Draft Genome Sequence of the *Aspergillus terreus* High-Itaconic-Acid-Productivity Strain IFO6365

Hiro Takahashi,[a,b,c] Tomoyuki Minami,[d] Mitsuyasu Okabe,[e] Enoch Y. Park,[f,g,h] Tsukasa Fujimoto,[b] Anna Takahashi,[i] Masataka Murase,[a] Shuichi Fukuyoshi,[j] Kenji Satou,[k] Shin Kanamasa[d,l]

aGraduate School of Medical Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

bGraduate School of Horticulture, Chiba University, Matsudo, Chiba, Japan

cFundamental Innovative Oncology Core Center, National Cancer Center, Tokyo, Japan

dGraduate School of Bioscience and Biotechnology, Chubu University, Kasugai, Aichi, Japan

eAureo Co., Ltd., Tokyo, Japan

fResearch Institute of Green Science and Technology, Shizuoka University, Shizuoka, Japan

gGraduate School of Integrated Science and Technology, Shizuoka University, Shizuoka, Japan

hGraduate School of Science and Technology, Shizuoka University, Shizuoka, Japan

iFaculty of Information Technologies and Control, Belarusian State University of Informatics and Radio Electronics, Minsk, Belarus

jInstitute of Medical, Pharmaceutical, and Health Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan

kFaculty of Biological Science and Technology, Institute of Science and Engineering, Kanazawa University, Kanazawa, Ishikawa, Japan

lCollege of Bioscience and Biotechnology, Chubu University, Kasugai, Aichi, Japan

Hiro Takahashi and Shin Kanamasa contributed equally to this work. Hiro Takahashi conducted the dry analysis and Shin Kanamasa conducted the wet laboratory experiments; author order was determined by taking into consideration the importance of the dry analysis in this study.

ABSTRACT Itaconic acid is an important organic acid used in the chemical industry. *Aspergillus terreus* strain IFO6365 is one of the highest-yielding itaconic acid-producing wild-type strains. Here, we report the draft genome sequence of IFO6365, enhancing the understanding of the role and biosynthesis of itaconic acid in this fungus.

Itaconic acid is an organic acid that is widely used as a raw material in the production of different chemicals. Industrially, the compound is produced by *Aspergillus terreus* due to its high yield (1). *A. terreus* IFO6365 is one of the strains with the highest yields of itaconic acid, despite being a wild-type strain (2). We previously cloned a gene encoding *cis*-aconitic acid decarboxylase that was essential for itaconic acid biosynthesis in this fungus (3). The Broad Institute sequenced the *A. terreus* NIH2624 genome at 11.05× coverage using whole-genome shotgun sequencing (4); the draft genome assembly, consisting of 26 scaffolds, is available in the EnsemblFungi database (https://fungi.ensembl.org/Aspergillus_terreus). We recently sequenced the genome of *A. terreus* TN-484 (5, 6), an itaconic acid productivity mutant derived from IFO6365. To obtain parental strain information for TN-484, genomic analysis of IFO6365 was performed.

IFO6365 cells, cultured in potato dextrose broth at 30°C, were homogenized using a µT-12 bead crusher (Taitec Co., Saitama, Japan), and genomic DNA was isolated using a NucleoSpin plant II kit (TaKaRa Bio, Inc., Shiga, Japan). Using a NEBNext Ultra II FS DNA library preparation kit for Illumina (New England Biolabs, Ipswich, MA, USA), DNA fragmentation (average fragment length, 600 bp) using DNA fragmentase, end repair, A tailing, adaptor ligation, PCR, and library purification were conducted for library preparation according to the manufacturer's instructions. Sequencing of the TruSeq DNA library (paired-end 300-bp reads) generated 17,711,412 reads. The adapter sequence removal and low-quality base trimming, followed by *de novo* assembly, were conducted using CLC Genomics Workbench version 12.0.3 (Qiagen, Valencia, CA, USA)

with default parameters. Then, reference-guided scaffolding of the draft genome was conducted using RaGOO version 1.1 (7), with default parameters, and the *A. terreus* NIH2624 genome sequence (GenBank accession number GCA_000149615). Among the RaGOO-generated scaffolds, a concatenated, unlocalized scaffold was discarded, because the scaffold consisted of short, highly fragmented contigs that had no homology to the reference sequence. The resulting genome assembly had a length of 28,633,404 bp and was divided into 23 scaffolds. The contig N_{50} value was 2,163,102 bp, the GC content was 53.1%, and the genome coverage was 170.3×. Of the 290 benchmarking universal single-copy ortholog (BUSCO) genes, 98.2% (including 1.0% duplicated genes) were found in the assembly, as calculated by BUSCO version 3.1.0 (8) using the parameter sp aspergillus_terreus. Coding region prediction for the scaffolds was conducted using AUGUSTUS version 3.2.2 (9) with the following parameters: noInFrameStop=true, genemodel=complete, and species=aspergillus_terreus. The estimated number of genes in the draft genome was 10,189. Gene annotation was also performed using OmicsBox version 1.2.4 (Qiagen) (10) with default parameters. This genomic information could provide insights into the genetic basis of itaconic acid production by IFO6365 origin mutants, such as TN-484 (5, 6).

Data availability. The draft genome sequence for strain IFO6365 has been deposited in GenBank/ENA/DDBJ under accession number BLJY00000000 (contig accession numbers BLJY01000001 to BLJY01000023). The SRA/DRA/ERA accession number is DRA009452.

REFERENCES

1. Calam CT, Oxford AE, Raistrick H. 1939. Studies in the biochemistry of micro-organisms: itaconic acid, a metabolic product of a strain of *Aspergillus terreus* Thom. Biochem J 33:1488–1495. https://doi.org/10.1042/bj0331488.
2. Okabe M, Lies D, Kanamasa S, Park EY. 2009. Biotechnological production of itaconic acid and its biosynthesis in *Aspergillus terreus*. Appl Microbiol Biotechnol 84:597–606. https://doi.org/10.1007/s00253-009-2132-3.
3. Kanamasa S, Dwiarti L, Okabe M, Park EY. 2008. Cloning and functional characterization of the *cis*-aconitic acid decarboxylase (CAD) gene from *Aspergillus terreus*. Appl Microbiol Biotechnol 80:223–229. https://doi.org/10.1007/s00253-008-1523-1.
4. Mabey JE, Anderson MJ, Giles PF, Miller CJ, Attwood TK, Paton NW, Bornberg-Bauer E, Robson GD, Oliver SG, Denning DW. 2004. C A DRE: the Central *Aspergillus* Data REpository. Nucleic Acids Res 32:D401–D405. https://doi.org/10.1093/nar/gkh009.
5. Yahiro K, Takahama T, Park Y, Okabe M. 1995. Breeding of *Aspergillus terreus* mutant TN-484 for an itaconic acid production with high yield. J Ferment Bioeng 79:506–508. https://doi.org/10.1016/0922-338X(95)91272-7.
6. Kanamasa S, Minami T, Okabe M, Park EY, Fujimoto T, Takahashi A, Murase M, Fukuyoshi S, Oda A, Satou K, Takahashi H. 2019. Draft genome sequence of *Aspergillus terreus* high-itaconic-acid-productivity mutant TN-484. Microbiol Resour Announc 8:e01170-19. https://doi.org/10.1128/MRA.01170-19.
7. Alonge M, Soyk S, Ramakrishnan S, Wang X, Goodwin S, Sedlazeck FJ, Lippman ZB, Schatz MC. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. Genome Biol 20:224. https://doi.org/10.1186/s13059-019-1829-6.
8. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
9. Hoff KJ, Stanke M. 2013. WebAUGUSTUS: a Web service for training AUGUSTUS and predicting genes in eukaryotes. Nucleic Acids Res 41:W123–W128. https://doi.org/10.1093/nar/gkt418.
10. BioBam Bioinformatics. 2019. OmicsBox: bioinformatics made easy (version 1.2.4). BioBam Bioinformatics, Valencia, Spain.