Draft Genome Sequence of the Yeast *Starmerella bacillaris* (syn., *Candida zemplinina*) FRI751 Isolated from Fermenting Must of Dried Raboso Grapes

Wilson José Fernandes Lemos Junior,[a] Laura Treu,[a,d] Vinícius da Silva Duarte,[b] Stefano Campanaro,[c] Chiara Nadai,[e] Alessio Giacomini,[a,e] Viviana Corich[a,e]

Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Legnaro, Italy[a]; Department of Microbiology, Universidade Federal de Viçosa, Viçosa, Brazil[b]; Department of Biology, University of Padova, Padua, Italy[c]; Department of Environmental Engineering, Technical University of Denmark, Kongens Lyngby, Denmark[d]; Interdepartmental Centre for Research in Viticulture and Enology, University of Padova, Conegliano, Italy[e]

ABSTRACT *Starmerella bacillaris* is an ascomycetous yeast commonly present in enological environments. Here, we report the first draft genome sequence of *S. bacillaris* FRI751, which will facilitate the study of the characteristics of this interesting enological yeast.

The ascomycetous yeast *Starmerella bacillaris* (syn., *Candida zemplinina*) is frequently found in spontaneous must fermentation, usually at a relatively high population level of 10^4 to 10^6 cells/ml (1), in grape marcs (2), and it is also normally present on botrytized grapes.

This species was isolated for the first time in Napa Valley (CA) in 2002 (3), and 1 year later, Sipiczki (4) assigned this *Candida* sp. to a novel species under the name *Candida zemplinina*, due to the significant differences observed in the rRNA sequence from that of the related species *Candida stellata* (5). For a long time, *C. zemplinina* has been confounded with its close species *C. stellata*, which shares similar ecological niches, particularly in grape and wine environments. Finally, it was established as *Starmerella bacillaris* (6).

S. bacillaris is able to ferment glucose, sucrose, and raffinose but not galactose, maltose, or lactose (6). Unable to grow in vitamin-free medium, it develops well in the presence of high glucose concentration, up to 50% (wt/vol) (6). It is highly fructophilic and a high-glycerol producer (7).

S. bacillaris is a psichrotolerant and osmotolerant species (4), and among the non-*Saccharomyces* yeasts of enological interest, *S. bacillaris* is considered one of the most promising species to satisfy modern market and consumer preferences. In particular, it produces less ethanol from must fermentation than *Saccharomyces cerevisiae*, low levels of biogenic amines, and average volatile acidity (8). It is also being tested in association with *Saccharomyces cerevisiae* in mixed or sequential fermentations to reduce alcohol content and to increase the organoleptic properties of wines (7), and its possible use in the vineyard as an antifungal agent against *Botrytis* is under study (8).

In this work, the first genome sequence for an *S. bacillaris* strain is released. Strain FRI751 was isolated from fermentation of dried grapes of Raboso wine, a vine variety cultivated mainly in the Northeast of Italy for the production of passito wines.

S. bacillaris FRI751 genomic DNA was prepared by zymolyase digestion, followed by standard phenol-chloroform extraction, as described by Vaughan-Martini and Martini

(9). The genome sequence was generated using an Illumina NextSeq 500 platform (1-kb mate-pair libraries) at the Ramaciotti Centre, Sydney, Australia. The sequencing generated 45-fold coverage with 1,435,554 paired-end (2 $\times$ 150 bp) and 102,368 unpaired reads (after quality filtering) that were used for the *de novo* assembly by SPAdes 3.10 software (10) (with option -k 21,33,55,77,99,127). The genome size of *S. bacillaris* FRI751 was 9.3 Mbp, divided into 106 contigs longer than 100 bp (N_{50} length, 208,744 bp), and the G+C content was 39.4%. Protein-coding gene (CDS) prediction was performed using GeneMark-ES (11) and resulted in 4,028 CDSs and a total of 4,315 exons. Gene annotation was obtained combining two strategies: (i) BlastKOALA (12) was used to search against a nonredundant set of KEGG genes, selecting Saccharomycetaceae as the taxonomy group; and (ii) RPS BLAST was used to compare protein sequences with Eukaryotic Orthologous Groups of proteins (KOG) (13).

The data reported here represent a useful resource to increase the knowledge of *S. bacillaris* metabolism and of its potential technological characteristics as applied to enology.

Accession number(s). The whole-genome shotgun project of *S. bacillaris* FRI751 has been deposited in DDBJ/ENA/GenBank under the accession no. MWSF00000000. The version described in this paper is the first version, MWSF01000000.

REFERENCES

1. Masneuf-Pomarede I, Juquin E, Miot-Sertier C, Renault P, Laizet Y, Salin F, Alexandre H, Capozzi V, Cocolin L, Colonna-Ceccaldi B, Englezos V, Girard P, Gonzalez B, Lucas P, Mas A, Nisiotou A, Sipiczki M, Spano G, Tassou C, Bely M, Albertin W. 2015. The yeast *Starmerella bacillaris* (synonym *Candida zemplinina*) shows high genetic diversity in winemaking environments. FEMS Yeast Res 15:fov045. https://doi.org/10.1093/femsyr/fov045.

2. Bovo B, Giacomini A, Corich V. 2011. Effects of grape marcs acidification treatment on the evolution of indigenous yeast populations during the production of grappa. J Appl Microbiol 111:382–388. https://doi.org/10.1111/j.1365-2672.2011.05060.x.

3. Mills DA, Johannsen EA, Cocolin L. 2002. Yeast diversity and persistence in *Botrytis*-affected wine fermentations. Appl Environ Microbiol 68:4884–4893. https://doi.org/10.1128/AEM.68.10.4884-4893.2002.

4. Sipiczki M. 2003. *Candida zemplinina* sp. nov., an osmotolerant and psychrotolerant yeast that ferments sweet botrytized wines. Int J Syst Evol Microbiol 53:2079–2083. https://doi.org/10.1099/ijs.0.02649-0.

5. Sipiczki M. 2004. Species identification and comparative molecular and physiological analysis of *Candida zemplinina* and *Candida stellata*. J Basic Microbiol 44:471–479. https://doi.org/10.1002/jobm.200410449.

6. Duarte FL, Pimentel NH, Teixeira A, Fonseca A. 2012. *Saccharomyces bacillaris* is not a synonym of *Candida stellata*: reinstatement as *Starmerella bacillaris* comb. nov. Antonie Van Leeuwenhoek 102:653–658. https://doi.org/10.1007/s10482-012-9762-7.

7. Englezos V, Rantsiou K, Cravero F, Torchio F, Ortiz-Julien A, Gerbi V, Rolle L, Cocolin L. 2016. *Starmerella bacillaris* and *Saccharomyces cerevisiae* mixed fermentations to reduce ethanol content in wine. Appl Microbiol Biotechnol 100:5515–5526. https://doi.org/10.1007/s00253-016-7413-z.

8. Lemos WJ, Bovo B, Nadai C, Crosato G, Carlot M, Favaron F, Giacomini A, Corich V. 2016. Biocontrol ability and action mechanism of *Starmerella bacillaris* (synonym *Candida zemplinina*) isolated from wine musts against gray mold disease agent *Botrytis cinerea* on grape and their effects on alcoholic fermentation. Front Microbiol 7:1249. https://doi.org/10.3389/fmicb.2016.01249.

9. Vaughan-Martini A, Martini A. 1996. Isolation, purification, and analysis of nuclear DNA in yeast taxonomy. Methods Mol Biol 53:89–102. https://doi.org/10.1385/0-89603-319-8:89.

10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

11. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an *ab initio* algorithm with unsupervised training. Genome Res 18:1979–1990. https://doi.org/10.1101/gr.081612.108.

12. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006.

13. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4:41. https://doi.org/10.1186/1471-2105-4-41.