



Complete Genome Sequence of *Kluyveromyces lactis* Strain GG799, a Common Yeast Host for Heterologous Protein Expression

Léa Chuzel, Mehul B. Ganatra, Kelly M. Schermerhorn, Andrew F. Gardner, Brian P. Anton, Christopher H. Taron

New England Biolabs, Ipswich, Massachusetts, USA

ABSTRACT We report the genome sequence of the dairy yeast *Kluyveromyces lactis* strain GG799 obtained using the Pacific Biosciences RS II platform. *K. lactis* strain GG799 is a common host for the expression of proteins at both laboratory and industrial scales.

Kluyveromyces lactis is an ascomycetous yeast from the family *Saccharomycetaceae* that has a long history in research and biotechnology. *K. lactis* produces lactase, an enzyme that breaks down milk sugar (lactose). It is a major source of lactase for dairy industry processes. Additionally, *K. lactis* is a popular host for protein expression. Over 100 proteins have been expressed in *K. lactis* (1, 2), including chymosin that is produced at industrial scale (3). Here we report the complete genome sequence of *K. lactis* strain GG799, a common host strain for protein expression.

A 10-kb SMRTbell library was prepared from genomic DNA extracted from *K. lactis* GG799 cells with a blood and cell culture midikit (Qiagen), size selected with a BluePippin (Sage Science) for fragments over 8 kb, and sequenced with an RS II using P6 chemistry, 6 single-molecule real-time (SMRT) cells, and 360-min movies (Pacific Biosciences). Reads from 3 SMRT cells were assembled *de novo* using the RS_HGAP_Assembly.2 protocol, generating 70 contigs with 137× mean coverage.

The haploid genome of GG799 comprises six chromosomes (A through F). Five of six chromosomes (A, B, C, E, and F) were complete, and each was assembled into a single contig. Chromosome D, although complete, was split into two contigs at the ribosomal DNA (rDNA) tandem repeat region. Our reconstruction of chromosome D included two 8.7-kb rDNA repeat copies, but we estimated from sequence coverage that ~35 repeats exist. Of the additional 63 *de novo* contigs not described above, 28 were from mitochondrial DNA (mtDNA) and the remainder were redundant (5 from subtelomeric chromosomal repeat regions and 30 from the rDNA repeat region of chromosome D). The six complete chromosomes and mtDNA were error corrected using data from all six SMRT cells, with mean coverages of ~300× and ~1,900×, respectively. The total length of the GG799 genome was 10.7 Mb, with a GC content of 38.7% (including mtDNA).

The assembled contigs were mapped to the previously sequenced *K. lactis* strain CLIB210 (4, 5). The *K. lactis* strains GG799 and CLIB210 have highly similar genome organization. The only major difference is at the mating-type locus (MAT) on chromosome C. Strain GG799 harbors an MAT α allele 250 kb into the 5' region of chromosome C, whereas CLIB210 contains an MAT α allele at the same position. This difference reflects the opposite mating phenotypes of both strains (6). Interestingly, GG799 lacks an extra copy of HML α at the 5' end of chromosome C that is present in CLIB210.

Received 15 May 2017 Accepted 20 May 2017 Published 27 July 2017

Citation Chuzel L, Ganatra MB, Schermerhorn KM, Gardner AF, Anton BP, Taron CH. 2017. Complete genome sequence of *Kluyveromyces lactis* strain GG799, a common yeast host for heterologous protein expression. *Genome Announc* 5:e00623-17. <https://doi.org/10.1128/genomeA.00623-17>.

Copyright © 2017 Chuzel 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 Christopher H. Taron, taron@neb.com.

L.C. and M.B.G. contributed equally to this work.

Both strains also have subtelomeric regions comprising many duplicated sequences, but their organization is not identical, a finding that is consistent with the known plasticity of these fast-evolving regions (7).

This study provides a second *K. lactis* genome sequence that will permit further comparative insights into the evolution of *Kluyveromyces* yeasts. Additionally, this work will facilitate more informed engineering of *K. lactis* GG799 for improved heterologous protein expression.

Accession number(s). The complete genome sequence has been deposited in GenBank under the accession no. [CP021239](#) to [CP021245](#). The version described in this paper is the first version.

ACKNOWLEDGMENTS

We thank Donald Comb of New England Biolabs for funding and a supportive research environment. We thank Rick Morgan, Barbara Taron, and Saulius Vainauskas for valuable comments and suggestions.

REFERENCES

1. van Ooyen AJJ, Dekker P, Huang M, Olsthoorn MMA, Jacobs DI, Colussi PA, Taron CH. 2006. Heterologous protein production in the yeast *Kluyveromyces lactis*. *FEMS Yeast Res* 6:381–392. <https://doi.org/10.1111/j.1567-1364.2006.00049.x>.
2. Spohner SC, Schaum V, Quitmann H, Czermak P. 2016. *Kluyveromyces lactis*: an emerging tool in biotechnology. *J Biotechnol* 222:104–116. <https://doi.org/10.1016/j.jbiotec.2016.02.023>.
3. van den Berg JA, van der Laken KJ, van Ooyen AJ, Renniers TC, Rietveld K, Schaap A, Brake AJ, Bishop RJ, Schultz K, Moyer D, Richman M, Shuster JR. 1990. *Kluyveromyces* as a host for heterologous gene expression: expression and secretion of prochymosin. *Biotechnology* 8:135–139. <https://doi.org/10.1038/nbt0290-135>.
4. Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, de Montigny J, Marck C, Neuvéglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisramé A, Boyer J, Cattolico L, Confanioleri F, de Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, Groppi A, Hantraye F, Hennequin C, Jauniaux N, Joyet P, Kachouri R, Kerrest A, Koszul R, Lemaire M, Lesur I, Ma L, Muller H, Nicaud J, Nikolski M, Oztas S, Ozier-Kalogeropoulos O, Pellenz S, Potier S, Richard G, Straub M, Suleau A, Swennen D, Tekaia F, Wésolowski-Louvel M, Westhof E, Wirth B, Zeniou-Meyer M, Zivanovic I, Bolotin-Fukuhara M, Thierry A, Bouchier C, Caudron B, Scarpelli C, Gaillardin C, Weissenbach J, Wincker P, Souciet J. 2004. Genome evolution in yeasts. *Nature* 430:35–44. <https://doi.org/10.1038/nature02579>.
5. Zivanovic Y, Wincker P, Vacherie B, Bolotin-Fukuhara M, Fukuhara H. 2005. Complete nucleotide sequence of the mitochondrial DNA from *Kluyveromyces lactis*. *FEMS Yeast Res* 5:315–322. <https://doi.org/10.1016/j.femsyr.2004.09.003>.
6. Colussi PA, Specht CA, Taron CH. 2005. Characterization of a nucleus-encoded chitinase from the yeast *Kluyveromyces lactis*. *Appl Environ Microbiol* 71:2862–2869. <https://doi.org/10.1128/AEM.71.6.2862-2869.2005>.
7. Fairhead C, Dujon B. 2006. Structure of *Kluyveromyces lactis* subtelomeres: duplications and gene content. *FEMS Yeast Res* 6:428–441. <https://doi.org/10.1111/j.1567-1364.2006.00033.x>.