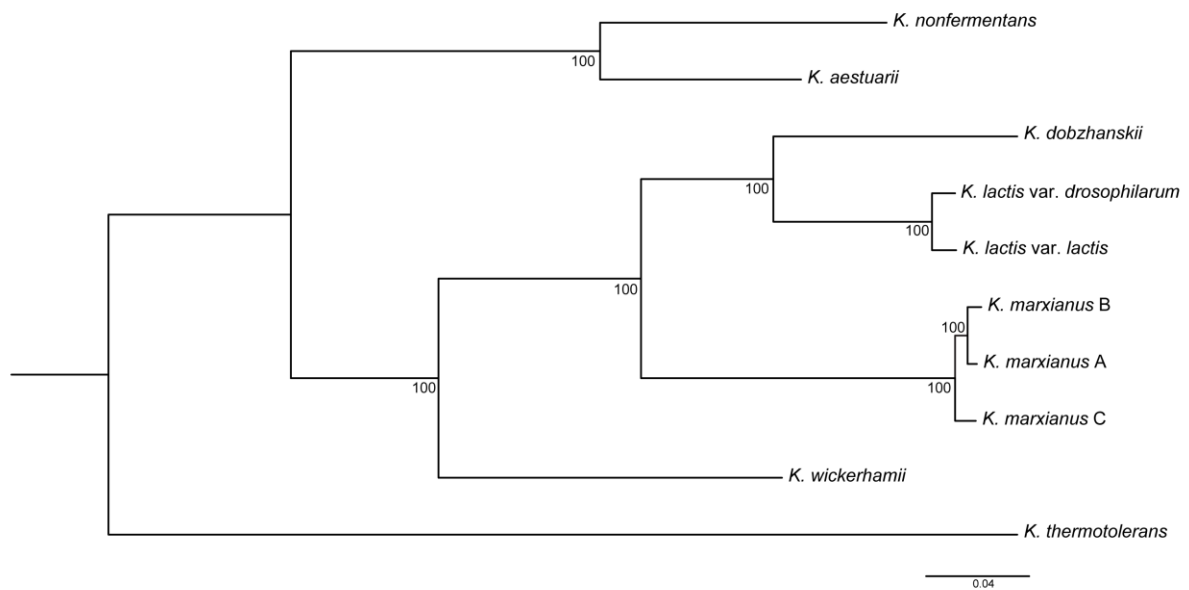


Current Biology, Volume 29

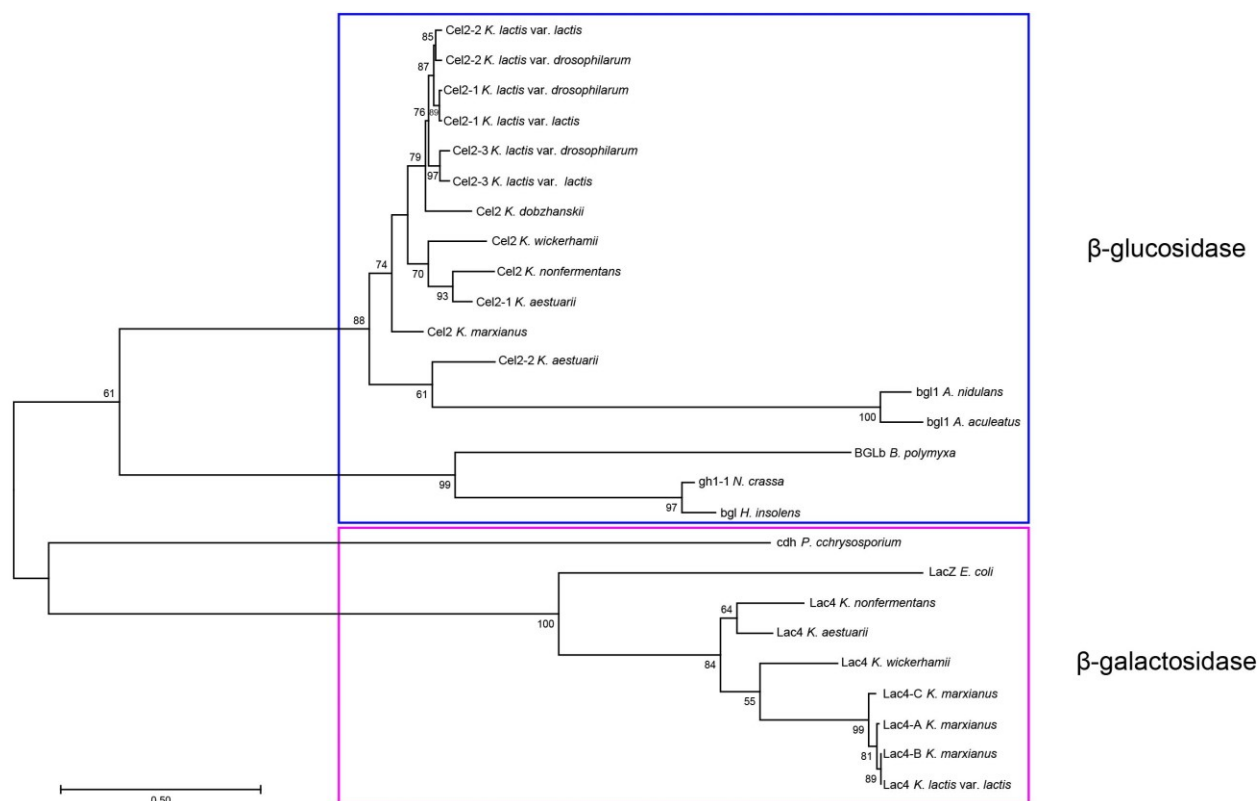
## Supplemental Information

### Origin of Lactose Fermentation in *Kluyveromyces lactis* by Interspecies Transfer of a Neo- functionalized Gene Cluster during Domestication

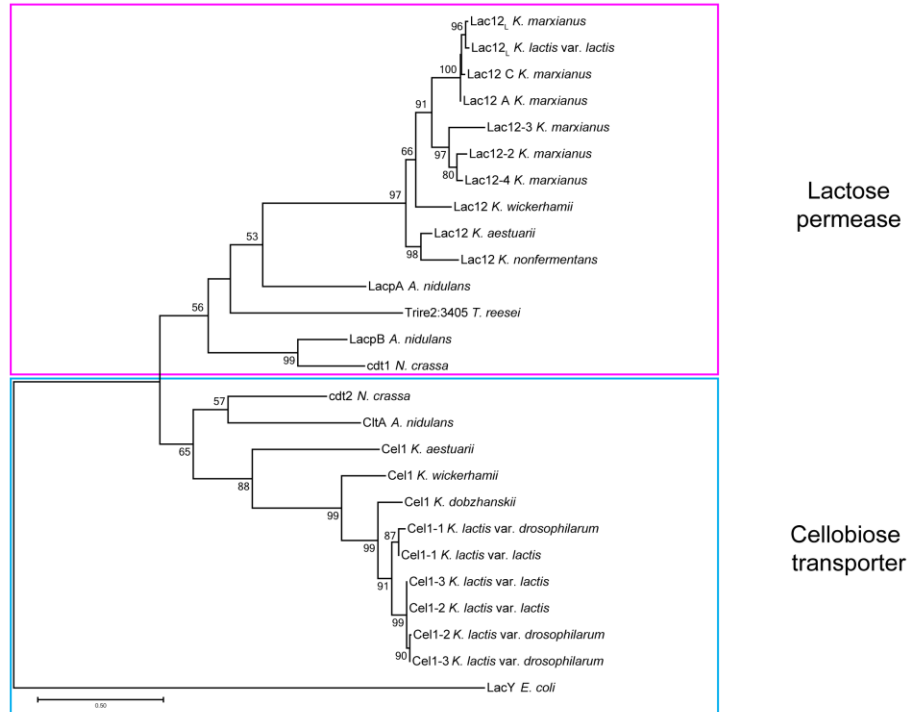
Javier A. Varela, Martina Puricelli, Raúl A. Ortiz-Merino, Romina Giacomobono, Stephanie Braun-Galleani, Kenneth H. Wolfe, and John P. Morrissey



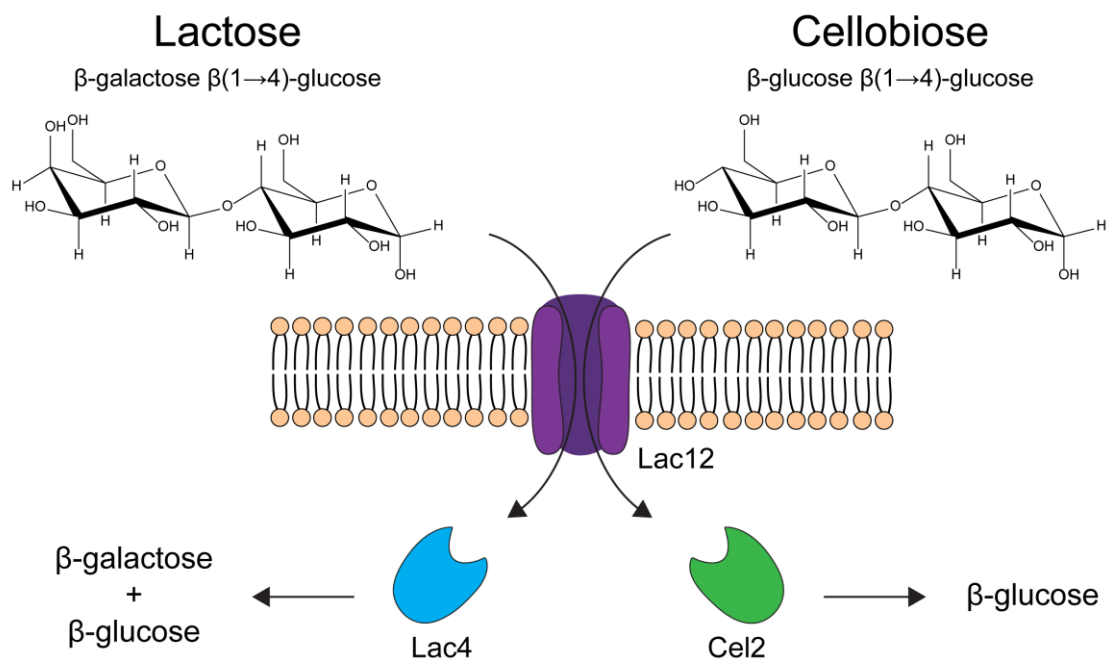
**Figure S1. Phylogeny of *Kluveromyces* species. Related to Figure 1.** The tree was constructed using 1,515 single-copy orthologous amino acid sequences from *Kluveromyces* genome assemblies (see STAR Methods). The values displayed on the branches correspond to bootstrapping values for the tree. The tree was computed with 100 bootstrap replicates.



**Figure S2. Phylogenetic tree of  $\beta$ -glucosidases and  $\beta$ -galactosidases. Related to Figure 3.** Cel2 proteins from *Kluyveromyces* species cluster with  $\beta$ -glucosidases from other fungi, whereas Lac4 proteins form a clade with the *E. coli* LacZ  $\beta$ -galactosidase. The maximum likelihood tree was computed using MEGA 7 [S1] with 500 bootstrap replicates. Branches with >50% bootstrap support are labeled.



**Figure S3. Phylogenetic tree of cellobiose and lactose transporters. Related to Figure 3.** The tree was drawn following a multiple sequence alignment of Lac12 and Cel1 proteins and their homologs in other species. Cel1 clusters with known fungal cellobiose transport genes, whereas Lac12 is in a clade with known fungal lactose permeases. The *E. coli* LacY sequence was used as outgroup. The maximum likelihood tree was computed using MEGA 7 [S1] with 500 bootstrap replicates. Branches with >50% bootstrap support are labeled.



**Figure S4. Scheme showing the bifunctional behavior of Lac12 from *K. marxianus*. Related to Figure 4.** The absence of Cel1, a dedicated cellobiose transporter in *K. marxianus*, is compensated by Lac12 that is able to transport either lactose or cellobiose. After being transported into the cell via Lac12, lactose and cellobiose are hydrolyzed into monomers by Lac4 (lactase) and Cel2 (cellobiase), respectively.

Species	Source
<i>K. aestuarii</i> CBS4438 <sup>†</sup>	Westerdijk Fungal Biodiversity Institute
<i>K. dozhanskii</i> CBS2104 <sup>†</sup>	Westerdijk Fungal Biodiversity Institute
<i>K. lactis</i> var. <i>lactis</i> CBS2359 <sup>†</sup>	Westerdijk Fungal Biodiversity Institute
<i>K. lactis</i> var. <i>drosophilarum</i> CBS2105 <sup>†</sup>	Westerdijk Fungal Biodiversity Institute
<i>K. marxianus</i> CBS397	Westerdijk Fungal Biodiversity Institute
<i>K. marxianus</i> NBRC1777	NITE Biological Research Centre
<i>K. marxianus</i> NBRC1777 <i>cel2</i>	This study
<i>K. marxianus</i> NBRC1777 <i>lac4</i>	This study
<i>K. marxianus</i> NBRC1777 <i>cel2 lac4</i>	This study
<i>K. marxianus</i> NBRC1777 <i>lac12 lac12-2 lac12-4</i>	[S2]
<i>K. wickerhamii</i> CBS2745 <sup>†</sup>	Westerdijk Fungal Biodiversity Institute
<i>S. cerevisiae</i> EBY.VW4000	Goethe University Frankfurt, Germany [S3]

**Table S1. Yeast strains used in this study. Related to Figure 1 - 4.** <sup>†</sup> indicates when the strain used is type strain for the species [S4].

Plasmid	Source
pGREG-505-TEF1	[S5]
pGREG- <i>KmLAC12</i> CBS397	[S5]
pGREG- <i>KmLAC12-2</i> CBS397	[S5]
pGREG- <i>KmLAC12-3</i> CBS397	[S5]
pGREG- <i>KmLAC12-4</i> CBS397	[S5]
pGREG- <i>KaLAC12</i> CBS4438	This study
pGREG- <i>KwLAC12</i> CBS2745	This study
pGREG- <i>KdCEL1</i> CBS2104	This study
p426	Addgene #43803
p426-TEF1- <i>KdCEL2</i> CBS2104	This study
p426-TEF1- <i>KwLAC4</i> CBS2745	This study
pUCC001	[S2]
pUCC001- <i>KmLAC4</i>	[S6]
pUCC001- <i>KmCEL2</i>	This study
pUCC001- <i>KmLAC12</i>	[S2]
pUCC001- <i>KmLAC12-2</i>	[S2]
pUCC001- <i>KmLAC12-4</i>	[S2]

**Table S2. Plasmids used in this study. Related to Figure 4.**

**Heterologous expression**

Name	Sequence
p426-F	cttagattagattgctatgctttc
p426-R	tcatgtaattagttatgtcacgc
pGREG505-F	cctcgagtcacgtaattagttatgtcacgc
pGREG505-R	cgacgggatcgataagcttgatcgaa
<i>KaLAC12</i> -F	gaattcgatatcaagcttatcgataccgctgatgctagaacagagttcaagttct
<i>KaLAC12</i> -R	gctgacataactaattacatgactcgaggtcgacttagaccttattctcgattgt tggt
<i>KwLAC12</i> -F	gaattcgatatcaagcttatcgataccgctgatgctccaagtccatcaaatca
<i>KwLAC12</i> -R	gctgacataactaattacatgactcgaggtcgactcatgctggattcggattcca
<i>KdCEL1</i> -F	gaattcgatatcaagcttatcgataccgctgatgctcaaaggattcaatcttagata gtg
<i>KdCEL1</i> -R	gctgacataactaattacatgactcgaggtcgactcaattaatcgcttcctcct t
<i>KdCEL2</i> -F	gcatagcaatctaataagatgtcaaatctgatgtcgaacaa
<i>KdCEL2</i> -R	tgacataactaattacatgactacaacccgggtccagtaag
<i>KwLAC4</i> -F	gcatagcaatctaataagatgtcaaccaattgattccttc
<i>KwLAC4</i> -R	tgacataactaattacatgactactcaaaaccaatggtgaaggt

**Mutant construction**

Name	Sequence
<i>KmCEL2</i> -gRNA-F	cgtcacggatgtaattgtccgac
<i>KmCEL2</i> -gRNA-R	aaacgctcggacattaacataccgt
<i>KmLAC4</i> -gRNA-F	cgtcgacatctcttaggacaagtt
<i>KmLAC4</i> -gRNA-R	aaacaactgtcctaagagatgtc
Diag- <i>KmCEL2</i> -F	ttgttaaaggatgcagggc
Diag- <i>KmCEL2</i> -R	ggatgctgaaccaccac
Diag- <i>KmLAC4</i> -F	cttaattagctgtacatggct
Diag- <i>KmLAC4</i> -R	tggatttaattggatctgatgg
BSA-R	tacacgcggttgtagagaaaaaaagaaaaattga

**Table S3. Primer sequences used in this study. Related to Figure 4.**



Strain	Mutation coordinates	Source
	<i>KmLAC12</i> :CHR III 16,869..-T	
NBRC1777 <i>lac12 lac12-2 lac12-4</i>	<i>LAC12-2</i> : CHR III 1,557,231..-T <i>LAC12-4</i> : CHR VIII 12,834..-GT	[S2]
NBRC1777 <i>cel2</i>	CHR III $\Delta$ 20,048 – 20,054	This study
NBRC1777 <i>lac4</i>	CHR III 11,255..-A	This study
NBRC1777 <i>cel2 lac4</i>	<i>CEL2</i> : CHR III $\Delta$ 20,048 – 20,054 <i>LAC4</i> : CHR III 11,255..-A	This study

**Table S4. Genotype of the *K. marxianus* mutants constructed in this study. Related to Figure 4.** Coordinates of the mutations in the NBRC1777 genome are shown. The nucleotides deleted are represented by a (-) symbol or  $\Delta$  in cases where several nucleotides were deleted.

## Supplemental References

- S1. Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27004904>
- S2. Varela, J.A., Puricelli, M., Montini, N., and Morrissey, J.P. (2019). Expansion and Diversification of MFS Transporters in *Kluyveromyces marxianus*. *Front. Microbiol.* **9**, 3330. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/30687296>
- S3. Wiczorke, R., Krampe, S., Weierstall, T., Freidel, K., Hollenberg, C.P., and Boles, E. (1999). Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett.* **464**, 123–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10618490>
- S4. Lachance (2011). *Kluyveromyces van der Walt*. In: Kurtzman C.P., Fell J.W., Boekhout T. (eds) *The yeasts : A taxonomic study*, Elsevier, Amsterdam, pp. 471-481
- S5. Varela, J.A., Montini, N., Scully, D., Van der Ploeg, R., Oreb, M., Boles, E., Hirota, J., Akada, R., Hoshida, H., and Morrissey, J.P. (2017). Polymorphisms in the *LAC12* gene explain lactose utilisation variability in *Kluyveromyces marxianus* strains. *FEMS Yeast Res.* **17**. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/28444380>
- S6. Rajkumar, A.S., Varela, J.A., Juergens, H., Daran, J.-M.G., and Morrissey, J.P. (2019). Biological Parts for *Kluyveromyces marxianus* Synthetic Biology. *Front. Bioeng. Biotechnol.*