





# Draft Genome Sequence of *Scheffersomyces spartinae* ARV011, a Marine Yeast Isolate

 Alex R. Villarreal,<sup>a</sup> Danielle E. Campbell,<sup>b,c</sup> Shanice S. Webster,<sup>d</sup>  Ryan C. Hunter<sup>a</sup>

<sup>a</sup>Department of Microbiology and Immunology, University of Minnesota, Minneapolis, Minnesota, USA

<sup>b</sup>Department of Medicine, Division of Infectious Diseases, Washington University School of Medicine, St. Louis, Missouri, USA

<sup>c</sup>Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri, USA

<sup>d</sup>Department of Microbiology and Immunology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA

**ABSTRACT** We report the draft genome sequence of *Scheffersomyces spartinae* ARV011, which was isolated from the Great Sippewissett Marsh in Falmouth, Massachusetts. Sequencing was performed using the Illumina NovaSeq 6000 platform, yielding 7,598,030 read pairs 250 bp in length. This resulted in a total draft genome size of 12,132,557 bp.

Previously assigned to the genus *Pichia* on the basis of phenotypic traits such as the ability to form ascospores, the species *spartinae* was assigned to its current genus *Scheffersomyces* in 2010 after large-subunit (LSU) rRNA (D1/D2) and small-subunit (SSU) rRNA gene sequencing provided support for its reassignment (1). Assignment of *Scheffersomyces spartinae* to its current genus remains controversial because of poor bootstrap support and, unlike other members of the genus, its inability to efficiently ferment D-xylose to ethanol (1, 2). However, *S. spartinae* is known to produce coenzyme Q9, fostering interest in the species for its potential biotechnological applications (1). The lack of sequenced *S. spartinae* genomes has limited progress in this area.

*S. spartinae* ARV011 was isolated from marine water and sediment collected from the Great Sippewissett Marsh in Falmouth, Massachusetts. The sample was composed of brackish water, sediment, plant matter, and macroscopic photosynthetic microbial aggregates known as “pink berries” (3). Microbial enrichment from this sediment was performed using yeast extract-peptone-dextrose (YPD) broth containing antibiotics (11 µg/ml carbenicillin, 25 µg/ml chloramphenicol, and 10 µg/ml tetracycline), with incubation at 30°C for 48 h prior to culture on YPD agar. An isolate of interest, based on the microscopic appearance of enlarged vacuoles containing fast moving volutin or “dancing bodies” (4) (Fig. 1), was selected for further characterization. Following overnight culture in YPD medium, genomic DNA was extracted using the Quick-DNA bacterial/fungal miniprep kit (Zymogen). Initial identification of the isolate as *S. spartinae* was performed by Sanger sequencing of the internal transcribed spacer (ITS) region using the primer set ITS1-30F/ITS1-217R (5) and analysis via BLASTn against the standard nonredundant nucleotide database (6). Both the forward and reverse sequences best matched the same *S. spartinae* ITS and ribosomal subunit sequence (GenBank accession number [KY105354.1](https://www.ncbi.nlm.nih.gov/nuclot/KY105354.1)), with the forward and reverse sequences exhibiting 98.94% and 99.20% identity, respectively. The same DNA extract was used to prepare a paired-end Illumina library using the TruSeq DNA PCR-free kit. Sequencing was performed using the Illumina NovaSeq 6000 platform, yielding 7,598,030 read pairs. Quality control was completed using KneadData v0.7.6 (<https://github.com/biobakery/kneaddata>), a wrapper for FastQC (7), Trimmomatic (8), and Bowtie2 (9), all with default settings. After quality control, 6,975,482 high-quality read pairs remained.

Adapter-free and quality-filtered sequences were assembled *de novo* using the SPAdes genome assembler v3.10.0 with default settings and no additional pipeline

**Citation** Villarreal AR, Campbell DE, Webster SS, Hunter RC. 2021. Draft genome sequence of *Scheffersomyces spartinae* ARV011, a marine yeast isolate. Microbiol Resour Announc 10:e00652-21. <https://doi.org/10.1128/MRA.00652-21>.

**Editor** Jason E. Stajich, University of California, Riverside

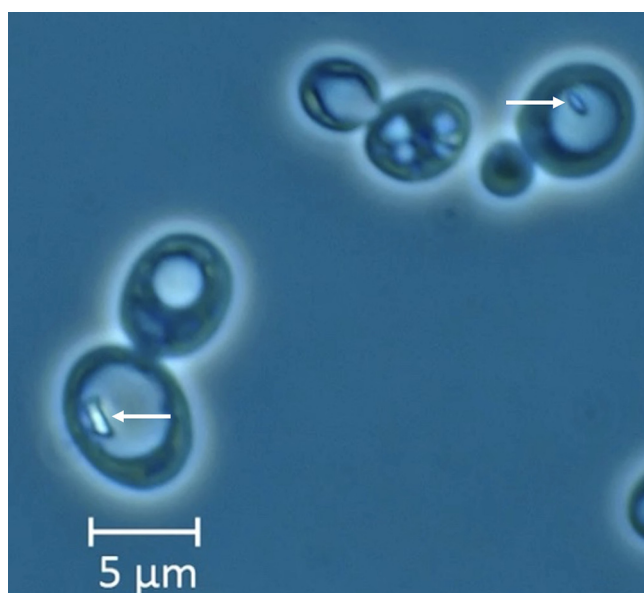
**Copyright** © 2021 Villarreal et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Alex R. Villarreal, [villa467@umn.edu](mailto:villa467@umn.edu), or Ryan C. Hunter, [rhunter@umn.edu](mailto:rhunter@umn.edu).

**Received** 6 July 2021

**Accepted** 19 October 2021

**Published** 11 November 2021



**FIG 1** Bright-field microscopy of isolate ARV011 at  $\times 1,000$  magnification. Yeast isolate cells ranged from 4 to 10  $\mu\text{m}$  in size, while the intracellular granules (white arrows) within enlarged vacuoles ranged from 1 to 2  $\mu\text{m}$  in size.

options (10). The resulting assembly graph was visualized via Bandage v0.8.1 (11) using default settings, which resulted in a total of 646 contigs (nodes), 858 connections between those contigs (edges), 19 dead ends, an  $N_{50}$  value of 375,049 bp, a GC content of 38.39%, and a total length of 12,201,803 bp, including overlaps. Annotation through NCBI estimated a draft genome size of 12,132,557 bp, with 103 contigs and 165 scaffolds. BLASTn was used via Bandage to reconfirm the isolate identification as *S. spartinae*, finding a 100% identity match to the published *S. spartinae* ATCC 18866 (12) partial 18S rRNA gene (GenBank accession number [HQ876044.1](#)) and a 100% match to the D1D2 region of the 26S rRNA gene (GenBank accession number [HQ876052.1](#)). Gene prediction and functional annotation were performed using default settings with Funannotate v1.8.1, a software package for gene prediction, annotation, and comparison of small eukaryotic genomes (13). Through Funannotate, Augustus, Augustus HiQ, GlimmerHMM, and snap identified a total of 14,401 genes. Also through Funannotate, Diamond mapped 550,947 proteins to the isolate genome, and 2,014 final alignments were made via Exonerate. Default databases were used for identification of genes and proteins. The full list of databases used and their version numbers are associated with the default settings of Funannotate v1.8.1. Functional annotation completeness was estimated to be above 89.6% using BUSCO v3 (14) run in protein mode and 91.3% run in genome mode with the Ascomycota lineage and *Saccharomyces cerevisiae* species as references. These data suggest that a nearly complete genome sequence was obtained for *Scheffersomyces spartinae* ARV011, which serves as a valuable resource for future studies of this organism.

**Data availability.** This whole-genome shotgun project for *Scheffersomyces spartinae* ARV011 has been deposited in DDBJ/ENA/GenBank under accession number [JAHMUF000000000](#). The version described in this paper is version [JAHMUF010000000](#). The draft genome assembly and annotation can be found under BioProject number [PRJNA738152](#) and BioSample number [SAMN19714210](#).

## ACKNOWLEDGMENTS

This work was supported by funds from the Marine Biological Laboratory, DOE (grant DE-SC0016127), NSF (grant MCB1822263), the Howard Hughes Medical Institute (grant 5600373), and a gift from the Simons Foundation.

We acknowledge the Woods Hole Microbial Diversity course directors (Rachel Whitaker and George O'Toole), course faculty Deborah Hogan, course staff, and lecturers for supporting this research. We acknowledge Alvaro Hernandez and Chris Wright at the Roy J. Carver Biotechnology Center at the University of Illinois in Urbana-Champaign for help with sequencing.

## REFERENCES

1. Kurtzman CP, Suzuki M. 2010. Phylogenetic analysis of ascomycete yeasts that form coenzyme Q-9 and the proposal of the new genera *Babjeviella*, *Meyerozyma*, *Millerozyma*, *Priceomyces*, and *Scheffersomyces*. *Mycoscience* 51:2–14. <https://doi.org/10.1007/S10267-009-0011-5>.
2. Jia R-R, Lv S-L, Chai C-Y, Hui F-L. 2020. Three new *Scheffersomyces* species associated with insects and rotting wood in China. *Mycologia* 71:87–99. <https://doi.org/10.3897/mycologia.71.56168>.
3. Wilbanks EG, Jaekel U, Salman V, Humphrey PT, Eisen JA, Facciotti MT, Buckley DH, Zinder SH, Druschel GK, Fike DA, Orphan VJ. 2014. Microscale sulfur cycling in the phototrophic pink berry consortia of the Sippewissett Salt Marsh. *Environ Microbiol* 16:3398–3415. <https://doi.org/10.1111/1462-2920.12388>.
4. Kharchuk MS, Glushenkov AN, Gromozova EN. 2019. Analysis of the motion of vacuolar volutin granules in *Saccharomyces cerevisiae*. *Folia Microbiol (Praha)* 64:207–213. <https://doi.org/10.1007/s12223-018-0646-8>.
5. Usyk M, Zolnik CP, Patel H, Levi MH, Burk RD. 2017. Novel ITS1 fungal primers for characterization of the mycobiome. *mSphere* 2:e00488-17. <https://doi.org/10.1128/mSphere.00488-17>.
6. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
7. Andrews S. 2016. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
8. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
9. Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
10. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Pribelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads, p 158–170. *In* Deng M, Jiang R, Sun F, Zhang X (ed), *Research in computational molecular biology*. Springer, Berlin, Germany. [https://doi.org/10.1007/978-3-642-37195-0\\_13](https://doi.org/10.1007/978-3-642-37195-0_13).
11. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>.
12. American Type Culture Collection. *Scheffersomyces spartinae* (Ahearn et al) Kurtzman et Suzuki:ATCC18866TM. <https://www.atcc.org/products/18866>.
13. Palmer JM, Stajich J. 2020. Funannotate v1.8.1: eukaryotic genome annotation. Zenodo. <https://zenodo.org/record/4054262#.YXZmHhrMLIU>.
14. Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol* 1962:227–245. [https://doi.org/10.1007/978-1-4939-9173-0\\_14](https://doi.org/10.1007/978-1-4939-9173-0_14).