

Genomic Sequence of the Yeast *Kluyveromyces marxianus* CCT 7735 (UFV-3), a Highly Lactose-Fermenting Yeast Isolated from the Brazilian Dairy Industry

Wendel B. Silveira,^a Raphael H. S. Diniz,^a M. Esperanza Cerdán,^b María I. González-Siso,^b Robson de A Souza,^a Pedro M. P. Vidigal,^c Otávio J. B. Brustolini,^a Emille R. B. de Almeida Prata,^d Alexandra C. Medeiros,^a Lílian C. Paiva,^a Moysés Nascimento,^e Éder G. Ferreira,^a Valdilene C. dos Santos,^a Caio R. S. Bragança,^a Tatiana A. R. Fernandes,^a Lívia T. Colombo,^a Flávia M. L. Passos^a

Departamento de Microbiologia,^a Núcleo de Análise de Biomoléculas (NuBioMol),^c and Departamento de Estatística,^e Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO), Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil; Grupo EXPRELA, Departamento de Biología Celular e Molecular, Universidade da Coruña, Campus da Zapateira, A Coruña, Spain^b; Instituto de Ciência, Engenharia e Tecnologia (ICET), Universidade Federal do Vale do Jequitinhonha e Mucuri (UFVJM), Brazil^d

Here, we present the draft genome sequence of *Kluyveromyces marxianus* CCT 7735 (UFV-3), including the eight chromosomes and the mitochondrial genomic sequences.

Received 24 September 2014 Accepted 28 September 2014 Published 6 November 2014

Citation Silveira WB, Diniz RHS, Cerdán ME, González-Siso MI, Souza RDA, Vidigal PMP, Brustolini OJB, de Almeida Prata ERB, Medeiros AC, Paiva LC, Nascimento M, Ferreira ÉG, dos Santos VC, Bragança CRS, Fernandes TAR, Colombo LT, Passos FML. 2014. Genomic sequence of the yeast *Kluyveromyces marxianus* CCT 7735 (UFV-3), a highly lactose-fermenting yeast isolated from the Brazilian dairy industry. *Genome Announc.* 2(6):e01136-14. doi:10.1128/genomeA.01136-14.

Copyright © 2014 Silveira et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Flávia M. L. Passos, flpassos@ufv.br.

Here, we present the draft genome of *Kluyveromyces marxianus* CCT 7735 (UFV-3), including the eight chromosomes and the mitochondrial genomic sequences. *Kluyveromyces marxianus* strains show both high metabolic diversity and a substantial degree of polymorphisms. This is likely to be explained by the different habitats from which it has been recovered (from plant sources and natural fermentations to ecological niches associated with warm-blooded animals, including dairy products). *K. marxianus* CCT 7735 (UFV-3), isolated from the regional Brazilian dairy industry, was selected from among other yeast isolates due to its high β -galactosidase activity, followed by a high flux of lactose assimilation and ethanol production (1). The high fermentative flux in *K. marxianus* UFV-3 seems to be favored by the high expression of key genes that encode lactose metabolism enzymes (2). It should be pointed out that opposed to other isolates and related taxa, including *Kluyveromyces lactis*, this strain is able to grow under anaerobic conditions (2). Therefore, the study of the Kluyver effect in *K. marxianus* UFV-3, i.e., the inability to transport certain sugars in the absence of oxygen, may provide new insights into carbon assimilation in yeasts of the *Kluyveromyces* genus.

On the other hand, high cell mass yields are attained when this yeast is grown in low-sugar concentrations, showing its potential as a cell factory. Besides β -galactosidase, *K. marxianus* UFV-3 exhibits high endopolygalacturonase activity. At a low growth rate, this enzyme is strongly secreted (our unpublished data).

Like the other *K. marxianus* strains, *K. marxianus* UFV-3 is capable of assimilating a variety of substrates and growing at elevated temperatures, which are desirable traits for use in biofuel production from inexpensive sources. Considering the little knowledge that exists about the genetics and physiology of *K. marxianus* UFV-3, as well as its potential to be used as a cell factory in the production of biofuel, chemicals, and recombinant

proteins, we sequenced its genome in order to gain insights about its biochemical and genetic traits.

Three libraries with insert sizes of 500 nucleotides (nt), 2 kb, and 5 kb were constructed, amplified, and sequenced in Illumina HiSeq 2000 by Sistemas Genómicos (Valencia, Spain), producing sequence data containing 51 million paired reads with 100 nt. To assemble the chromosomes and mitochondrial genomic sequences of *K. marxianus* UFV-3, the sequences of *K. marxianus* DMKU3-1042 (Genbank accession no. AP012213 to AP012221) were selected as references. The assembly was performed using a variant calling approach. The reads were mapped in the references using CLC Genomics Workbench version 6.5.1 (CLC bio), and the contigs were reconstructed using Bcfutils of SAMtools version 1.4 (3). This assembly produced nine contigs with overall G+C contents of 39% for chromosomes and 13.5% for the mitochondrial sequences and mean coverages of 385-fold for the chromosomes and 841-fold for the mitochondrial sequences. The contigs were analyzed using Augustus version 2.5.5 (4, 5) to predict the genes with *K. lactis* trained data. The proteins encoded by the 4,787 predicted genes were functionally annotated using BLAST searches (<http://blast.ncbi.nlm.nih.gov/>) and HMMER (6). Most of the predicted proteins (~97%) presented high similarity to *K. marxianus* DMKU3-1042 proteins, and some (~3%) were similar to *K. lactis* NRRL Y-1140.

Nucleotide sequence accession numbers. The complete genomic sequences of the *K. marxianus* CCT 7735 (UFV-3) chromosomes and mitochondrial genome have been deposited in GenBank under the accession numbers CP009303 to CP009311.

ACKNOWLEDGMENTS

The Brazilian agencies CAPES, CNPq, and FAPEMIG supported this work. General support to the Grupo EXPRELA during 2012 to 2014 was funded by Xunta de Galicia (contract no. 2012/118) and by FEDER.

We thank the Department of Information Technology (Federal University of Viçosa) for supporting the data analysis on the computational cluster.

REFERENCES

1. Silveira WB, Passos FJV, Mantovani HC, Passos FML. 2005. Ethanol production from cheese whey permeate by *Kluyveromyces marxianus* UFV-3: a flux analysis of oxido-reductive metabolism as a function of lactose concentration and oxygen levels. *Enzyme Microb. Technol.* **36**:930–936. <http://dx.doi.org/10.1016/j.enzmictec.2005.01.018>.
2. Diniz RHS, Silveira WB, Fietto LG, Passos FML. 2012. The high fermentative metabolism of *Kluyveromyces marxianus* UFV-3 relies on the increased expression of key lactose metabolic enzymes. *Antonie Van Leeuwenhoek* **101**:541–550. <http://dx.doi.org/10.1007/s10482-011-9668-9>.
3. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPPD. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**:2078–2079. <http://dx.doi.org/10.1093/bioinformatics/btp352>.
4. Stanke M, Schöffmann O, Morgenstern B, Waack S. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**:62. <http://dx.doi.org/10.1186/1471-2105-7-62>.
5. Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**(Suppl 2):ii215–ii225. <http://dx.doi.org/10.1093/bioinformatics/btg1080>.
6. Finn RD, Clements J, Eddy SR. 2011. HMMER Web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**:W29–W37. <http://dx.doi.org/10.1093/nar/gkr367>.