Draft Genome Sequences of the Highly Halotolerant Strain Zygosaccharomyces rouxii ATCC 42981 and the Novel Allodiploid Strain Zygosaccharomyces sapae ABT301T Obtained Using the MinION Platform

Melissa Bizzarri, Stefano Cassanelli, Leszek P. Pryszcz, Jan Gawor, Robert Gromadka, Lisa Solieri

*Department of Life Sciences, University of Modena and Reggio Emilia, Reggio Emilia, Italy
bInternational Institute of Molecular and Cell Biology, Warsaw, Poland
cInstitute of Biochemistry and Biophysics PAS, Warsaw, Poland

ABSTRACT  Here, we report draft genome sequences of the halotolerant and allodiploid strains Zygosaccharomyces rouxii ATCC 42981 and Zygosaccharomyces sapae ABT301T. Illumina and Oxford Nanopore MinION sequencing revealed genome sizes of 20.9 and 24.7 Mb, respectively. This information will be useful for deciphering the genetics of hybrid adaptation to high salt and sugar concentrations in nonconventional yeasts.

The halotolerant yeasts of the genus Zygosaccharomyces find relevant applications in food spoilage and fermentation (1). They exhibit high diversity in response to high solute concentrations, tendency to hybridization, and ectopic recombination at the mating type loci, leading to ploidy and karyotype variation (2, 3). Zygosaccharomyces rouxii ATCC 42981 is an allodiploid strain isolated from Japanese miso, which grows at NaCl and dextrose concentrations up to 3.0 M and 70% (wt/vol), respectively (4). Zygosaccharomyces sapae represents a novel species, first described in high-sugar traditional balsamic vinegar (TBV), for which ABT301 (CBS 12607T = MUCL 54092T = UMCC 152T) is the type strain (5). ABT301T is a sugar-resistant and slow-growing strain more sensitive to salt than is ATCC 42981. Under standard conditions, ATCC 42981 produces more glycerol than does ABT301T and better retains it in the cell under conditions of salt stress (6). ATCC 42981 is thought to have arisen from hybridization between two divergent parents (3, 5, 7–9), while no evidence about the origin of strain ABT301T is available. Here, we present the draft genome sequences of ATCC 42981 and ABT301T.

Single-colony isolates were obtained from the Unimore Microbial Culture Collection (UMCC) of the University of Modena and Reggio Emilia in Italy. ABT301T was isolated from a TBV sample in May to June 2004 (10). DNA was extracted by using the phenol-chloroform-isooamyl alcohol method (11) after cell wall enzymatic lysis with 300 U lyticase (Sigma, St. Louis, MO) and subjected to short-read and long-read sequencing by using the MiSeq (Illumina) and MinION (ONT) platforms. Illumina libraries were prepared with an average insert size of ~600 bp and sequenced in paired-end mode on a MiSeq instrument using a v3 600-cycle chemistry kit. In total, 2,234,027 and 3,452,971 short paired-end reads were generated for ATCC 42981 and ABT301T, respectively.

MinION libraries were prepared from unsheared genomic DNA using a 1D ligation sequencing kit with modifications included in the One-Pot ligation manual (https://doi.org/10.17504/protocols.io.k9ac2ze). Genomes were sequenced separately on a MinION Mkib instrument using SQK-LSK108 chemistry and R9.4.1 flow cells. Total numbers of 260,559 and 197,963 long reads were generated for ATCC 42981 and ABT301T, respec-
tively. They were basecalled with Albacore v2.1.7, quality trimmed with PoreChop v0.2.1 (https://github.com/rrwick/Porechop) and error corrected with Canu v1.7 (12). Platanus v1.2.4 (13) was used to assemble the initial contigs, which were subsequently scaffolded with corrected MinION reads using DBG2OLC (14). Finally, scaffolds were polished with long reads using Racon v1.2.0 (15) and with short reads using Pilon v1.22 (16) and then reduced using Redundans v.014 (17). All software programs were used at default settings. Genes were annotated by similarity to the closest haploid relative, Z. rouxii CBS 732T (18), using Exonerate v2.2.0 (19). Assembly completeness was assessed by BUSCO v3.0.2 (20).

Comparison with haploid CBS 732T showed that the ATCC 42981 and ABT301T assembled genomes had a 2.14 and 2.53 times larger assembly size and contained a 2.11 and 2.46 times higher number of protein-coding genes, respectively (Table 1). For both genomes we dissected three haplotypes, and one of them was identical to that of CBS 732T (identity cutoff, 0.92). The data suggest a recursive hybridization model (21). The reported assemblies will decipher how hybridization, followed by functional genome stabilization, may offer a rapid adaptation strategy to salt stress environments in yeasts.

Data availability. The BioProject has been deposited in GenBank under number PRJEB26771. All sequencing reads of Z. rouxii ATCC 42981 and Z. sapae ABT301T have been deposited at EMBL/GenBank under the accession numbers UEMZ01000001 to UEMZ01000033 and UEGL01000001 to UEGL01000052, respectively.

ACKNOWLEDGMENTS

This work was partially supported by the Italian Ministry of Education, University and Research (MIUR), within the framework of the Italian National Grant for Fundamental Research (FFABR 2017). L.P.P. is funded by the National Science Centre (Poland) Polonez-1 framework from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement number 665778.

We acknowledge Paolo Giudici for his helpful advice and discussions and Luciana De Vero, Ph.D., for maintaining the UMCC strains.

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. Dakal TC, Solieri L, Giudici P. 2014. Adaptive response and tolerance to sugary and salt stress in the food yeast Zygosacharomyces rouxii. Int J Food Microbiol 185:140–157. https://doi.org/10.1016/j.ijfoodmicro.2014.05.015.
2. Watanabe J, Uehara K, Mogi Y. 2013. Diversity of mating-type chromosome structures in the yeast Zygosacharomyces rouxii caused by ectopic exchanges between MAT-like loci. PLoS One 8:e62121. https://doi.org/10.1371/journal.pone.0062121.
3. Bizzarri M, Giudici P, Cassanelli S, Solieri L. 2016. Chimeric sex-determining chromosomal regions and dysregulation of cell-type identity in a sterile Zygosacharomyces allodiploid yeast. PLoS One 11:e0152558. https://doi.org/10.1371/journal.pone.0152558.
4. Solieri L, Dakal TC, Biccianti S. 2014. Quantitative phenotypic analysis of multistress response in Zygosacharomyces rouxii complex. FEMS Yeast Res 14:586–600. https://doi.org/10.1111/1567-1364.12146.
5. Solieri L, Dakal TC, Giudici P. 2013. Zygosacharomyces sapae sp. nov., isolated from Italian traditional balsamic vinegar. Int J Syst Evol Microbiol 63:364–371. https://doi.org/10.1099/ijis.0.043323-0.

TABLE 1 Assembly metrics and annotation completeness obtained by using the BUSCO universal fungal genes (fungi_odb9) data set

| Feature                                  | Data for strain: |
|------------------------------------------|-----------------|
|                                          | CBS 732T | ABT301T | ATCC 42981 |
| Assembly size (bp)                       | 9,764,635 | 24,741,993 | 20,910,059 |
| No. of scaffolds                         | 7        | 52      | 33         |
| G+C content (%)                         | 39.13    | 39.57   | 39.65      |
| N90 contig size (bp)                     | 1,496,342 | 1,409,619 | 1,393,912 |
| N90 contig size (bp)                     | 1,114,666 | 146,869  | 400,395   |
| No. of gaps                              | 1,269    | 0       | 0         |
| Longest scaffold (bp)                    | 1,865,392 | 1,913,612 | 1,903,919 |
| No. of genes                            | 4,991    | 12,300  | 10,524    |
| No. of BUSCO complete genes (%)          | 285 (98.28) | 282 (97.24) | 285 (98.28) |
| No. of BUSCO duplicated genes (%)        | 0 (0)    | 240 (85.11) | 264 (92.63) |
6. Solieri L, Vezzani V, Cassanelli S, Dakal TC, Pazzini J, Giudici P. 2016. Differential hypersaline stress response in Zygosaccharomyces rouxii complex yeasts: a physiological and transcriptional study. FEMS Yeast Res 16:fow063. https://doi.org/10.1093/femsyr/fow063.

7. James SA, Bond CJ, Stratford M, Roberts IN. 2005. Molecular evidence for the existence of natural hybrids in the genus Zygosaccharomyces. FEMS Yeast Res 5:747–755. https://doi.org/10.1016/j.femsyr.2005.02.004.

8. Gordon JL, Wolfe KH. 2008. Recent allopolyploid origin of Zygosaccharomyces rouxii strain ATCC 42981. Yeast 25:449 – 456. https://doi.org/10.1002/yea.1598.

9. Watanabe J, Uehara K, Mogi Y, Tsukkioka Y. 2017. Mechanism for restoration of fertility in hybrid Zygosaccharomyces rouxii generated by interspecies hybridization. Appl Environ Microbiol 83:e01187-17. https://doi.org/10.1128/AEM.01187-17.

10. Solieri L, Landi S, De Vero L, Giudici P. 2006. Molecular assessment of indigenous yeast population from traditional balsamic vinegar. J Appl Microbiol 101:63–71. https://doi.org/10.1111/j.1365-2672.2006.02906.x.

11. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. 1985. Ty elements transpose through an RNA intermediate. Cell 40:491–500. https://doi.org/10.1016/0092-8674(85)90197-7.

12. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.

13. Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T. 2014. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res 24:1384 –1395. https://doi.org/10.1101/gr.170720.113.

14. Ye C, Hill CM, Wu S, Ruan J, Ma Z. 2016. DBG2OLC: Efficient assembly of large genomes using long erroneous reads of the third generation sequencing technologies. Sci Rep 6:31900. https://doi.org/10.1038/srep31900.

15. Vaser R, Sovic I, Nagarajan N, Škic M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.

16. Walker BJ, Abeel T, Shea T, Priest M, Abouelellai A, Sakthikumar S, Curom CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.

17. Przytycz L, Gabałdón T. 2016. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res 44:e113. https://doi.org/10.1093/nar/gkw294.

18. Souciet JL, Dujon B, Gaillardin C, Johnston M, Baret PV, Cliften P, Sherman DJ, Weisenbach J, Westhof E, Wincker P, Rubinin J, Barbe V, Segurens B, Artigueneave F, Anthouard V, Vacherie B, Val M-E, Fulton RS, Minx P, Wilson R, Durrens P, Jean G, March C, Martin T, Nikola M, Rolland T, Seret M-L, Casarègola S, Despons L, Fairhead C, Fischer G, Lafontaine I, Leh V, Lemaire M, de Montigny J, Neuveglin C, Thierry A, Blanc-Lenffe I, Bleykasten C, Diffels J, Fritsch E, Franjul L, Goëffon A, Jainaux N, Kachouri-Lafond R, Payen C, Potier S, Przybylova L, Ozanne C, Richard G-F, Sacerdot C, Straub M-L, Tall E. 2009. Comparative genomics of protoploid Saccharomyces cerevisiae. Genome Res 19:1696–1709. https://doi.org/10.1101/gr.091546.109.

19. Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31. https://doi.org/10.1186/1471-2105-6-31.

20. Simão 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.

21. Dakal TC, Giudici P, Solieri L. 2016. Contrasting patterns of rDNA homogenization within the Zygosaccharomyces rouxii species complex. PLoS One 11:e0160744. https://doi.org/10.1371/journal.pone.0160744.