Whole-Genome Sequence of an Isogenic Haploid Strain, *Saccharomyces cerevisiae* IR-2idA30(MATa), Established from the Industrial Diploid Strain IR-2

Kazuhiro E. Fujimori, a Yosuke Kobayashi, a* Taisuke Seike, a* Takehiko Sahara, a Satoru Ohgiya, b Yoichi Kamagata a

a Bioproduction Research Institute (BPRI), National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan
b Bioproduction Research Institute (BPRI), National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, Hokkaido, Japan

**ABSTRACT** We present the draft genome sequence of an isogenic haploid strain, IR-2idA30(MATa), established from *Saccharomyces cerevisiae* IR-2. Assembly of long reads and previously obtained contigs from the genome of diploid IR-2 resulted in 50 contigs, and the variations and sequencing errors were corrected by short reads.

*Saccharomyces cerevisiae* IR-2 was first isolated in 1985 from fermented food in Indonesia, which showed unusual flocculation properties even under normal cultivation (1–4). In our previous report, we constructed a draft genome sequence of the original IR-2 diploid strain, which comprised ~300 contigs with a number of heterogeneous variations on both chromosomes (5). To overcome these problems, we used consanguineous inbreeding to establish isogenic strain pairs from the IR-2 diploid with improved efficiency in sporulation and spore germination.

IR-2idA30 (MATa and MATa) is a representative isogenic (identical except for the mating-type region) strain pair that we established from the original IR-2 diploid by repetitive syngensis (zygoses, sporulation, and haploid selection) as follows. To prevent homothallism, two HO genes were disrupted by *kanMX* (G418*) and *bleMX* (Zeocin®) using a traditional lithium acetate (LiAc) transformation method and homologous recombination (6, 7). After four rounds of syngensis, IR-2idA30(MATa Δho:kanMX), which exhibited loss of flocculation, was isolated. The antibiotic resistance gene in the HO locus was removed by expression of Cre recombinase via the pSH47bla plasmid (6), in which *URA3* was replaced with the blasticidin S resistance gene *bsd* from the pCAG-Bsd plasmid, resulting in IR-2idA30(MATa Δho::loxP). Finally, the mating-type region of IR-2idA30(MATa) was converted from MATa to MATa by HO expression via the pHO-bla plasmid, in which Cre was replaced with HO, resulting in IR-2idA30(MATa).

We constructed a draft genome sequence of IR-2idA30(MATa) as follows. The cells were grown in yeast extract-peptone-dextrose (YPD) medium (10 g/liter yeast extract, 20 g/liter Bacto peptone, and 20 g/liter D-glucose) at 30°C for 24 hours. Genomic DNA was isolated using the Dr. GenTLE (for yeast) high recovery system. Short reads from IR-2idA30(MATa) were generated on the Ion Personal Genome Machine (PGM) platform using the Ion Xpress Plus fragment library kit. The total number of reads was ~4.26 million with a mean read length of 378 bases. Approximately 133 million bases of long reads (mean length, 9,643 bases) from IR-2idA30(MATa) were generated using the PacBio RS II sequencer and the PacBio single-molecule real-time (SMRT) kit. Next, the long reads and contigs of the previously obtained IR-2 draft genome sequence (GenBank accession number BAUI00000000) were assembled to create longer contigs using PBJelly2 14.9.9, resulting in 50 contigs totaling 12,013,145 bases. Furthermore, by mapping the short reads using CLC Genomics Workbench 12.0 with the default conditions, 20,176 variations and short indels were...
corrected (mapping efficiency, 98.72%; coverage, 97.2×). Finally, we obtained 50 contigs of 12,013,960 bases and identified 7,458 potential open reading frames using the CLC Genomics Workbench. A dot plot created using D-GENIES for large genome alignment revealed that 95.61% of IR-2 idA30(MATa) bases are highly identical to those of S. cerevisiae S288c (48/50 contigs mapped to S288c chromosomes or mitochondrial sequences) \( ^{8} \). The sequence similarities between S288c and IR-2 idA30(MATa) are summarized in Table 1. Interestingly, a large deletion in the IR-2 idA30(MATa) genome was observed, corresponding to chromosome I of S288c (approximate positions, 2800 to 11750) \( ^{8} \).

Highly accurate genetic information from an isogenic strain with a homogeneous genetic background would provide a useful foundation for developing genetically and metabolically engineered strains for the production of valuable chemical compounds.

**Data availability.** The draft genome sequence of IR-2 idA30(MATa) has been deposited in DDBJ under the accession numbers BIMU01000001 to BIMU01000050, DRA accession numbers DRA007804 (Ion PGM) and DRA007805 (PacBio RS II), and BioProject number PRJDB7860.

**ACKNOWLEDGMENTS**

We thank N. Aoyagi, M. Hamajima, and M. Kabasawa for technical assistance, sample preparation, and processing.

This study was supported through project number P13011 by the New Energy and Industrial Technology Development Organization (NEDO).

Y. Kobayashi, T. Seike, T. Sahara, S. Ohgiya, and Y. Kamagata participated in project conception and coordination; K. E. Fujimori participated in the study design, carried out all the experiments and data analyses, and wrote the manuscript. All authors read and approved the final manuscript.

**REFERENCES**

1. Kuriyama H, Seiko Y, Murakami T, Kobayashi H, Sonoda Y. 1985. Continuous ethanol fermentation with cell recycling using flocculating yeast. J Ferment Technol 63:159–165.

2. Kobayashi Y, Sahara T, Suzuki T, Kamachi S, Matsushika A, Hoshino T, Ohgiya S, Kamagata Y, Fujimori KE. 2017. Genetic improvement of xylose metabolism by enhancing the expression of pentose phosphate pathway genes in Saccharomyces cerevisiae IR-2 for high-temperature ethanol production. J Ind Microbiol Biotechnol 44:879–891. https://doi.org/10.1007/s10295-017-1912-5.

3. Kobayashi Y, Sahara T, Ohgiya S, Kamagata Y, Fujimori KE. 2018. Systematic optimization of gene expression of pentose phosphate pathway genes in Saccharomyces cerevisiae IR-2 for high-temperature ethanol production. J Ind Microbiol Biotechnol 44:879–891. https://doi.org/10.1007/s10295-017-1912-5.

4. Inoue H, Hashimoto S, Matsushika A, Watanabe S, Sawayaama S. 2014. Breeding of a xylose-fermenting hybrid strain by mating genetically engineered haploid strains derived from industrial Saccharomyces cerevisiae. J Ind Microbiol Biotechnol 41:1773–1781. https://doi.org/10.1007/s10295-014-1531-3.

5. Sahara T, Fujimori KE, Nezuo M, Tsukahara M, Tochigi Y, Ohgiya S, Kamagata...
Y. 2014. Draft genome sequence of *Saccharomyces cerevisiae* IR-2, a useful industrial strain for highly efficient production of bioethanol. Genome Announc 2:e01160-13. https://doi.org/10.1128/genomeA.01160-13.

6. Güldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH. 1996. A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Res 24:2519–2524. https://doi.org/10.1093/nar/24.13.2519.

7. Gietz RD, Schiestl RH. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc 2:31–34. https://doi.org/10.1038/nprot.2007.13.

8. Cabanettes F, Klopp C. 2018. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ 6:e4958. https://doi.org/10.7717/peerj.4958.