

# The Genome Sequence of *Saccharomyces eubayanus* and the Domestication of Lager-Brewing Yeasts

EmilyClare Baker,<sup>1</sup> Bing Wang,<sup>1</sup> Nicolas Bellora,<sup>2</sup> David Peris,<sup>1,3</sup> Amanda Beth Hulfachor,<sup>1</sup> Justin A. Koshalek,<sup>4</sup> Marie Adams,<sup>4</sup> Diego Libkind,<sup>2</sup> and Chris Todd Hittinger<sup>\*1,3</sup>

<sup>1</sup>Laboratory of Genetics, Genome Center of Wisconsin, Wisconsin Energy Institute, J. F. Crow Institute for the Study of Evolution, University of Wisconsin–Madison

<sup>2</sup>Laboratorio de Microbiología Aplicada y Biotecnología, Instituto de Investigaciones en Biodiversidad y Medioambiente (INIBIOMA), Universidad Nacional del Comahue-CONICET, Bariloche, Argentina

<sup>3</sup>DOE Great Lakes Bioenergy Research Center, University of Wisconsin–Madison

<sup>4</sup>Biotechnology Center, University of Wisconsin–Madison

\*Corresponding author: E-mail: cthittinger@wisc.edu.

Associate editor: Jianzhi Zhang

## Abstract

The dramatic phenotypic changes that occur in organisms during domestication leave indelible imprints on their genomes. Although many domesticated plants and animals have been systematically compared with their wild genetic stocks, the molecular and genomic processes underlying fungal domestication have received less attention. Here, we present a nearly complete genome assembly for the recently described yeast species *Saccharomyces eubayanus* and compare it to the genomes of multiple domesticated allopolyploid hybrids of *S. eubayanus* × *S. cerevisiae* (*S. pastorianus* syn. *S. carlsbergensis*), which are used to brew lager-style beers. We find that the *S. eubayanus* subgenomes of lager-brewing yeasts have experienced increased rates of evolution since hybridization, and that certain genes involved in metabolism may have been particularly affected. Interestingly, the *S. eubayanus* subgenome underwent an especially strong shift in selection regimes, consistent with more extensive domestication of the *S. cerevisiae* parent prior to hybridization. In contrast to recent proposals that lager-brewing yeasts were domesticated following a single hybridization event, the radically different neutral site divergences between the subgenomes of the two major lager yeast lineages strongly favor at least two independent origins for the *S. cerevisiae* × *S. eubayanus* hybrids that brew lager beers. Our findings demonstrate how this industrially important hybrid has been domesticated along similar evolutionary trajectories on multiple occasions.

**Key words:** domestication, hybridization, *Saccharomyces eubayanus*, lager brewing, genome assembly.

## Introduction

The molecular evolutionary processes of domestication have been extensively studied in plants and animals (Doebly et al. 2006; Lu et al. 2006; Cruz et al. 2008; Wang et al. 2014), but relatively few examples have been investigated in fungi (Rokas 2009; Borneman et al. 2011). Although domesticated microbes achieve tremendous population sizes, lineages can adaptively lose the ability to reproduce sexually when passaged vegetatively (Lang et al. 2009), while other lineages are derived from sterile interspecies hybrids (Querol and Bond 2009; Borneman et al. 2012; Gibson and Liti 2015). The loss of sexual reproduction and potentially extreme bottlenecks could lead to especially dramatic elevations in the fixation of deleterious alleles through Muller's Ratchet (Muller 1964; Felsenstein 1974), with each lineage of domesticated microbe potentially following its own domestication trajectory.

The yeasts used to ferment lager-style beers are examples of highly successful, domesticated interspecies hybrids. Although *Saccharomyces cerevisiae* has been used for millennia to brew ale-style beers and other alcoholic beverages, lager brewing originated more recently in the 15th century central

Europe (Meussdoerffer 2009). Distinguished by their low fermentation temperatures and the settling of yeasts during fermentation, lagers are characterized as having a crisp taste that is distinctive from ale-style beers, which are often associated with relatively pure strains of *S. cerevisiae*. Although ale strains of *S. cerevisiae* are the prototypical brewing yeasts, hybrid lager strains account for 94% of the world market (Riese and Eßlinger 2009).

Historically, lager yeasts have been referred to as *S. pastorianus* syn. *S. carlsbergensis*, but, in the early 1980s, they were shown to be interspecies hybrids of *S. cerevisiae* and a second parental species (Nilsson-Tillgren et al. 1981; Martini and Kurtzman 1985; Martini AV and Martini A 1987). Although the phylogenetic placement of this missing parent had been hypothesized since the early 2000s (Casaregola et al. 2001; Naumova et al. 2005; Nguyen and Gaillardin 2005; Rainieri et al. 2006; Nakao et al. 2009; Nguyen et al. 2011), *S. eubayanus* was first described as an independent species in 2011 when nonhybrid strains were found in association with *Nothofagus* trees in Patagonia (Libkind et al. 2011). Since then, rare isolates have also been isolated in North America (Peris et al. 2014)

© The Author 2015. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

Open Access

and China (Bing et al. 2014). However, the only evidence of *S. eubayanus* in Europe remains the hybrid lager-brewing strains of *S. cerevisiae* × *S. eubayanus* and other genetically complex interspecies hybrids (Almeida et al. 2014; Gibson and Liti 2015). *Saccharomyces eubayanus* has not been definitively associated with human-controlled fermentations, except as interspecies hybrids (Rodríguez et al. 2014; Gibson and Liti 2015).

Lager yeasts consist of two distinct lineages, both of which were isolated in the late 19th century and form the basis of lager brewing today (Gibson et al. 2013; Gibson and Liti 2015). These lineages, Saaz and Frohberg, were named for the areas in which they were isolated, and are also known as Group I and Group II strains, respectively (Dunn and Sherlock 2008). Other authors distinguish Saaz/Group I strains taxonomically as *S. carlsbergensis* (Wendland 2014) and Frohberg/Group II strains as *S. pastorianus*. Each lineage has unique flavor and brewing profiles, as well as genetic compositions, but both lineages are *S. cerevisiae* × *S. eubayanus* allopolyploid hybrids (Dunn and Sherlock 2008; Gibson et al. 2013; Walther et al. 2014). Saaz strains have a slightly greater capacity to grow at low temperatures than Frohberg strains, and the two lineages also produce differing amounts of various esters important to the taste of the final product (Gibson et al. 2013; Walther et al. 2014). Saaz strains also have relatively poor fermentation performance compared with Frohberg strains, at least partly due to their poor utilization of maltose and maltotriose, which comprise 45–65% and 16–26% of all available sugars, respectively (Boulton and Quain 2007). This deficiency likely contributes to the predominance of Frohberg strains in modern industrial-scale brewing (Gibson et al. 2013).

Greater tolerance of lower temperatures and poor utilization of maltose and maltotriose make Saaz strains more physiologically similar to their *S. eubayanus* parent than Frohberg strains (Gibson et al. 2013). Perhaps not surprisingly, Saaz strains contain a higher proportion of *S. eubayanus* DNA in their genome, approximately one full diploid *S. eubayanus* genome and a haploid *S. cerevisiae* genome (i.e., allotriploids) (Walther et al. 2014). In comparison, Frohberg strains have a more equitable composition of parental genomes so that Frohberg genomes comprise approximately one full diploid *S. eubayanus* genome and one full diploid *S. cerevisiae* genome, making them approximately allotetraploid hybrids (Nakao et al. 2009; Walther et al. 2014).

The origin of these lineages is still contentious. One hypothesis proposes that the Saaz and Frohberg lineages resulted from independent hybridization events (Liti et al. 2005; Dunn and Sherlock 2008; Bond 2009). The most extensive study to date (Dunn and Sherlock 2008) examined the genome composition of multiple lager yeast strains using aCGH (array comparative genomic hybridization) with probes from *S. cerevisiae* S288c and *S. uvarum* CBS 7001 (previously *S. bayanus* var. *uvarum*), the sister species of *S. eubayanus*, as well as limited sequence data. However, concerns have remained about the use of *S. uvarum* as a proxy for *S. eubayanus* and the low resolution of aCGH, and subsequent findings have cast doubt on the multiple origin hypothesis. Specifically, three translocations between the *S. eubayanus*

and *S. cerevisiae* chromosomes of lager yeasts share identical breakpoints between the Saaz and Frohberg lineages (Hewitt et al. 2014; Walther et al. 2014). This led Jürgen Wendland et al. to propose that both lineages were derived ultimately from a single hybridization event between a diploid *S. eubayanus* parent and a diploid *S. cerevisiae* parent (or two haploids followed by endoreduplication) (Walther et al. 2014; Wendland 2014). Under this hypothesis, the Frohberg strains represent the ancestral state of all lager-brewing yeast lineages, while Saaz strains were derived later from a rare viable allodiploid meiotic spore that subsequently mated with a haploid *S. eubayanus* spore to form an allotriploid lineage. Other groups, however, have argued on the basis of experimental evidence and limited gene sequence data from lager strains that identical breakpoints could have resulted from independent events at recombination hotspots or fragile sites in *Saccharomyces* chromosomes (Hewitt et al. 2014; Monerawela et al. 2015). Consequently, it is still unclear which hypothesis best explains the origins of lager-brewing yeasts.

Understanding the origins and evolution of lager yeasts, as well as the genetics of what makes them such successful industrial strains compared with their nonhybrid parents, is still an active area of research. Although a draft assembly of *S. eubayanus* has been available since 2011, this initial assembly has a number of drawbacks (Libkind et al. 2011). Most critically, it was assembled with a relatively low coverage of 36-bp single-end reads using the *S. uvarum* genome as a reference. Although *S. uvarum* is the sister species to *S. eubayanus*, they are, nevertheless, approximately 7% diverged at the sequence level (Libkind et al. 2011), roughly the same distance as between humans and macaques (Gibbs et al. 2007). This reference-based assembly cannot account for any translocations in the *S. eubayanus* genome relative to the *S. uvarum* genome and lacks crucial information on *S. eubayanus* subtelomeric sequences. Complete and accurate assembly of the subtelomeric regions is of special interest because they often harbor novel and highly polymorphic genes, such as many of those important for brewing (Brown et al. 2010), but these regions have been traditionally difficult to assemble (Brown et al. 2010; Ellegren 2014). Recently, an improved genome assembly for *S. eubayanus* was published by Heby et al. (2015), but this assembly still lacks critical coverage of the *S. eubayanus* subtelomeric sequences and fails to provide scaffolded *S. eubayanus* chromosomes.

In the absence of a high-quality reference genome, putative *S. eubayanus* genes within lager-brewing yeasts have been described based on the apparent absence of close homologs within *S. cerevisiae* and, when their genomic location is known, their physical proximity to non-*S. cerevisiae* portions of lager yeast genomes (Dietvorst et al. 2005; Nakao et al. 2009; Vidgren et al. 2010; Cousseau et al. 2013). With our assembly of the *S. eubayanus* genome, it is now possible to study lager yeasts in the context of the complete genomes of both parental species. Not only does our *S. eubayanus* assembly fill this important information gap, its quality and coverage, both in terms of depth and completeness, exceeds that of any *Saccharomyces* genome outside of *S. cerevisiae*. In

addition to setting a new benchmark for genome assembly quality, the *S. eubayanus* genome also demonstrates what is possible with current Illumina sequencing technology and assembly algorithms. Here, we analyze the genome evolution of *S. eubayanus* and its domesticated hybrids to begin to infer how these genomes have changed during domestication, yielding unprecedented insights into the complex origins and evolution of these industrially important hybrids.

## Results/Discussion

### A High-Quality Annotated De Novo Genome Assembly of *Saccharomyces eubayanus*

To study the evolution of the genome sequence and structure of *S. eubayanus* and its hybrid descendants, we assembled de novo the genome sequence of FM1318—a monosporic derivative of the species type strain (CRUB 1568<sup>T</sup>=PYCC 6148<sup>T</sup>=CBS 12357<sup>T</sup>). To enable the use of the ALLPATHS-LG genome assembler, we built two specialized Illumina libraries: A “fragment library” with paired-end 300-bp reads (i.e., 2 × 300 bp) and a “jumping library” with mate-pair reads with an average insert size of approximately 6.5 kb. Briefly, ALLPATHS-LG first joins paired-end reads from the fragment library that overlap to create longer reads, from which it builds a de Bruijn graph to construct contigs; the longer insert jumping library is then incorporated into the de Bruijn graph to scaffold the contigs, resolve repeats, and flatten the graph (Ribeiro et al. 2012). Because all *Saccharomyces* genomes contain Ty retrotransposons that are approximately 6 kb, duplicate gene families, and several other large repeats, a long-read or long-insert scaffolding strategy is critical to providing physical evidence that spans gaps to order and orient contigs. ALLPATHS-LG performed well on the nuclear genome but poorly on the two-micron plasmid and the mitochondrial genome, both of which are circular and present at higher copy numbers. Instead, we assembled these genomes using SPAdes (Bankevich et al. 2012) which produced complete contigs that we manually circularized. We performed several additional procedures and quality checks to combine and improve the final assembly (see Materials and Methods).

This assembly strategy resulted in a genome assembly of higher quality than any published for a *Saccharomyces* species other than *S. cerevisiae* (Liti et al. 2009, 2013; Scannell et al. 2011; Engel et al. 2014). Specifically, the 11.66-Mb nuclear genome was found on 22 scaffolds, and the scaffold N50 was 896 kb. These scaffolds were built from 144 contigs, with a contig N50 of 198 kb. Manual inspection of gaps suggested that most were due to Ty elements or other repetitive sequences. With the exception of six scaffolds ranging from 2 to 13 kb in length, all scaffolds were placed onto chromosomes. Only chromosome XII had an internal gap, which corresponded to the ribosomal DNA (*rDNA*) repeats and is conserved with *S. cerevisiae*. Indeed, the genome of *S. eubayanus* is nearly syntenic with *S. cerevisiae* with the exception of a handful of small inversions and two previously documented reciprocal translocations that it shares with its sister species, *S. uvarum* (fig. 1) (Fischer et al. 2000; Scannell et al. 2011). One assembled chromosome (chromosome VIII) contained

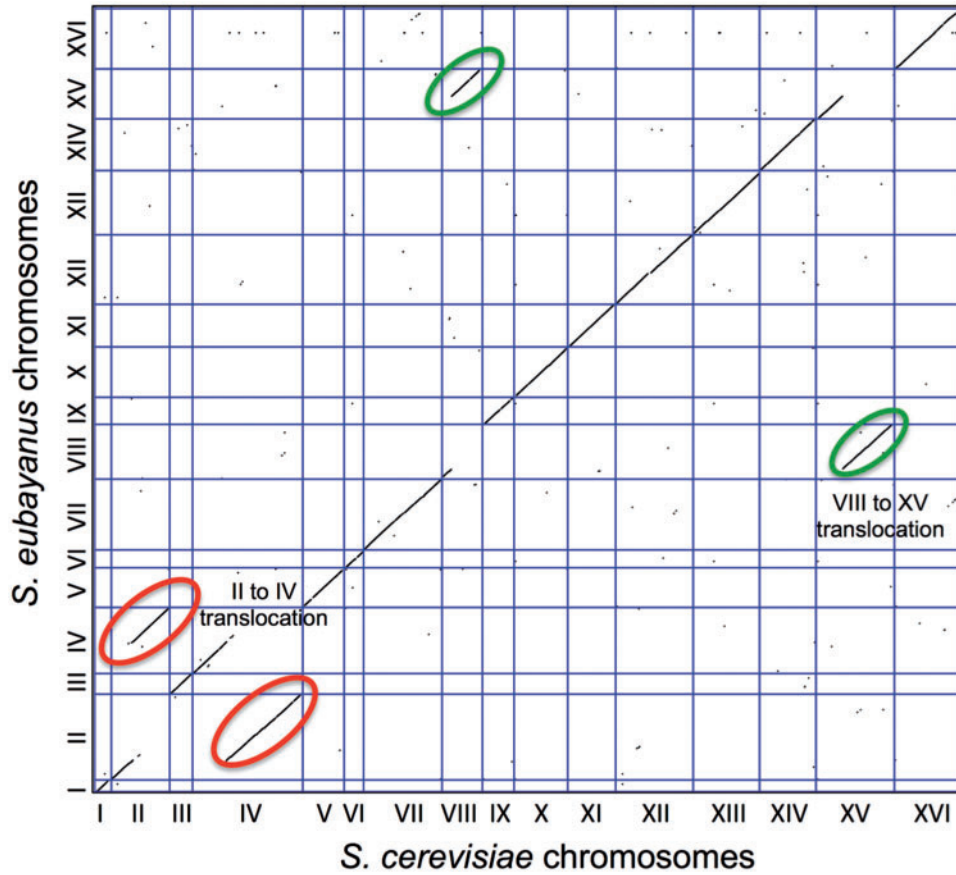
telomeric repeats, and several contained well-resolved multi-copy subtelomeric gene families.

We adapted the Yeast Genome Annotation Pipeline (YGAP) to annotate our high-quality de novo assembly of the *S. eubayanus* genome (see Materials and Methods). This pipeline resulted in 5,515 predicted protein-coding genes for *S. eubayanus*, which is similar to the current draft genomes of other *Saccharomyces* species (Liti et al. 2009, 2013; Scannell et al. 2011). Of the predicted protein-coding genes, 4,993 were unambiguous 1:1:1 orthologs among *S. cerevisiae*, *S. uvarum*, and *S. eubayanus*. Because the Saaz and Froberg lineages of lager-brewing yeasts contain largely complete but sometimes nonoverlapping copies of both the *S. cerevisiae* and *S. eubayanus* genomes (Walther et al. 2014), we separately considered 1:1:1:1 orthologs for each lineage. Specifically, the *S. cerevisiae*:*S. cerevisiae* (Froberg):*S. uvarum*:*S. eubayanus*:*S. eubayanus* (Froberg) set had 3,649 orthologs, while the *S. cerevisiae*:*S. cerevisiae* (Saaz):*S. uvarum*:*S. eubayanus*:*S. eubayanus* (Saaz) set had only 3,102 orthologs due to the loss of more *S. cerevisiae* genes, principally through the loss of chromosomes (Dunn and Sherlock 2008). A total of 2,268 orthologs were common between the Saaz 1:1:1:1 ortholog set and the Froberg 1:1:1:1 ortholog set. For phylogenetic analyses, we also included *S. paradoxus* sequences (1:1:1:1:1 ortholog sets), as an outgroup for *S. cerevisiae* lager and nonlager sequences, leaving 2,194 orthologs.

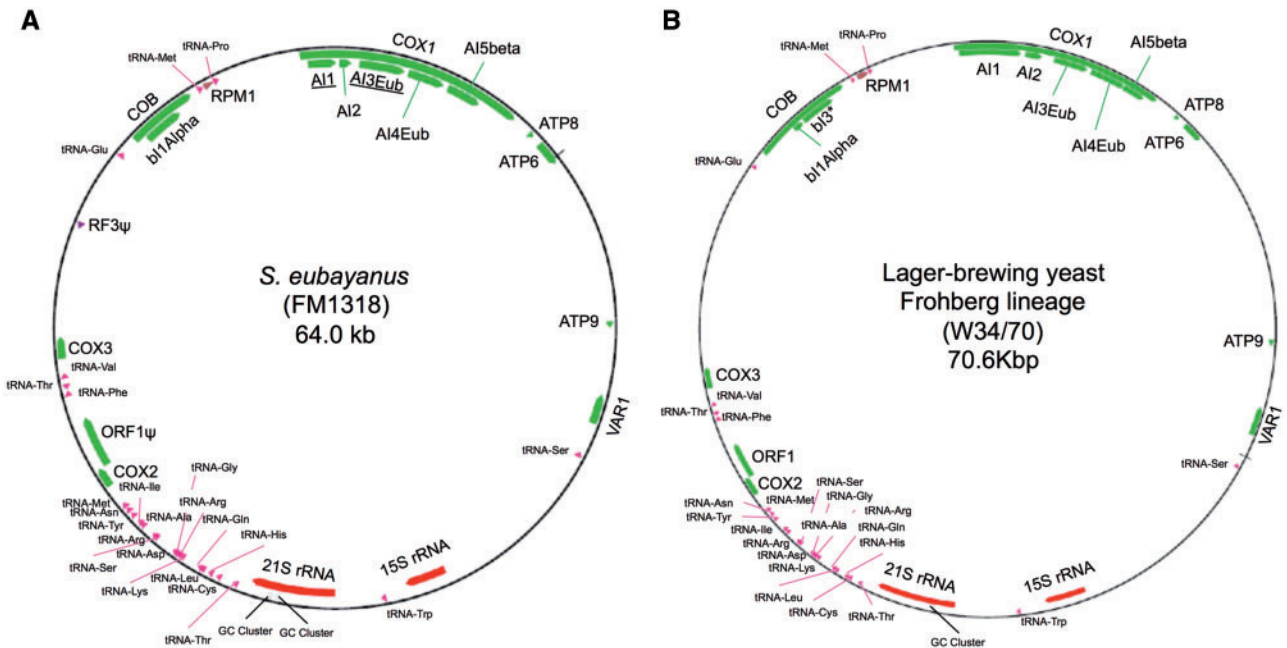
Based on the shared set of 1:1:1:1 orthologs (coding sequences [CDSs]) between the Saaz and Froberg genomes, the *S. cerevisiae* subgenome of the Saaz lineage is 99.57% identical to S288c, while in the Froberg lineage, the *S. cerevisiae* subgenome is 99.60% identical. The *S. eubayanus* subgenomes of both lager strains are 99.55% identical to FM1318 (supplementary table S1, Supplementary Material online), similar to a previous estimate of 99.56% (Libkind et al. 2011).

### The *Saccharomyces eubayanus* Mitochondrial Genome

The *S. eubayanus* mitochondrial genome (mtDNA) is 64 kb, which is 6.6 kb smaller than the mtDNA of Froberg lager-brewing yeast (Nakao et al. 2009) and 21.8 kb smaller than the mtDNA of *S. cerevisiae* (S288c). *Saccharomyces eubayanus* mtDNA has a similar gene order to the Froberg representative, but it differs in the number of introns in COB and the locations of COX1 introns (fig. 2). Both *S. eubayanus* and Froberg mtDNA show two rearrangements relative to *S. cerevisiae* (S288c), one involving 15S rRNA and tRNA-Trp loci and a second involving the tRNA-Glu and COB loci (supplementary fig. S1, Supplementary Material online). These structural and sequence similarities establish *S. eubayanus* as the main donor of mtDNA for lager yeasts of the Froberg lineage. In addition, our analyses also revealed a dynamic history of selfish elements and possible localized introgression with other *Saccharomyces* species. Overall, the Froberg mtDNA CDSs are approximately 98.56% identical to the *S. eubayanus* mtDNA (supplementary results S1 and supplementary figs. S1 and S2, Supplementary Material online).



**Fig. 1.** Dot plot comparing the location of genes in the *Saccharomyces eubayanus* (FM1318) genome assembly with their location in *S. cerevisiae* (S288c). Lines circled in the same color indicate reciprocal translocations.



**Fig. 2.** Schematic representation of *Saccharomyces eubayanus* (FM1318) and lager-brewing yeast (Frohberg strain W34/70) annotated mitochondrial genomes. Mitochondrial genes, rRNAs, tRNAs, and noncoding RNAs are represented in green, red, pink, and brown, respectively. Genes with asterisks are elements or gene sequences not shared by both *S. eubayanus* and the lager yeast mitochondrial genomes. Underlined names are intronic regions located in different positions between *S. eubayanus* and Frohberg lager yeast.

Downloaded from https://academic.oup.com/mbe/article/32/1/1/2818/982141 by guest on 23 April 2024

## Characterization of the Maltose (MAL) Utilization Genes of *Saccharomyces eubayanus*

The MAL genes allow the sugars maltose, maltotriose, and related sugars to be utilized as carbon sources. They are typically subtelomeric and often found in clusters consisting of genes encoding a maltose permease (MALT), a maltase (MALS), and a transcriptional regulator of the pathway (MALR) (Needleman 1991). *Saccharomyces cerevisiae* also contains additional maltose utilization genes, including genes encoding the isomaltases IMA1–IMA5 (Teste et al. 2010) and the maltose and maltotriose transporters MPH2 and MPH3 (Day et al. 2002). These genes are also subtelomeric but are not necessarily found in close association with other maltose utilization genes. For simplicity, we refer to genes related to maltose utilization collectively as MAL genes. The industrial importance of understanding maltose utilization, particularly the *S. eubayanus* versions of MAL genes in lager-brewing yeasts, and the historic difficulty of assembling the subtelomeric regions, makes these genes excellent candidates to test the quality and completeness of our *S. eubayanus* assembly.

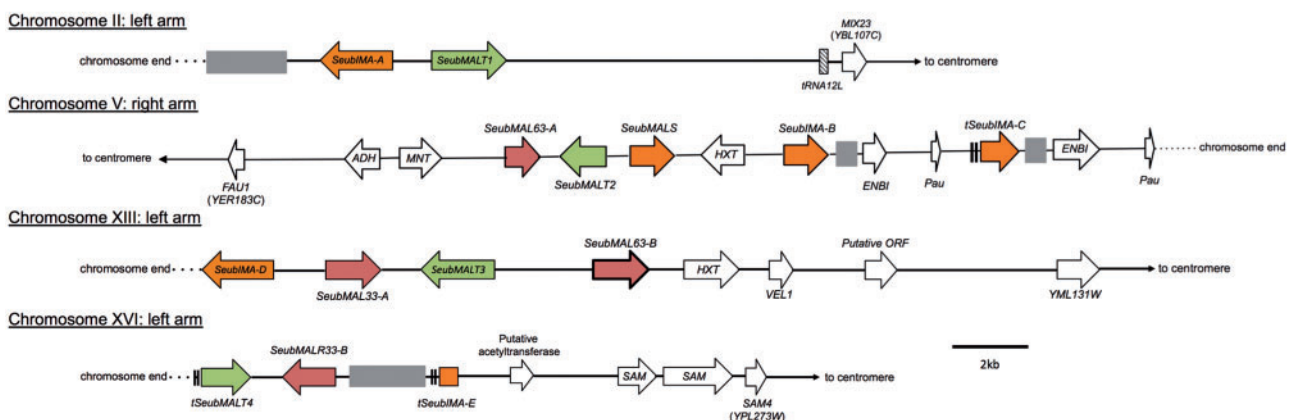
An extensive search of the *S. eubayanus* genome revealed 14 genes related to maltose utilization spread across four chromosomes (fig. 3 and supplementary table S2, Supplementary Material online), as well as two putatively pseudogenized MAL genes on an unplaced scaffold (supplementary results S2, Supplementary Material online). For comparison, in the most recent assembly of the *S. eubayanus* genome, which was based entirely on Illumina paired-end reads, only three of these MAL genes could be identified,

and none could be placed on chromosomes (Hebly et al. 2015). In our analysis, all MAL genes were found in subtelomeric clusters (fig. 3) and, in most cases, were the last genes before the end of the assembled chromosome sequence. The MAL genes were often found in association with other genes known to be related to fermentation and brewing, such as genes encoding hexose transporters and alcohol dehydrogenases, consistent with the tendency of subtelomeric regions to harbor brewing-related genes (Brown et al. 2010).

## The Origins of the MAL Genes of Lager-Brewing Yeasts

To infer the origins of MAL genes in lager yeasts, we extracted homologs from the publicly available genome assemblies of the two lager yeast lineages, Saaz and Froberg (Walther et al. 2014), by using both the *S. cerevisiae* (S288c) (Cherry et al. 2012; Engel et al. 2014) and *S. eubayanus* sequences as BLAST queries. Because ale strains are the most likely *S. cerevisiae* parent of lager-brewing yeasts (Dunn and Sherlock 2008; Querol and Bond 2009; Monerawela et al. 2015), MAL genes were also extracted from the publicly available genomes of the ale strains Foster's O and Foster's B (Borneman et al. 2011). Many ale MAL genes were incomplete or, in some cases, completely missing from previously published assemblies, so we relied primarily on the MAL gene sequences from S288c for our comparisons. Where possible, lager MAL genes were assigned as either of *S. cerevisiae* or *S. eubayanus* origin based on sequence similarity (supplementary table S2, Supplementary Material online).

### *S. eubayanus* (FM1318)



**FIG. 3.** Genome regions in *Saccharomyces eubayanus* (FM1318) with clusters of genes related to maltose (MAL) utilization. Regions are represented from chromosome ends to the first gene that is syntenic with *S. cerevisiae* (S288c). Gene sizes and distances are approximately to scale. Arrows show the direction of transcription and direction to the centromeres. Gray boxes represent sequence gaps. MAL genes are colored: Orange genes encode maltases and isomaltases, green genes encode maltose transporters (permeases), and red genes encode transcription factors that regulate other MAL genes. MAL genes are named for their closest *S. cerevisiae* homolog in S288c and prefixed with “Seub”. *S. eubayanus* MAL genes that are most similar to the same *S. cerevisiae* MAL gene are distinguished by letters (*SeubIMA-A*, *SeubIMA-B*, etc.). Double lines before or after a gene represent incomplete sequence due to poor sequence resolution in those areas and their names are marked with a “t” for truncated. All non-MAL genes in *S. eubayanus* are named for their closest gene family in *S. cerevisiae* using standard names with the exception of genes that are the first gene syntenic with *S. cerevisiae* (S288c). The first *S. eubayanus* genes syntenic with S288c are named using the standard names of their *S. cerevisiae* syntenic homologs, where available, along with their systematic names.

Although many lager *MAL* genes had very high sequence identity with either *S. cerevisiae* (S288c) or *S. eubayanus* (FM1318) genes, in a number of cases, the identity was insufficient to identify them as being of *S. eubayanus* or *S. cerevisiae* origin. In most cases, the similarity between the ale strain *MAL* genes and their lager homologs was about the same as for S288c. In several cases, however, the ale strain *MAL* genes were much more similar to the lager yeast genes than the sequences from S288c. Several ale *MAL* genes had over 98% nucleotide identity to their lager yeast counterparts, whereas the closest S288c homolog had less than 85% identity; in one case, the nearest homolog had even been identified as a *S. eubayanus* gene. Because many of the *MAL* genes extracted from the ale genomes were incomplete, were of poor sequence quality, or had frameshifts that were likely 454 sequencing artifacts, these results are probably an underestimate of the true similarity of ale strains to the *S. cerevisiae* genomes of lager-brewing yeasts. In some cases, the percent identity of a lager *MAL* gene to its closest *S. cerevisiae* or *S. eubayanus* homolog was low enough that its true origin remained ambiguous. Despite relatively low sequence identities, synteny sometimes supported the *S. eubayanus* origin of lager yeast genes where sequence identity was ambiguous (fig. 4A and supplementary table S2, Supplementary Material online). For example, the gene order within *MAL* clusters from *S. eubayanus*, *S. cerevisiae*, and the Saaz and Frohberg lager yeast lineages revealed one cluster in both lager yeast lineages that is largely syntenic with the *S. eubayanus* *MAL* cluster on chromosome V and another *MAL* cluster in the Saaz lineage that is syntenic with the cluster on chromosome X of S288c (fig. 4A). The common origin of many genes within these clusters is further supported by phylogenetic analysis (fig. 4B). Although we could establish the parent species for a number of lager yeast *MAL* genes, the origin of many *MAL* genes remains ambiguous, including the previously identified lager-brewing yeast-specific genes, *MTT1* and *Lager-AGT1* (supplementary results S3 and supplementary tables S2 and S3, Supplementary Material online) (Dietvorst et al. 2005; Nakao et al. 2009; Vidgren et al. 2010; Cousseau et al. 2013). To improve identification of the source of *MAL* genes and other brewing-related genes in lager yeasts, it will be necessary to identify strains more closely related to the parental strains, both in *S. cerevisiae* and *S. eubayanus*.

### A Genome-Wide Signature of Domestication in Lager-Brewing Yeasts

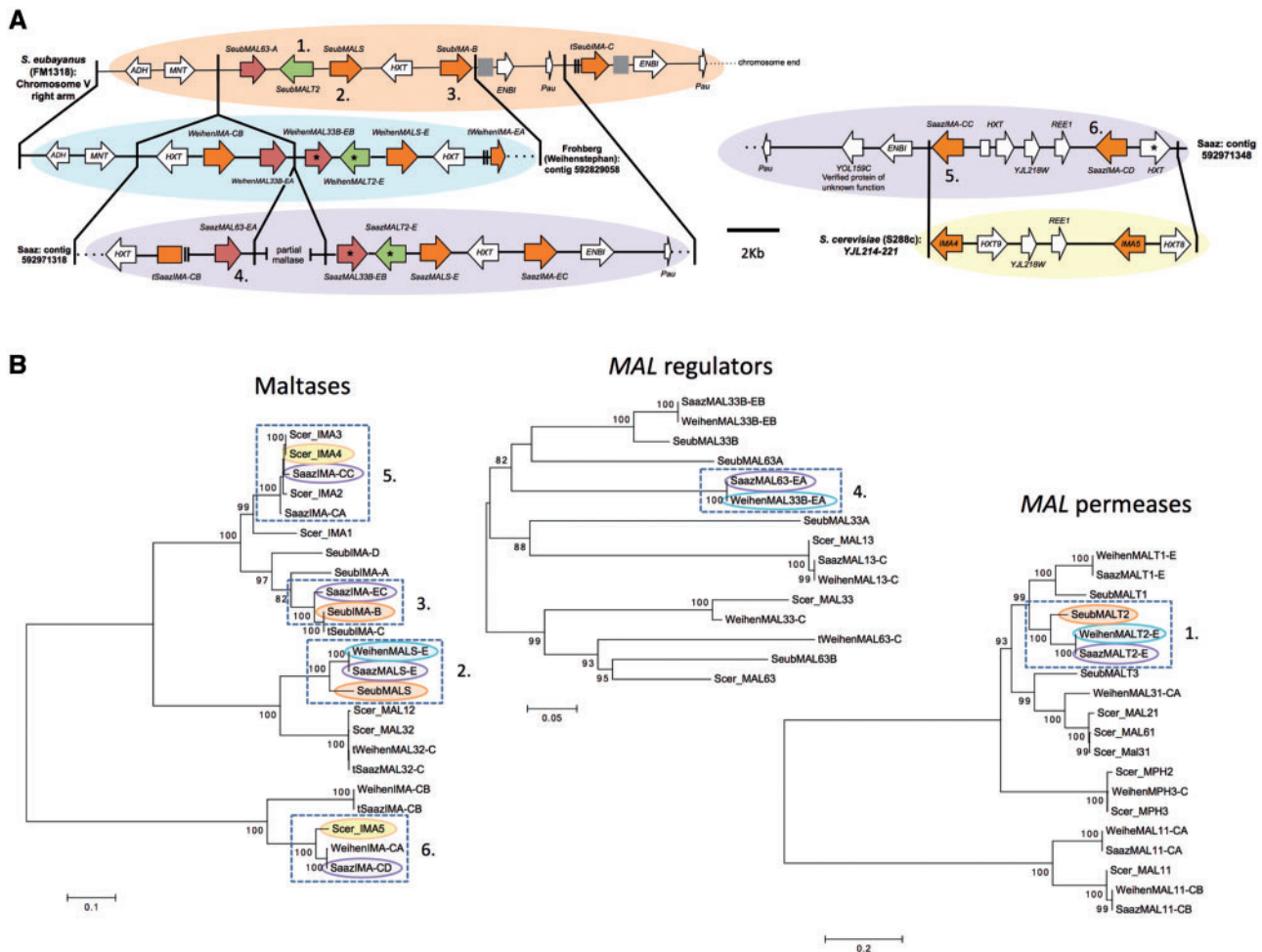
To test for the genome-wide consequences of domestication, we compared the genome-wide distributions of the ratio of nonsynonymous substitutions to total substitutions per gene between different lineages of *Saccharomyces* and the subgenomes of lager-brewing yeast hybrids. Specifically, we considered genes with homologs in *S. cerevisiae*, *S. paradoxus*, *S. eubayanus*, *S. uvarum*, and where both the *S. cerevisiae* and *S. eubayanus* homologs had been retained in both the lager-brewing yeast lineages (1:1:1:1:1 ortholog sets). Using PAML (Yang 2007), the rate of synonymous substitution ( $d_S$ ) and the rate of nonsynonymous substitution ( $d_N$ ) were computed for

each gene along each lineage, from which we estimated the number of substitutions. Saaz and Frohberg were considered separately as the lager yeast representative (supplementary data S1, Supplementary Material online).

Consistent with domestication having had a genome-wide impact, we found that the proportion of nonsynonymous changes out of all substitutions per gene was significantly greater ( $P < 10^{-21}$  for both Frohberg and Saaz) in the *S. eubayanus* subgenome of lager lineages than in the wild Patagonian isolate (fig. 5). For all comparisons, the results were not dependent on whether the Saaz or Frohberg lineage was used as the lager yeast representative, and the Saaz and Frohberg lineages themselves did not differ from one another ( $P \gg 0.05$ ). These results strongly suggest that purifying selection has played a much more limited role in the evolution of the *S. eubayanus* subgenome of lager yeasts than in the genome of its nondomesticated relative. Although the polyploid state of lager yeast could also have contributed, relaxation of purifying selection under domestication conditions was probably the primary cause of the increase in the fixation of nonsynonymous substitutions.

Demographic factors outside of domestication, such as a hypothesized population bottleneck in the lineage leading to the *S. eubayanus* strains of the Northern Hemisphere (Almeida et al. 2014; Peris et al. 2014), could also have contributed to the observed increase in nonsynonymous substitutions. Prior to their hybridization with *S. cerevisiae*, these strains were likely part of a European subpopulation of *S. eubayanus* with very low genetic diversity. By comparing the number of substitutions since the divergence of these lineages from each other (supplementary data S2, Supplementary Material online) with the number of substitutions since their divergence with FM1318, we estimate that at least 3% of the observed divergence can be attributed to processes that occurred after the *S. eubayanus* parents of the two lager yeast lineages diverged from one another. Importantly, this subset of substitutions appears to show an even stronger bias toward nonsynonymous changes than the rest of the genome (nonsynonymous changes represent 55% of substitutions from the common ancestor of Saaz and Frohberg versus 36% from their ancestor with FM1318). Thus, the *S. eubayanus* subgenomes in these interspecies hybrids seem to have experienced a substantial relaxation of purifying selection beyond any demographic factors that may have occurred within wild *S. eubayanus* during the divergence of the relevant Patagonian and Northern Hemisphere subpopulations.

We have likely underestimated the impact of domestication on the evolution of the *S. eubayanus* subgenomes in lager yeasts because our estimate assumes that domestication began only after the divergence of the *S. eubayanus* strains that would give rise to the Saaz and Frohberg lineages. The population of *S. eubayanus* that gave rise to lager hybrids could have already begun to adapt to brewing prior to the divergence of those two lineages. Nonetheless, it is unlikely that *S. eubayanus* strains were as strongly associated with brewing as ale strains of *S. cerevisiae* prior to hybridization. Consistent with the hypothesis that *S. eubayanus*



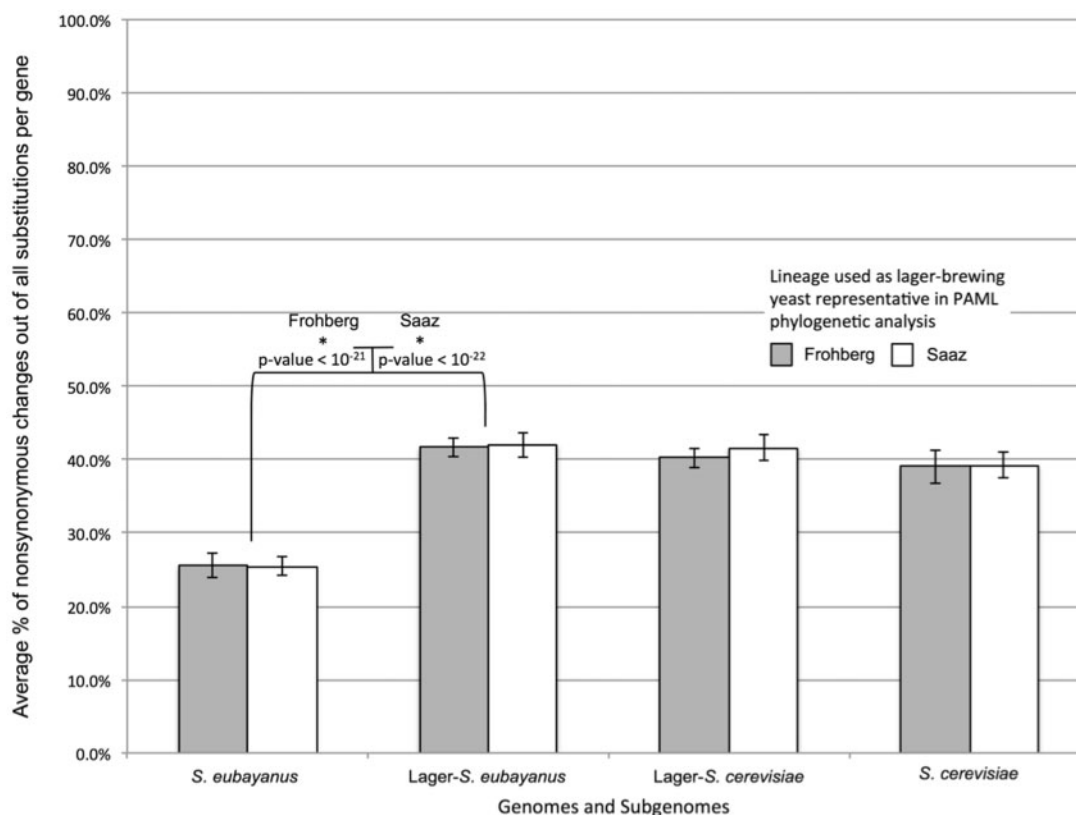
**Fig. 4.** Synteny and phylogenetic analysis of MAL genes from *Saccharomyces eubayanus* (FM1318), *S. cerevisiae* (S288c), and the Saaz (CBS 1513) and Froberg (W34/70) lineages of lager-brewing yeasts. Numbers in (A) and (B) indicate genes whose orthology is supported both by synteny and phylogenetic analysis. (A) Solid lines connecting genomes designate blocks of synteny. Chromosome and contig locations are indicated to the left or right of genome segments. The location of the *S. cerevisiae* segment is indicated by the systematic names of the genes within the syntenic region. Asterisks indicate genes with complete sequences but putatively inactivating mutations. The inactivated Saaz *HXT* in the region syntenic to *S. cerevisiae* is divided into two due to an insertion within the gene. Genes are colored as in figure 3. Gene sizes and distances are approximately to scale. Arrows show the directions of transcription. Gray boxes represent gaps in the sequence. Double lines before or after a gene represent incomplete sequence due to poor resolution in those areas, and their names are marked with a “t” for truncated. Dotted lines represent the end of a chromosome or contig. (B) Maximum likelihood trees for maltases, regulators of MAL genes, and maltose permeases (transporters) based on nucleotide sequences. Branch lengths are based on the number of substitutions per site. Bootstrap support values of 70 or higher are shown at nodes. Genes present in the synteny analysis are highlighted by an oval of the same color as their genome blocks in (A). Dashed boxes indicate groups of genes whose orthology is also supported by synteny. More details on these genes can be found in [supplementary table S2, Supplementary Material](#) online.

subgenomes were more affected by domestication, the *S. cerevisiae* subgenomes of lager hybrids do not show the same degree of an increase in nonsynonymous changes after the divergence of the two lager lineages (fig. 5).

### Some Metabolism Genes May Have Been Shaped by Domestication to Brewing

Genes that have experienced elevated rates of protein sequence evolution in both the Saaz and Froberg lineages but that acquired different nonsynonymous changes are of special interest because they may indicate independent responses to the same environmental and domestication pressures. The *S. eubayanus* alleles of *NOT3* have experienced an

independent elevation of the relative rate of nonsynonymous substitution (Saaz  $\omega = 2.10$ , Froberg  $\omega = 1.57$ ) ( $\omega = \text{rate of nonsynonymous substitution } (d_N) / \text{rate of synonymous substitution } (d_S)$ ). This global regulator of transcription (Collart et al. 2013) has further acquired a large number of changes along both *S. cerevisiae* lager branches, including five unshared nonsynonymous differences between the Saaz and Froberg alleles. Similarly, *ADR1* encodes a transcription factor required for the expression of the alcohol dehydrogenase specialized for ethanol oxidation (*ADH2*), rather than fermentation, as well as the regulation of other genes related to the utilization of ethanol (Young et al. 2003). In the Saaz lineage, the *ADR1* allele derived from *S. cerevisiae* appears to have undergone rapid evolution ( $\omega = 2.23$ ), while the



**Fig. 5.** Genome-wide averages of the percent of nonsynonymous substitutions out of all substitutions per gene for the genomes of *Saccharomyces cerevisiae*, *S. eubayanus*, and the *S. cerevisiae* and *S. eubayanus* subgenomes of lager-brewing yeast hybrids when using either the Frohberg (gray) or Saaz (white) lineage as the representative of lager-brewing yeasts. Substitutions were estimated by PAML phylogenetic analysis using orthologs from *S. cerevisiae*, *S. paradoxus*, *S. eubayanus*, and *S. uvarum* and orthologs that are present in both the *S. cerevisiae* and *S. eubayanus* subgenomes of the Saaz and Frohberg lineages. Only the 2,194 genes shared by both 1:1:1:1:1 ortholog sets were included in the analyses. Error bars represent 99% binomial confidence intervals. Comparisons between genomes and subgenomes (and between using Saaz versus Frohberg as the lager yeast representative) were made using logistic regression. Asterisks represent statistically significant comparisons. Note that the enrichment of nonsynonymous changes in lager lineages of *S. eubayanus* is not due to the inverse correlation between  $\omega$  and  $d_s$ : The average  $d_s$  for each branch left to right is 0.0045, 0.0046, 0.0061, 0.0062, 0.0046, 0.0050, 0.0046, and 0.0047; the average  $d_n$  for each branch is 0.0006, 0.0006, 0.0017, 0.0019, 0.0010, 0.0013, 0.0010, and 0.0010.

*S. eubayanus* allele also appears to have an elevated rate of evolution ( $\omega = 0.94$ ). In the Frohberg lineage, the estimated rates of evolution for both the *S. cerevisiae* allele ( $\omega = 0.87$ ) and the *S. eubayanus* allele ( $\omega = 1.45$ ) are moderately elevated. Finally, *REG2*, an important regulator of glucose-repressed genes (Frederick and Tatchell 1996), especially the *MAL* genes (Jiang et al. 2000), experienced five nonsynonymous substitutions along the *S. eubayanus* Frohberg branch. Many caused radical amino acid changes that may have reduced the function of the protein, allowing it to adjust to the needs of the new glucose-poor but maltose-rich environment.

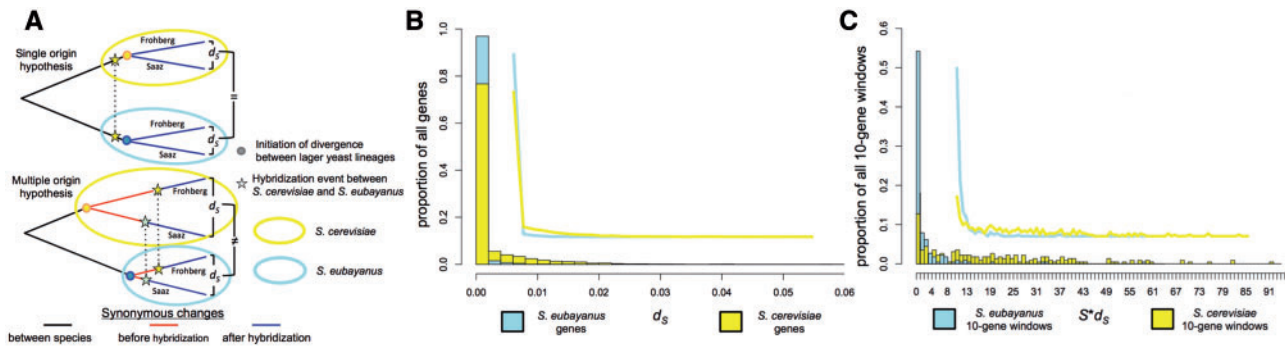
### Multiple Independent Origins of Lager-Brewing Yeasts

The single origin (Walther et al. 2014; Wendland 2014) and multiple origin (Liti et al. 2005; Dunn and Sherlock 2008; Bond 2009) hypotheses for the origins of lager yeasts make distinct predictions about the divergences that would be expected between the *S. eubayanus* and *S. cerevisiae* subgenomes of the Saaz and Frohberg lineages (fig. 6A). In the case of a single origin, both the *S. eubayanus* and *S. cerevisiae* subgenomes

would have begun with identical sequences in the two lineages. Once the *S. eubayanus* and *S. cerevisiae* subgenomes were present in the same hybrid nucleus, they would be expected to have accumulated neutral substitutions at the same rate. As a result, the single origin hypothesis predicts that the neutral divergence between the *S. cerevisiae* subgenome of the Saaz lineage and the *S. cerevisiae* subgenome of the Frohberg lineage should be equivalent to the divergence of the *S. eubayanus* subgenomes of the Saaz and Frohberg lineages. In contrast, the multiple origin hypothesis predicts that the parental genomes could have already begun to diverge from each other before hybridizing, leading to potentially different levels of neutral divergence.

Consistent with the multiple origin hypothesis, the rate of synonymous substitution ( $d_s$ ) and the estimated number of synonymous substitutions ( $S \times d_s$ ) were over ten times higher for the *S. cerevisiae* subgenome than for the *S. eubayanus* subgenome (supplementary data S2 and supplementary table S4, Supplementary Material online). This pattern was not driven by a few genes but was rather distributed across the entire genome (fig. 6B [ $P < 10^{-206}$ ] and C [ $P < 10^{-44}$ ]). The relatively even distribution of divergence was also





**FIG. 6.** (A) Hypothetical models for the origin of hybrid lager-brewing yeast lineages and the relative neutral divergence ( $d_s$  and  $S \times d_s$ ) between the subgenomes of each lineage under each model. (B) and (C) are the distributions of the estimated synonymous rates of evolution ( $d_s$ ) and the estimated number, respectively, of synonymous substitutions ( $S \times d_s$ ) for all genes and ten gene windows respectively, between either the *Saccharomyces cerevisiae* subgenomes (yellow) or the *S. eubayanus* subgenomes (blue) of the two lager-brewing yeast lineages. Line drawings within graphs represent outlines of the histograms to better show their overall distributions.

inconsistent with more complex models involving the differential loss of heterozygosity between lineages (supplementary fig. S3, Supplementary Material online). This result provides a clear indication that the *S. cerevisiae* subgenomes of the two lineages of lager-brewing yeasts had begun to diverge from each other well before the *S. eubayanus* subgenomes. The directionality of this result is consistent with population genetic studies that suggest that there was very limited genetic diversity among the *S. eubayanus* strains that would have produced lager-brewing yeasts (Peris et al. 2014) and that standing variation among *S. cerevisiae* ale strains was higher (Dunn and Sherlock 2008). In this context, these results suggest that the Saaz and Frohberg lineages were created by at least two distinct hybridization events between nearly identical strains of *S. eubayanus* with relatively more diverse ale strains of *S. cerevisiae*.

## Conclusions

### The Origins of Lager-Brewing Yeasts

The finding that multiple hybridization events gave rise to allopoloid lager-brewing yeasts is in concert with previous observations that suggested complex reticulate evolutionary events for many *Saccharomyces* strains used in human-controlled fermentations, including multiple origins of *S. cerevisiae*  $\times$  *S. kudriavzevii* hybrids and the existence of manifold *S. uvarum* strains with introgressions from multiple species (Le Jeune et al. 2007; Peris, Belloch, et al. 2012; Peris, Lopes, Arias et al. 2012; Peris, Lopes, Belloch et al. 2012). Monerawela et al. (2015) recently made a similar argument that the amount of diversity seen between the *S. cerevisiae* portions of the Saaz and Frohberg lineages cannot be accounted for by 500 years of divergence from a single hybridization event at the historical start of lager brewing. Although fungal molecular clocks and the historical record of brewing are open to interpretation, no plausible molecular mechanism is available to explain how the *S. cerevisiae* and *S. eubayanus* subgenomes could have evolved at such different rates at nearly neutral sites once they were present in the same nucleus. Although both the Saaz and Frohberg lineages show similar patterns of relaxed purifying selection, the difference in the amount of

neutral divergences between their subgenomes is incompatible with the hypothesis that the lineages arose from a common origin as the result of a single hybridization event.

The existence of shared translocations between the two lineages must still be reconciled with the multiple origins of the lager-brewing yeast lineages. A variety of plausible models have been proposed to explain how shared translocations could have arisen between independent lineages. Mitotic recombination hotspots or fragile sites have been observed in experiments involving pure and hybrid strains, including lager-brewing yeasts (Dunham et al. 2002; Bond et al. 2004; James et al. 2008; Dunn et al. 2013; Hewitt et al. 2014; Monerawela et al. 2015). Indeed, one of the three shared translocations is near the *MAT* locus, a known target of the *Ho* endonuclease. Alternatively, shared translocations might have occurred in one of the parental strains prior to the hybridizations that produced the modern lager-brewing yeast lineages. Hybridization and introgression between different *Saccharomyces* species is common in strains associated with human-controlled fermentation (Le Jeune et al. 2007; Novo et al. 2009; Libkind et al. 2011; Dunn et al. 2012; Peris, Belloch, et al. 2012; Peris, Lopes, Arias et al. 2012; Peris, Lopes, Belloch et al. 2012; Almeida et al. 2014), so a partially domesticated lineage of *S. eubayanus* could have acquired *S. cerevisiae* genetic material or vice versa. Although the precise mechanism, or mechanisms, for the independent formation of identical translocations remain unknown, the balance of evidence strongly favors multiple origins for lager-brewing yeasts over a single origin.

### The Genomic Response of Lager Yeasts to Domestication

Once hybridization events occurred in the Saaz and Frohberg lineages, both genomes appear to have experienced similar rates of increased evolution. Although positive selection likely contributed to some number of nonsynonymous substitutions, relaxation of purifying selection was probably the main driver of the increased rates of protein sequence evolution in these lineages. The different responses between the

*S. cerevisiae* and *S. eubayanus* subgenomes could reflect the fact that the *S. cerevisiae* parental strains had more previous exposure to brewing environments.

Domestication may have led to a particularly striking relaxation of selection in the case of lager yeasts because hybridization likely prevented the new lineages from mating and recombining with their parental species through conventional meiotic means. During selective sweeps and passaging bottlenecks, the lineages would have been exposed to Muller's Ratchet (Muller 1964; Felsenstein 1974), increasing the average  $\omega$  for the entire genome. Clonal interference is also known to occur in experimentally evolved populations of yeasts (Kao and Sherlock 2008; Lang et al. 2013) and the meiotic counterpart, the Hill–Robertson effect (or interference), is often seen in plant and animal domestication along with elevated evolutionary rates (Hill and Robertson 1966; Doebley et al. 2006; Lu et al. 2006; Cruz et al. 2008; Wang et al. 2014). Many questions remain about the genetic process of domesticating lager yeasts, but the analysis of the near-complete genome sequence of *S. eubayanus* has substantially clarified the origins of the major lineages of the interspecies *S. eubayanus* × *S. cerevisiae* hybrids used to brew lagers and provided a roadmap for future research.

## Materials and Methods

### Strains and Genomes

The genome for *S. cerevisiae* strain S288c was accessed through *Saccharomyces* Genome Database (SGD) (Cherry et al. 2012; Engel et al. 2014, last accessed May 21, 2014). The ale strain genomes, Foster's O and Foster's B, which were deposited by Borneman et al. (2011), were also accessed through SGD. The lager-brewing yeast genomes CBS 1513 (Saaz lineage) and Weinstephan34/70 (WS34/70, W34/70; Frohberg lineage) deposited by Walther et al. (2014) were accessed from DDBJ/EMBL/GenBank (accessions AZCJ01000000 and AZAA01000000, respectively). The assemblies for *S. uvarum* strain CBS 7001 and *S. paradoxus* strain CBS 432 were accessed through [www.saccharomycessensu-tricto.org](http://www.saccharomycessensu-tricto.org) (Liti et al. 2009; Scannell et al. 2011, last accessed May 21, 2014). Mitochondrial sequences for *S. cerevisiae*, *S. paradoxus*, and the Frohberg lineage of lager yeast were accessed through GenBank (accessions NC\_001224, YP\_006460229, and NC0012145, respectively) (Foury et al. 1998; Nakao et al. 2009; Procházka et al. 2012). It is unlikely that using a monosporic derivative of the type strain of *S. eubayanus* (FM1318), thereby removing any heterozygosity in the type strain of *S. eubayanus* (CRUB 1568<sup>T</sup> = PYCC 6148<sup>T</sup> = CBS 12357<sup>T</sup>), had a significant impact on our analyses because the type strain itself has very low heterozygosity (0.0021%) (Hebly et al. 2015).

### Genome Sequencing

The fragment library was prepared according to our previously published Illumina genomic DNA library preparation protocol (Hittinger et al. 2010), with the exception that Illumina paired-end adapters and polymerase chain reaction (PCR) primers were used. From this paired-end library,

35,394,604 Illumina MiSeq 300-bp reads were obtained with an insert size of  $380 \pm 100$  bp. The jumping library was prepared from genomic DNA using Illumina's Nextera Mate Pair Sample Preparation Guide (Illumina Part # 15035209, Rev. C, January 2013) and Nextera Mate Pair Sample Preparation Kit (Illumina, Inc., San Diego, CA) with the following modifications. For each sample, 4.5  $\mu$ g of high molecular weight DNA was introduced to 12  $\mu$ L of tagmentation enzyme and incubated at 55°C for 27 min. After tagmentation, strand displacement and cleanup were performed as described in the protocol, and the DNA was loaded on a 0.6% agarose gel impregnated with GelGreen Nucleic Acid Stain (Biotinium, Inc., Hayward, CA). The DNA was excised from the gel between 6 and 8 kb. Gel cleanup and size selection were verified using an Agilent Bioanalyzer High Sensitivity Chip (Agilent Technologies, Santa Clara, CA). Purified samples were circularized, and the remaining linear DNA was digested as described in the manufacturer's protocol. After circularization, the DNA was sheared to an average size of 400 bp using a Diagenode Bioruptor (Diagenode, Inc., Denville, NJ) for 33 min, with cycles of 30 s on, 30 s off, on high power. Samples were end-repaired, a 3'-A was added to each fragment, Illumina adapters were ligated, and PCR was performed as described in the protocol. The final products were purified using Agencourt AMPure XP beads. The quality and quantity of the finished library were assessed using an Agilent DNA1000 series chip assay (Agilent Technologies) and Invitrogen Qubit HS Kit (Invitrogen, Carlsbad, CA), respectively, and the library was standardized to 2  $\mu$ M. The library was sequenced using an Illumina HiSeq 2000, and images were analyzed using CASAVA version 1.8.2. From this jumping library, 14,144,070 100-bp mate-pair reads were obtained with an apparent insert size of  $6593 \pm 914$  bp.

### Nuclear Genome Assembly

The Illumina mate-pair and paired-end libraries were both trimmed with CUTADAPT v. 1.3 (Martin 2011). Reads below 20 bp were discarded. Reads whose mates were discarded were retained as single-end reads. These processed reads were assembled using ALLPATHS-LG v. r44837 with the following options: frag\_size = 500, frag\_stddev = 100, insert\_size = 3000, insert\_stddev = 300, and PLOIDY = 1. Insert sizes were estimated by ALLPATHS-LG from the data during the assembly process. The ALLPATHS-LG assembly used 79.3% of reads from the fragment library, including 3,251,698 valid pairs, while it used 81.5% of reads from the jumping library, including 5,542,587 valid pairs. We also attempted an assembly by adding an additional jumping library that underwent no size selection, but the scaffold N50 declined from 896 to 764 kb, and several misassemblies were introduced. For these discrepancies, the validity of the original assembly scaffolding was verified by PCR (supplementary fig. S4, Supplementary Material online). In one case, the alternative assembly completely closed a gap, which was verified by PCR; this 2.8-kb segment was the only portion of the second assembly retained in the final assembly. Two adapter

sequence contaminants at the edges of contigs were replaced with Ns.

Because there were so few unplaced scaffolds greater than 1 kb in the initial ALLPATHS-LG assembly, we considered each individually. We deleted four scaffolds that were completely contained elsewhere in the final genome, including two ALLPATHS-LG scaffolds that contained partial mitochondrial genomes redundant with the complete mitochondrial assembly generated by SPAdes (see below). BLAST searches against the genomes of *S. cerevisiae* and our *S. eubayanus* assembly suggested that three scaffolds likely belonged in gaps at the *GAL2*, *PPH22*, and *PEP1* loci. To evaluate these possible placements, we mapped the mate-pair reads using BOWTIE2 with the default settings and compared how many paired mates supported joining all possible pairs of contigs. By ignoring joins supported by fewer than 100 paired mates, strong support was found to place each of these scaffolds (supplementary tables S5 and S6, Supplementary Material online). Based on this evidence, the *PPH22* scaffold was fully placed into a gap, creating two smaller gaps. The *GAL2* scaffold was placed across and fully closed two small gaps—an assembly challenge caused by a recent tandem duplication shared with *S. uvarum* (Hittinger et al. 2004). Partial genes encoding *PEP1/YBL017C* appeared at the edges of two different contigs in the middle of large scaffolds, an apparent segmental duplication of a small portion of chromosome II onto chromosome XIV. The BOWTIE2 mapping provided similar support for placing the unplaced *PEP1* scaffold in both places. Because the BLAST overlap with chromosome II was stronger, we placed the unplaced *PEP1* scaffold there, although the sequence likely is present in both places. Two unplaced scaffolds contain interesting gene clusters (one with *MAL* pseudogenes and one with *FRE* genes involved in iron metabolism), but the BOWTIE2 mapping supported a handful of alternative placement scenarios that we did not investigate further. The remaining four unplaced scaffolds were short (less than 3,500 kb) and only had homology to uncharacterized or dubious open reading frames (ORFs).

### Nuclear Genome Annotation

We annotated the high-quality de novo genome assembly of *S. eubayanus* with a pipeline based on the YGAP (Proux-Wéra et al. 2012). After obtaining a draft annotation from YGAP, we imposed several criteria to do the following: 1) Remove predicted protein-coding genes that did not encode complete ORFs, 2) to enforce GenBank-compatible annotations, and 3) to make orthology assignments explicit when the evidence was unambiguous. YGAP is designed to annotate yeast species that are closely related to *S. cerevisiae* and were similarly derived from a whole-genome duplication event about 100 Ma, as well as more distant relatives that were not subject to this whole-genome duplication. Consequently, YGAP produces a Yeast Gene Order Browser (Proux-Wéra et al. 2012) “pillar” file that does not explicitly assign orthology, but rather lists both ohnologs (paralogs created by the whole-genome duplication) that are present in *S. cerevisiae* as possible annotations. We assigned orthology by imposing a strict

synteny requirement: Orthology was assigned only when the gene with two pillar annotations was adjacent to a gene whose orthology had already been assigned and when synteny was conserved between *S. cerevisiae* and *S. eubayanus*. Only three genes could not be annotated unambiguously by this procedure, all of which are adjacent to translocations or inversions.

We also corrected some annotations to conform to GenBank conventions and standards. Specifically, if a CDS was predicted to continue into a gap, it was marked as partial to exclude the gap from the CDS region. If a CDS had an internal stop codon, its annotation was deleted. Transfer RNA (tRNA) annotations whose cognate amino acid could not be identified were deleted. If a predicted CDS did not begin with a start codon, we scanned upstream until we found a start codon, encountered a stop codon, or reached a gap; the longest possible ORF was then annotated; in some cases, the frame was corrected. For a CDS that did not end in a stop codon, we simply scanned downstream until we reached the nearest in-frame stop codon. We also deleted annotations for all CDS with fewer than ten amino acid residues, duplicate annotations, and annotations for which at least half of the sequence was a gap or ambiguous sequence. We modified three predicted introns to match splice donor and acceptor consensus sites and deleted one predicted intron where no splice donor and acceptor consensus sites could be found. We also deleted two predicted novel genes that completely overlapped with *S. cerevisiae* homologs.

To make the dot plot figure showing the synteny between *S. eubayanus* and *S. cerevisiae* (fig. 1), we concatenated every chromosome sequence together to make one continuous sequence and generated an index file to indicate where each chromosome started and ended. We then used the program LASTZ (Harris 2007) with the default settings to perform the genome alignment and generate a tab-delimited file. We filtered out any alignment less than 1 kb and used “R” to make the dot plot.

### Mitochondrial Genome Assembly and Annotation

To recover the mitochondrial genome, the trimmed paired-end reads were assembled with SPAdes 3.1.1 (Bankevich et al. 2012). Because mtDNA has a higher adenine and thymine (AT) content than nuclear DNA and has multiple copies per cell, we analyzed the distribution of guanine and cytosine (GC) content and the coverage among the resulting contigs with lengths > 5 kb. One contig of about 64 kb was detected with a GC content of 17.5% and 14.6-fold higher coverage than most other contigs, all of which had > 35% GC content (supplementary fig. S5, Supplementary Material online). Because this contig matched the expected size and characteristics of *S. eubayanus* mitochondrial genome, inferred from the Froberg lager yeast mitochondrial genome (Nakao et al. 2009), the contig was selected as the putative mitochondrial genome. We found evidence of 41 *S. cerevisiae* mitochondrial genes using BLASTN (Altschul et al. 1997), including 25 tRNAs and 2 rRNAs. Location of ribosomal RNA (rRNA) genes were confirmed with RNAMmer1.2 (Lagesen et al. 2007). Additional

searches using BLASTX-Q3 were performed using *Saccharomyces paradoxus* protein sequences as queries. The genome arrangement of the genes were plotted and compared with *S. cerevisiae* and Froberg lineage lager-brewing yeast mitochondrial genomes (fig. 2 and supplementary figs. S1 and S2 and supplementary results S1, Supplementary Material online).

### Analysis of Genome Evolution

The sequences used to analyze both nonsynonymous and synonymous changes during genome evolution (fig. 5) were the 1:1:1:1:1 ortholog sets produced using YGAP and the procedure described above, while those used to analyze the rate of synonymous change between the lager lineages (fig. 6) were the 1:1:1:1:1 ortholog sets. To prepare sequences for evolutionary analysis, the translated amino acid sequences of orthologs were first aligned by MUSCLE v.3.8.31 (Edgar 2004). Alignments with PRANK (Löytynoja 2014) produced nearly identical results for the genes *ADR1*, *ERT1*, *FET3*, and *NOT3*. Indels between orthologs were then removed from the original untranslated sequences using PAL2NALv.14 (Suyama et al. 2006). The evolutionary analysis for each set of orthologs was performed using the CODEML package of PAMLv.4.7 (Yang 2007) assuming the F3x4 codon frequency model, and implementing the free ratio branch model. The free ratio branch model assumes an independent  $\omega$  for each branch and does not assume a molecular clock. The number of synonymous and nonsynonymous substitutions was estimated for each gene by multiplying  $d_n$  by the number of nonsynonymous sites in the gene ( $N$ ) and  $d_s$  by the number of synonymous sites in the gene ( $S$ ). We treated each gene as an independent event in order to make the data parametric, and for each gene, we calculated the proportion of nonsynonymous changes out of all substitutions ( $N \times d_n / (S \times d_s + N \times d_n)$ ). The distributions of these values were compared between genomes and subgenomes by logistic regression as implemented in R version 3.1.0. Ninety-nine percent binomial confidence intervals were calculated in the MSTATv.6.1.1 statistical package (<http://mcardle.wisc.edu/mstat/>) to establish the statistical difference in the quantity of synonymous changes between the *S. eubayanus* and *S. cerevisiae* subgenomes of the lager-brewing yeasts (supplementary table S4, Supplementary Material online) and to construct error bars in figure 5. MSTAT was also used to perform Wilcoxon rank-sum tests on the data in figure 6.

### Gene Analyses

For sequences from ale strains and *S. eubayanus*, genome locations were found by BLAST using S288c and, where applicable, known lager yeast gene sequences as queries. The sequences of interest were then extracted manually with custom scripts. BLAST searches were also performed on over 100 publicly available *S. cerevisiae* genomes (Bergström et al. 2014; Strobe et al. 2015), but searches of these genomes failed to further clarify the origins of any lager MAL genes. The percent identity between gene sequences, either amino acid or nucleotide as applicable, was calculated from MUSCLE

alignments using Clustal2.1 accessed from the MUSCLE server (<http://www.ebi.ac.uk/Tools/msa/muscle/>). Gene trees were made in MEGA 6.06 (<http://www.megasoftware.net/>) (Tamura et al. 2013). The parameters to construct each tree were determined by the “Find Best-Fit Substitution Model (ML)” package within MEGA for nucleotide sequences using default parameters. Trees were constructed using the indicated best fitting model and parameters; robustness was assessed using 1,000 bootstrap replicates.

### Supplementary Material

Supplementary results S1–S3, data S1 and S2, figures S1–S5, and tables S1–S6 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

### Acknowledgments

We thank Paula Gonçalves and José Paulo Sampaio for critical comments on the manuscript, as well as Dana A. Oplente for providing guidance for statistical analyses. This work was supported by the USDA National Institute of Food and Agriculture (Hatch project 1003258 to C.T.H.); by the National Science Foundation (grant no. DEB-1253634 to C.T.H.); in part by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494 to C.T.H.); by Project ANPCyT PICT2011-1814 (D.L.); by Project UNComahue B171 (D.L.); and by a NSF-CONICET Argentina bilateral cooperation agreement (5055/14). C.T.H. is an Alfred Toepfer Faculty Fellow, supported by the Alexander von Humboldt Foundation. C.T.H. is a Pew Scholar in the Biomedical Sciences, supported by the Pew Charitable Trusts. The raw Illumina reads for *Saccharomyces eubayanus* FM1318 were deposited in NCBI's SRA as BioProject PRJNA243390. The final genome assembly and annotations were deposited in NCBI's GenBank as accession JMCK00000000.

### References

- Almeida P, Gonçalves C, Teixeira S, Libkind D, Bontrager M, Masneuf-Pomarede I, Albertin W, Durrens P, Sherman DJ, Marullo P, et al. 2014. A Gondwanan imprint on global diversity and domestication of wine and cider yeast *Saccharomyces uvarum*. *Nat Commun*. 5:4044.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25:3389–3402.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 19:455–477.
- Bergström A, Simpson JT, Salinas F, Barré B, Parts L, Zia A, Ba ANN, Moses AM, Louis EJ, Mustonen V, et al. 2014. A high-definition view of functional genetic variation from natural yeast genomes. *Mol Biol Evol*. 31:872–888.
- Bing J, Han PJ, Liu WQ, Wang QM, Bai FY. 2014. Evidence for a Far East Asian origin of lager beer yeast. *Curr Biol*. 24:R380–R381.
- Bond U. 2009. The genomes of lager yeasts. In: Gadd AILSSGM, editor. *Advances in applied microbiology*. Vol. 69. Chapter 6. Waltham (MA): Academic Press. p. 159–182.
- Bond U, Neal C, Donnelly D, James TC. 2004. Aneuploidy and copy number breakpoints in the genome of lager yeasts mapped by microarray hybridisation. *Curr Genet*. 45:360–370.

- Borneman AR, Desany BA, Riches D, Affourtit JP, Forgan AH, Pretorius IS, Egholm M, Chambers PJ. 2011. Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet*. 7:e1001287.
- Borneman AR, Desany BA, Riches D, Affourtit JP, Forgan AH, Pretorius IS, Egholm M, Chambers PJ. 2012. The genome sequence of the wine yeast VIN7 reveals an allotriploid hybrid genome with *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii* origins. *FEMS Yeast Res*. 12:88–96.
- Boulton C, Quain D, editors. 2007. *Brewing yeast and fermentation*. Oxford: Wiley-Blackwell.
- Brown CA, Murray AW, Verstrepen KJ. 2010. Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr Biol*. 20:895–903.
- Casaregola S, Nguyen HV, Lapathitis G, Kotyk A, Gaillardin C. 2001. Analysis of the constitution of the beer yeast genome by PCR, sequencing and subtelomeric sequence hybridization. *Int J Syst Evol Microbiol*. 51:1607–1618.
- Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, et al. 2012. *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res*. 40:D700–D705.
- Collart MA, Panasenko OO, Nikolaev SI. 2013. The Not3/5 subunit of the Ccr4-Not complex: a central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. *Cell Signal*. 25:743–751.
- Cousseau FEM, Alves SL, Trichez D, Stambuk BU. 2013. Characterization of maltotriose transporters from the *Saccharomyces eubayanus* sub-genome of the hybrid *Saccharomyces pastorianus* lager brewing yeast strain Weihenstephan 34/70. *Lett Appl Microbiol*. 56:21–29.
- Cruz F, Vila C, Webster MT. 2008. The legacy of domestication: accumulation of deleterious mutations in the dog genome. *Mol Biol Evol*. 25:2331–2336.
- Day RE, Higgins VJ, Rogers PJ, Dawes IW. 2002. Characterization of the putative maltose transporters encoded by YDL247w and YJR160c. *Yeast* 19:1015–1027.
- Dietvorst J, Londesborough J, Steensma HY. 2005. Maltotriose utilization in lager yeast strains: *MTT1* encodes a maltotriose transporter. *Yeast* 22:775–788.
- Doebly JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* 127:1309–1321.
- Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D. 2002. Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*. 99:16144–16149.
- Dunn B, Paulish T, Stanbery A, Piotrowski J, Koniges G, Kroll E, Louis EJ, Liti G, Sherlock G, Rosenzweig F. 2013. Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. *PLoS Genet*. 9:e1003366.
- Dunn B, Richter C, Kvitek DJ, Pugh T, Sherlock G. 2012. Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res*. 22:908–924.
- Dunn B, Sherlock G. 2008. Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Res*. 18:1610–1623.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792–1797.
- Ellegren H. 2014. Genome sequencing and population genomics in non-model organisms. *Trends Ecol Evol*. 29:51–63.
- Engel SR, Dietrich FS, Fisk DG, Binkley G, Balakrishnan R, Costanzo MC, Dwight SS, Hitz BC, Karra K, Nash RS, et al. 2014. The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3 (Bethesda)* 4:389–398.
- Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737–756.
- Fischer G, James SA, Roberts IN, Oliver SG, Louis EJ. 2000. Chromosomal evolution in *Saccharomyces*. *Nature* 405:451–454.
- Foury F, Roganti T, Lecrenier N, Purnelle B. 1998. The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. *FEBS Lett*. 440:325–331.
- Frederick DL, Tatchell K. 1996. The *REG2* gene of *Saccharomyces cerevisiae* encodes a type 1 protein phosphatase-binding protein that functions with Reg1p and the Snf1 protein kinase to regulate growth. *Mol Cell Biol*. 16:2922–2931.
- Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, et al. 2007. Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 316:222–234.
- Gibson B, Liti G. 2015. *Saccharomyces pastorianus*: genomic insights inspiring innovation for industry. *Yeast* 32:17–27.
- Gibson BR, Storgårds E, Krogerus K, Vidgren V. 2013. Comparative physiology and fermentation performance of Saaz and Froberg lager yeast strains and the parental species *Saccharomyces eubayanus*. *Yeast* 30:255–266.
- Harris RS. 2007. Improved pairwise alignment of genomic DNA. [Ph.D. thesis]. [State College (PA)]: Pennsylvania State University.
- Hebly M, Brickwedde A, Bolat I, Driessen MRM, de Hulster EAF, van den Broek M, Pronk JT, Geertman JM, Daran JM, Daran-Lapujade P. 2015. *Saccharomyces cerevisiae* x *Saccharomyces eubayanus* interspecific hybrid, the best of both worlds and beyond. *FEMS Yeast Res*. 15:fov005.
- Hewitt SK, Donaldson IJ, Lovell SC, Delneri D. 2014. Sequencing and characterisation of rearrangements in three *S. pastorianus* strains reveals the presence of chimeric genes and gives evidence of breakpoint reuse. *PLoS One* 9:e92203.
- Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet Res*. 8:269–294.
- Hittinger CT, Gonçalves P, Sampaio JP, Dover J, Johnston M, Rokas A. 2010. Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature* 464:54–58.
- Hittinger CT, Rokas A, Carroll SB. 2004. Parallel inactivation of multiple GAL pathway genes and ecological diversification in yeasts. *Proc Natl Acad Sci U S A*. 101:14144–14149.
- James TC, Usher J, Campbell S, Bond U. 2008. Lager yeasts possess dynamic genomes that undergo rearrangements and gene amplification in response to stress. *Curr Genet*. 53:139–152.
- Jiang H, Tatchell K, Liu S, Michels CA. 2000. Protein phosphatase type-1 regulatory subunits Reg1p and Reg2p act as signal transducers in the glucose-induced inactivation of maltose permease in *Saccharomyces cerevisiae*. *Mol Gen Genet*. 263:411–422.
- Kao KC, Sherlock G. 2008. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat Genet*. 40:1499–1504.
- Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 35:3100–3108.
- Lang GI, Murray AW, Botstein D. 2009. The cost of gene expression underlies a fitness trade-off in yeast. *Proc Natl Acad Sci U S A*. 106:5755–5760.
- Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, Desai MM. 2013. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* 500:571–574.
- Le Jeune C, Lollier M, Demuyter C, Erny C, Legras JL, Aigle M, Masneuf-Pomarède I. 2007. Characterization of natural hybrids of *Saccharomyces cerevisiae* and *Saccharomyces bayanus* var. *uvarum*. *FEMS Yeast Res*. 7:540–549.
- Libkind D, Hittinger CT, Valério E, Gonçalves C, Dover J, Johnston M, Gonçalves P, Sampaio JP. 2011. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc Natl Acad Sci U S A*. 108:14539–14544.
- Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanou V, et al. 2009. Population genomics of domestic and wild yeasts. *Nature* 458:337–341.
- Liti G, Nguyen Ba AN, Blythe M, Müller CA, Bergström A, Cubillos FA, Dafnis-Calas F, Khoshraftar S, Malla S, Mehta N, et al. 2013. High

- quality *de novo* sequencing and assembly of the *Saccharomyces arboricolus* genome. *BMC Genomics* 14:69.
- Liti G, Peruffo A, James SA, Roberts IN, Louis EJ. 2005. Inferences of evolutionary relationships from a population survey of LTR-retrotransposons and telomeric-associated sequences in the *Saccharomyces sensu stricto* complex. *Yeast* 22:177–192.
- Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. *Methods Mol Biol.* 1079:155–170.
- Lu J, Tang T, Tang H, Huang J, Shi S, Wu CI. 2006. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. *Trends Genet.* 22:126–131.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17:10–12.
- Martini AV, Kurtzman CP. 1985. Deoxyribonucleic acid relatedness among species of the genus *Saccharomyces Sensu Stricto*. *Int J Syst Bacteriol.* 35:508–511.
- Martini AV, Martini A. 1987. Three newly delimited species of *Saccharomyces sensu stricto*. *Antonie Van Leeuwenhoek* 53:77–84.
- Meussdoerffer FG. 2009. A comprehensive history of beer brewing. In: Esslinger HM, editor. *Handbook of brewing processes, technology, markets*. Weinheim (Germany): Wiley-VCH. p. 1–42.
- Monerawela C, James TC, Wolfe KH, Bond U. 2015. Loss of lager specific genes and subtelomeric regions define two different *Saccharomyces cerevisiae* lineages for *Saccharomyces pastorianus* Group I and II strains. *FEMS Yeast Res.* 15:fou008.
- Muller HJ. 1964. The relation of recombination to mutational advance. *Mutat Res.* 106:2–9.
- Nakao Y, Kanamori T, Itoh T, Kodama Y, Rainieri S, Nakamura N, Shimonaga T, Hattori M, Ashikari T. 2009. Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Res.* 16:115–129.
- Naumova ES, Naumov GI, Masneuf-Pomarède I, Aigle M, Dubourdiou D. 2005. Molecular genetic study of introgression between *Saccharomyces bayanus* and *S. cerevisiae*. *Yeast* 22:1099–1115.
- Needleman R. 1991. Control of maltase synthesis in yeast. *Mol Microbiol.* 5:2079–2084.
- Nguyen HV, Gaillardin C. 2005. Evolutionary relationships between the former species *Saccharomyces uvarum* and the hybrids *Saccharomyces bayanus* and *Saccharomyces pastorianus*; reinstatement of *Saccharomyces uvarum* (Beijerinck) as a distinct species. *FEMS Yeast Res.* 5:471–483.
- Nguyen HV, Legras JL, Neuvéglise C, Gaillardin C. 2011. Deciphering the hybridisation history leading to the Lager lineage based on the mosaic genomes of *Saccharomyces bayanus* strains NBRC1948 and CBS380. *PLoS One* 6:e25821.
- Nilsson-Tillgren T, Gjermansen C, Kielland-Brandt MC, Litske Petersen JG, Holmberg S. 1981. Genetic differences between *Saccharomyces carlsbergensis* and *S. cerevisiae*. Analysis of chromosome III by single chromosome transfer. *Carlsberg Res Commun.* 46:65–76.
- Novo M, Bigey F, Beyne E, Galeote V, Gavory F, Mallet S, Cambon B, Legras JL, Wincker P, Casaregola S, et al. 2009. Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proc Natl Acad Sci U S A.* 106:16333–16338.
- Peris D, Belloch C, Lopandić K, Álvarez-Pérez JM, Querol A, Barrio E. 2012. The molecular characterization of new types of *Saccharomyces cerevisiae* × *S. kudriavzevii* hybrid yeasts unveils a high genetic diversity: high genetic diversity in *S. cerevisiae* × *S. kudriavzevii* hybrids. *Yeast* 29:81–91.
- Peris D, Lopes CA, Arias A, Barrio E. 2012. Reconstruction of the evolutionary history of *Saccharomyces cerevisiae* × *S. kudriavzevii* hybrids based on multilocus sequence analysis. *PLoS One* 7:e45527.
- Peris D, Lopes CA, Belloch C, Querol A, Barrio E. 2012. Comparative genomics among *Saccharomyces cerevisiae* × *Saccharomyces kudriavzevii* natural hybrid strains isolated from wine and beer reveals different origins. *BMC Genomics* 13:407.
- Peris D, Sylvester K, Libkind D, Gonçalves P, Sampaio JP, Alexander WG, Hittinger CT. 2014. Population structure and reticulate evolution of *Saccharomyces eubayanus* and its lager-brewing hybrids. *Mol Ecol.* 23:2031–2045.
- Procházka E, Franko F, Poláková S, Sulo P. 2012. A complete sequence of *Saccharomyces paradoxus* mitochondrial genome that restores the respiration in *S. cerevisiae*. *FEMS Yeast Res.* 12:819–830.
- Proux-Wéra E, Armisen D, Byrne KP, Wolfe KH. 2012. A pipeline for automated annotation of yeast genome sequences by a conserved-synteny approach. *BMC Bioinformatics* 13:237.
- Querol A, Bond U. 2009. The complex and dynamic genomes of industrial yeasts. *FEMS Microbiol Lett.* 293:1–10.
- Rainieri S, Kodama Y, Kaneko Y, Mikata K, Nakao Y, Ashikari T. 2006. Pure and mixed genetic lines of *Saccharomyces bayanus* and *Saccharomyces pastorianus* and their contribution to the lager brewing strain genome. *Appl Environ Microbiol.* 72:3968–3974.
- Ribeiro FJ, Przybylski D, Yin S, Sharpe T, Gnerre S, Abouelleil A, Berlin AM, Montmayeur A, Shea TP, Walker BJ, et al. 2012. Finished bacterial genomes from shotgun sequence data. *Genome Res.* 22:2270–2277.
- Riese JC, Eßlinger HM. 2009. World beer market. In: Esslinger HM, editor. *Handbook of brewing processes, technology, markets*. Weinheim (Germany): Wiley-VCH. p. 497–513.
- Rodríguez ME, Pérez-Través L, Sangorrín MP, Barrio E, Lopes CA. 2014. *Saccharomyces eubayanus* and *Saccharomyces uvarum* associated with the fermentation of *Araucaria araucana* seeds in Patagonia. *FEMS Yeast Res.* 14:948–965.
- Rokas A. 2009. The effect of domestication on the fungal proteome. *Trends Genet.* 25:60–63.
- Scannell DR, Zill OA, Rokas A, Payen C, Dunham MJ, Eisen MB, Rine J, Johnston M, Hittinger CT. 2011. The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. *G3 (Bethesda)* 1:11–25.
- Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, Dietrich FS, McCusker JH. 2015. The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res.* 25:762–774.
- Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34:W609–W612.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol.* 30:2725–2729.
- Teste MA, François JM, Parrou JL. 2010. Characterization of a new multi-gene family encoding isomaltases in the yeast *Saccharomyces cerevisiae*, the IMA family. *J Biol Chem.* 285:26815–26824.
- Vidgren V, Multanen JP, Ruohonen L, Londesborough J. 2010. The temperature dependence of maltose transport in ale and lager strains of brewer's yeast. *FEMS Yeast Res.* 10:402–411.
- Walther A, Hesselbart A, Wendland J. 2014. Genome sequence of *Saccharomyces carlsbergensis*, the world's first pure culture lager yeast. *G3 (Bethesda)* 4:783–793.
- Wang GD, Xie HB, Peng MS, Irwin D, Zhang YP. 2014. Domestication genomics: evidence from animals. *Annu Rev Anim Biosci.* 2:65–84.
- Wendland J. 2014. Lager yeast comes of age. *Eukaryot Cell.* 13:1256–1265.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.
- Young ET, Dombek KM, Tachibana C, Ideker T. 2003. Multiple pathways are co-regulated by the protein kinase Snf1 and the transcription factors Adr1 and Cat8. *J Biol Chem.* 278:26146–26158.