Murphy RA. 7 October 2011. Biotechnological use of fungal enzymes. *Fungi* doi:10.1002/9781119976950.ch7.
7. Zhao G, et al. 2012. Construction of the mutant strain in *Aspergillus oryzae* 3.042 for abundant proteinase production by the N<sup>+</sup> ion implantation mutagenesis. *Int. J. Food Sci. Technol.* 47:504–510.

Received 6 June 2012 Accepted 27 June 2012

Address correspondence to Xiaohong Cao, [zhaoguozhongsun@gmail.com](mailto:zhaoguozhongsun@gmail.com).

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

doi:10.1128/EC.00160-12