Draft Genome Sequence of *Saccharomyces cerevisiae* Strain P-684, Isolated from *Prunus verecunda*

Hiro Takahashi,a,b,c Takayuki Yoshizaki,d Hisashi Kondo,d Taichiro Motomura,a Masataka Murase,a Anna Takahashi,a,f Shuichi Fukuyoshi,g Chiyoko Machida,h Shin Kanamasa,f,h Hiromi Shimizu,i Shin-ichi Iwaguchi;j

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
dFaculty of Life Science and Biotechnology, Fukuyama University, Fukuyama, Hiroshima, Japan
eFaculty of Information Technologies and Control, Belarusian State University of Informatics and Radio Electronics, Minsk, Belarus
fCollege of Bioscience and Biotechnology, Chubu University, Kasugai, Aichi, Japan
gInstitute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kanazawa, Ishikawa, Japan
hGraduate School of Bioscience and Biotechnology, Chubu University, Kasugai, Aichi, Japan
iNara Prefecture Institute of Industrial Development, Nara, Nara, Japan
jFaculty of Science, Nara Women's University, Nara, Nara, Japan

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

ABSTRACT *Saccharomyces cerevisiae* strain P-684 is a yeast isolated from the flowers of *Prunus verecunda* 'Antiqua,' producing high quantities of malic and succinic acids in sake brewing. Here, we report the draft genome sequence of P-684, enlightening the mechanisms of biosynthesis of these organic acids by this strain.

Sake is a traditional Japanese alcoholic beverage made from fermented rice. One of its characteristics is that it is composed of various organic acids, especially malic acid and succinic acid, which are mostly produced by yeast and whose acidity greatly affects the taste of the beverage (1). We isolated *Saccharomyces cerevisiae* strain P-684 from the flowers of *Prunus verecunda* 'Antiqua' (cherry blossom, naranoyaezakura) using an enriched culture made from koji extract and used this yeast for sake brewing (2). The products brewed using P-684 had higher malic and succinic acid contents than did the sake brewed using other yeast strains (2). Thus, in this study, to characterize the above features at a molecular level, we analyzed the genome sequence of the yeast strain P-684.

P-684 cells were cultured in yeast extract-peptone-dextrose (YPED) medium at 25°C. Then, they were collected, and their genomic DNA was isolated using a GenTLE (from yeast) high-recovery kit (TaKaRa Bio, Inc.). Using the DNA library prep kit of Beijing Genomics Institute (BGI), DNA fragmentation (average fragment length, 300 bp) using a g-TUBE device (Covaris, Inc.), end repair, A tailing, adaptor ligation, PCR, and library purification were conducted to prepare the library, according to the manufacturers’ instructions. The sequencing of the DNBseq DNA library (paired-end 150-bp reads) generated 16,621,910 reads on the DNBSEQ-G400 platform (MGI Tech, Shenzhen, China). The adapter sequence removal and low-quality base trimming, followed by *de novo* assembly, were conducted using the “trim read” and “*de novo* assembly” modules on the CLC Genomics Workbench, version 20.0.4 (Qiagen, Valencia, CA, USA), using the default parameters. Then, reference-guided scaffolding of the draft genome sequence was conducted using RaGOO version 1.1 (3), using the default parameters, and the *S. cerevisiae* S288c genome sequence (GenBank accession number GCF_000146045.2) as the reference.
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 was 11,492,082 bp long and was divided into 17 scaffolds comprising 901 contigs. The N50 values (contig and scaffold), GC content, and genome coverage were 34,008 bp and 883,925 bp, 38.1%, and 207.7×, respectively. Of the 1,759 benchmarking universal single-copy ortholog (BUSCO) genes, 97.2% (including 5.6% of duplicated genes) were found in the assembly, as calculated using BUSCO version 3.1.0 (4) and by using the parameter “sp=saccharomyces_cerevisiae_S288C” and the “Saccharomyces_cerevisiae_S288C” data set. The prediction of the coding region of the scaffolds was conducted using AUGUSTUS version 3.3.3 (5) with the following parameters: “noInFrameStop=true,” “genemodel=complete,” and “species=saccharomyces_cerevisiae_S288C.” The estimated number of genes in the draft genome was 5,277. Gene annotation was also performed using Trinotate version 3.2.1 (6) with the default parameters.

This genomic information could provide insights into the genetic basis of the characteristics of this yeast in sake brewing.

Data availability. The draft genome sequence and gene annotation for strain P-684 have been deposited in GenBank/ENA/DDBJ under accession number BLZQ01000001 (BLZQ01000001 to BLZQ01000017). The SRA/DRA/ERA accession number is DRA010485.

ACKNOWLEDGMENTS

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant numbers JP18H03330, JP18K06297, JP18KT0047, and JP19K22892) and the Research Foundation for the Electrotechnology of Chubu.

REFERENCES