



Genome Assembly of the Ty1-Less *Saccharomyces paradoxus* Strain DG1768

Jingxuan Chen,^a Holly McQueary,^{b*} David W. Hall,^{b§} Peter Philippsen,^c David J. Garfinkel,^{b,d} Casey M. Bergman^{a,b}

^aInstitute of Bioinformatics, University of Georgia, Athens, Georgia, USA

^bDepartment of Genetics, University of Georgia, Athens, Georgia, USA

^cMolecular Microbiology, Biozentrum, University of Basel, Basel, Switzerland

^dDepartment of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia, USA

ABSTRACT Here, we report an essentially complete genome assembly for the Ty1-less *Saccharomyces paradoxus* strain DG1768 (derivative of strain 337) based on PacBio and Illumina shotgun sequence data. We also document the genetic alterations that make this yeast strain a key resource for Ty1 mobility studies.

Saccharomyces yeasts have played essential roles in elucidating the biology of eukaryotic retrotransposons (1–4). Here, we report a genome assembly for *Saccharomyces paradoxus* strain DG1768, a representative of a “Ty1-less” lineage frequently used to study Ty1 retrotransposition (5–11). DG1768 [*MAT* α *ura3-100* *Ty1-155- Δ ura3(Scer)-101 his3- Δ 200hisG*] is derived from a wild *Saccharomyces* strain collected in California (UCDFST 51-186) that was obtained from the Ron Davis laboratory (renamed 155). Michael Ciriacy mated strain 155 with *Saccharomyces cerevisiae* GRF18 and then repeatedly backcrossed to strain 155 to create a nearly isogenic strain (called 276) carrying *ho* (*Scer*), one full-length Ty1 element (called Ty1-155), and a *ura3* mutation at the endogenous *URA3* locus. Homologous recombination was used to replace Ty1-155 in strain 276 with *S. cerevisiae* *URA3* plus ~600 bp of vector sequence, followed by selection of a *ura3* mutation in the new heterologous *URA3* locus to create the ρ - strain 337. Finally, *HIS3* was deleted (replaced with *Salmonella* *HisG*), and functional *S. cerevisiae* mitochondria were introduced via cytoduction to create strain DG1768 (6). Subsequently, DG1768 was shown to be *S. paradoxus* (8).

To prepare DNA for sequencing, single colonies of strain DG1768 were inoculated in 7 mL of yeast extract-peptone-dextrose (YPD) liquid broth and cultured for ~24 h at 30°C. Table 1 provides details on how PacBio and Illumina data sets were generated. For assembly, unfiltered PacBio reads were input to HGAP 4 (smrtlink-release_5.1.0.26412) (12), which performed read quality control, error correction, and adapter trimming. The PacBio-only assembly was polished five times with Pilon (v1.24) (13) using unfiltered paired-end 150-bp Illumina reads as input. Polished contigs were scaffolded with RagTag (v2.0.1) (14) using *S. paradoxus* YPS138 as the reference (15). Assembly statistics were calculated using QUASt (v5.0.2) (16) and BUSCO (v5.0.0, saccharomycetes_odb10 database) (17). Ty elements were annotated using a RepeatMasker-based pipeline (11). Telomere-associated sequences were annotated with LRSDAY (18). Default software parameters were used except where otherwise noted.

Our DG1768 assembly is 11,932,051 bp in length (scaffold N_{50} value of 917,222 bp), with an overall GC content of 38.43%. The assembly has 18 contigs in 17 scaffolds; 15 chromosomes and the mitochondrial DNA (mtDNA) are completely assembled in 1 contig each, and chromosome XII is assembled in 1 scaffold with 2 contigs, with a gap at the ribosomal DNA (rDNA) locus. A total of 97.5% of *Saccharomycetes* benchmarking universal single-copy orthologs (BUSCOs) are present and single copy. Core X elements

Editor Vincent Bruno, University of Maryland School of Medicine

Copyright © 2022 Chen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jingxuan Chen, jingxuanchen@uga.edu, or Casey M. Bergman, cbergman@uga.edu.

*Present address: Holly McQueary, Cornell University, Ithaca, New York, USA.

§Present address: David W. Hall, Exact Sciences, Redwood City, California, USA.

The authors declare no conflict of interest.

Received 2 September 2021

Accepted 30 December 2021

Published 20 January 2022

TABLE 1 Methods for and properties of the PacBio and Illumina whole-genome sequencing data sets used to assemble and polish the DG1768 genome

Parameter	PacBio reads	Illumina reads
DNA isolation method	Wizard genomic DNA purification kit (Promega)	YeaStar genomic DNA kit (Zymo Research)
Library prepn method	SMRTbell Express template preparation kit >15-kb size-selection protocol with Covaris g-TUBE shearing	Method of Urich et al. (19) with no bisulfite conversion
Sequencing chemistry	Sequel sequencing kit v2.1	S4 reagent kit (300 cycles)
Sequencing instrument	PacBio Sequel II	NovaSeq 6000
Avg read length (bp)	5,912	150
N_{50} (bp)	10,652	150
No. of reads	615,827	2,357,654

or Y' elements were annotated in 29 of 32 scaffold ends, indicating the near completeness of all chromosomes. We confirmed that no full-length copies of Ty1 are present in the DG1768 genome. The scar of the Ty1-155 knockout was localized to chromosome VI from nucleotide position 292,634 to 294,347. The scar of the *HIS3* knockout was localized to chromosome XV from nucleotide position 686961 to 688113. We also identified a segment of *S. cerevisiae* DNA, spanning chromosome IV from nucleotide position 1 to 135828, caused by introduction of the *ho* allele, and we confirmed that DG1768 mtDNA is from *S. cerevisiae*. The provenance and essentially complete genome assembly for DG1768 presented here should provide a useful resource for studies on Ty1 retrotransposon biology.

Data availability. PacBio data used to generate the assembly and Illumina data used to polish the assembly are available under BioProject accession number [PRJNA748953](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA748953). The assembly was deposited at NCBI under GenBank accession numbers [CP081969](https://www.ncbi.nlm.nih.gov/genbank/CP081969) to [CP081986](https://www.ncbi.nlm.nih.gov/genbank/CP081986).

ACKNOWLEDGMENTS

We dedicate this article to Michael Ciriacy, who performed all genetic steps essential for the Ty1-155 knockout experiment at the University of Düsseldorf and tragically died before seeing the fruition of his work. We thank Kyria Boundy-Mills (University of California, Davis) and Jasper Rine (University of California, Berkeley) for assistance determining the provenance of strain 155. We thank Marissa Howard, Noah Workman, and Magdy Alabady at the University of Georgia Genomics and Bioinformatics Core for assistance with PacBio library preparation and sequencing, Tyler Earp for assistance with Illumina library preparation, GeneWiz for Illumina sequencing, the Georgia Advanced Computing Resource Center for computational resources, and two members of the *Microbiology Resource Announcements* editorial board for comments on the manuscript.

This work was funded by the University of Georgia Research Foundation (C.M.B.) and NIH grant R01GM124216 (D.J.G. and C.M.B.).

REFERENCES

- Cameron JR, Loh EY, Davis RW. 1979. Evidence for transposition of dispersed repetitive DNA families in yeast. *Cell* 16:739–751. [https://doi.org/10.1016/0092-8674\(79\)90090-4](https://doi.org/10.1016/0092-8674(79)90090-4).
- Boeke JD. 1999. Retrotransposons of fungi, p 1503–1507. In Granoff A, Webster RG (ed), *Encyclopedia of virology*, 2nd ed. Academic Press, Cambridge, MA. <https://doi.org/10.1006/rwvi.1999.0249>.
- Curcio MJ, Lutz S, Lesage P. 2015. The Ty1 LTR-retrotransposon of budding yeast, *Saccharomyces cerevisiae*. *Microbiol Spectr* 3:MDNA3-0053-2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0053-2014>.
- Sandmeyer S, Patterson K, Bilanchone V. 2015. Ty3, a position-specific retrotransposon in budding yeast. *Microbiol Spectr* 3:MDNA3-0057-2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0057-2014>.
- Wilke CM, Adams J. 1992. Fitness effects of Ty transposition in *Saccharomyces cerevisiae*. *Genetics* 131:31–42. <https://doi.org/10.1093/genetics/131.1.31>.
- Garfinkel DJ, Nyswaner K, Wang J, Cho JY. 2003. Post-transcriptional cosuppression of Ty1 retrotransposition. *Genetics* 165:83–99. <https://doi.org/10.1093/genetics/165.1.83>.
- Blanc VM, Adams J. 2004. Ty1 insertions in intergenic regions of the genome of *Saccharomyces cerevisiae* transcribed by RNA polymerase III have no detectable selective effect. *FEMS Yeast Res* 4:487–491. [https://doi.org/10.1016/S1567-1356\(03\)00199-5](https://doi.org/10.1016/S1567-1356(03)00199-5).
- Moore SP, Liti G, Stefanisko KM, Nyswaner KM, Chang C, Louis EJ, Garfinkel DJ. 2004. Analysis of a Ty1-less variant of *Saccharomyces paradoxus*: the gain and loss of Ty1 elements. *Yeast* 21:649–660. <https://doi.org/10.1002/yea.1129>.
- Garfinkel DJ, Nyswaner KM, Stefanisko KM, Chang C, Moore SP. 2005. Ty1 copy number dynamics in *Saccharomyces*. *Genetics* 169:1845–1857. <https://doi.org/10.1534/genetics.104.037317>.
- Saha A, Mitchell JA, Nishida Y, Hildreth JE, Ariberre JA, Gilbert WV, Garfinkel DJ. 2015. A *trans*-dominant form of Gag restricts Ty1 retrotransposition and

- mediates copy number control. *J Virol* 89:3922–3938. <https://doi.org/10.1128/JVI.03060-14>.
11. Czaja W, Bensasson D, Ahn HW, Garfinkel DJ, Bergman CM. 2020. Evolution of Ty1 copy number control in yeast by horizontal transfer and recombination. *PLoS Genet* 16:e1008632. <https://doi.org/10.1371/journal.pgen.1008632>.
 12. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
 13. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo 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>.
 14. 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>.
 15. Yue JX, Li J, Aigrain L, Hallin J, Persson K, Oliver K, Bergstrom A, Coupland P, Warringer J, Lagomarsino MC, Fischer G, Durbin R, Liti G. 2017. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat Genet* 49:913–924. <https://doi.org/10.1038/ng.3847>.
 16. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
 17. 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>.
 18. Yue JX, Liti G. 2018. Long-read sequencing data analysis for yeasts. *Nat Protoc* 13:1213–1231. <https://doi.org/10.1038/nprot.2018.025>.
 19. Urich MA, Nery JR, Lister R, Schmitz RJ, Ecker JR. 2015. MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. *Nat Protoc* 10:475–483. <https://doi.org/10.1038/nprot.2014.114>.