Letter

Industrial fuel ethanol yeasts contain adaptive copy number changes in genes involved in vitamin B1 and B6 biosynthesis

Boris U. Stambuk,[1,2,3] Barbara Dunn,[1] Sergio L. Alves, Jr.,[2] Eduarda H. Duval,[2] and Gavin Sherlock[1,3]

[1]Department of Genetics, Stanford University, Stanford, California 94305-5120, USA; [2]Departamento de Bioquímica, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina 88040-900, Brazil

Fuel ethanol is now a global energy commodity that is competitive with gasoline. Using microarray-based comparative genome hybridization (aCGH), we have determined gene copy number variations (CNVs) common to five industrially important fuel ethanol *Saccharomyces cerevisiae* strains responsible for the production of billions of gallons of fuel ethanol per year from sugarcane. These strains have significant amplifications of the telomeric *SNO* and *SNZ* genes, which are involved in the biosynthesis of vitamins B6 (pyridoxine) and B1 (thiamin). We show that increased copy number of these genes confers the ability to grow more efficiently under the repressing effects of thiamin, especially in medium lacking pyridoxine and with high sugar concentrations. These genetic changes have likely been adaptive and selected for in the industrial environment, and may be required for the efficient utilization of biomass-derived sugars from other renewable feedstocks.

[Supplemental material is available online at http://www.genome.org. The microarray data from this study have been submitted to the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession no. GSE13875.]

The use of the budding yeast *Saccharomyces cerevisiae* as a producer of fuel ethanol is becoming increasingly important as concerns regarding the depletion, environmental impact, and security of nonrenewable fossil fuel sources make renewable fuel alternatives highly attractive. Fuel ethanol, which is already produced in large amounts in Brazil and the USA, is by far the most common renewable fuel; it is thus poised to contribute greatly toward goals of energy independence and environmental sustainability (Farrell et al. 2006; Goldemberg 2007). Indeed, Brazilian sugarcane ethanol is now considered a global energy commodity that is fully competitive with gasoline due to several technological achievements in the industrial sector (Amorim 2005; Andrietta et al. 2007).

The fermentation process in Brazil uses very high yeast cell densities in a semicontinuous fed-batch mode to ferment broths (cane juice and/or diluted molasses) containing up to 150–200 g/L of total sugar, producing high ethanol concentrations (9%–12% v/v) with high yield (90%–92% of the theoretical maximum) and productivity (each fermentation cycle lasts 6–10 h, allowing two to three fermentations per day). At the end of each fermentation cycle, the yeast cells are collected by centrifugation, washed in dilute sulfuric acid, then recycled back for each subsequent fermentation during the entire 6- to 9-mo crop season (Andrietta et al. 2007; Basso et al. 2008). Thus, the fermentative capacity and viability of the yeast cells, as well as their ability to withstand stressful industrial conditions—high temperatures and osmotic pressures, high ethanol concentrations, low nutrients and pH, process interruptions, etc.—are important traits required for the efficient production of fuel ethanol.

The predominant microorganisms responsible for the efficient production of fuel ethanol are selected strains of the yeast *S. cerevisiae*. Studies of the microbiological dynamics of the industrial fermentors revealed a very rapid succession of yeast strains during fuel ethanol fermentations, and consequently the original "starter" yeast (usually commercial baker's yeast strains) was completely replaced by other strains in a matter of weeks (Silva-Filho et al. 2005; Basso et al. 2008). A few highly productive yeast strains then tended to dominate the fermentor during the entire production season, allowing efficient and stable fermentations. Some of these selected strains became commercially available in the late 1990s, and, at present, more than half of the distilleries in Brazil use one, or more commonly a mixture of two or more, of these selected strains as starters in the fermentation process, which collectively produce billions of gallons of fuel ethanol per year (Basso et al. 2008). Despite their economic importance as highly efficient fuel ethanol producers, almost nothing is known about the genes, traits, and/or physiological adaptations that allow these selected yeast strains to produce large amounts of ethanol while apparently thriving in the competitive and stressful fermentor environment.

Analysis of genomic variation among a fairly large number of wild-type, laboratory, and industrial *S. cerevisiae* strains has revealed not only significant genetic variability, but has also shown that, in general, these strains cluster according to technological application rather than geographical origin (Fay and Benavides 2005; Legras et al. 2007; Liti et al. 2009; Schacherer et al. 2009). We wished to identify sources of genomic variation in the fuel yeast strains that might be linked to (and possibly responsible for) important industrial traits. We used microarray-based comparative genomic hybridization, also known as aCGH or microarray karyotyping, which has been used previously to distinguish between yeast strains (Dunn et al. 2005; Carreto et al. 2008; for review, see Gresham et al. 2008), to assay the genomes of five industrially important Brazilian *S. cerevisiae* fuel ethanol strains for

[3]Corresponding authors.
E-mail sherlock@genome.stanford.edu; fax (650) 724-3701.
E-mail bstambuk@mbox1.ufsc.br; fax +5548-3721-9672.

copy number variations (CNVs). Our results indicate that relative to the reference laboratory *S. cerevisiae* strain, S288C, all five fuel strains showed significant amplifications of the telomeric *SNO* and *SNZ* genes, which are involved in the biosynthesis of vitamins B6 (pyridoxine) and B1 (thiamin). We further show that increased copy number of these genes confers on the laboratory strain the ability to grow more efficiently under the repressing effect of thiamin, especially in the absence of pyridoxine and in high sugar concentrations. These data suggest that these gene amplifications provide an important adaptive advantage under the industrial fermentation conditions in which the yeast are propagated.

Results

Microarray karyotyping of industrial yeast strains

We performed aCGH analysis of five industrial fuel ethanol yeast strains that are commonly used in Brazil (BG-1, CAT-1, PE-2, SA-1, and VR-1; Supplemental Table S1), assaying two independent isolates from each strain. Visualization of the data within the Java TreeView program (Saldanha 2004) revealed that all yeast strains analyzed showed hybridization patterns with signals across most of the array corresponding to R/G ratios near a value of 1.0, indicating that the genomes of these industrial diploid yeast strains lacked chromosomal aneuploidies (data not shown). The CGH-Miner program (Wang et al. 2005) revealed that the genomes of these five fuel ethanol yeast strains contained 225 genetic elements showing significant CNV with a consensus false discovery rate (FDR) of <0.05 (Fig. 1A). Most CNVs were observed at telomeric or subtelomeric chromosomal regions and generally correspond to

genes that are depleted in copy number relative to the laboratory strain. Clustering all the aCGH data for the fuel yeast strains with data obtained from other representative industrial *S. cerevisiae* strains (a baker's, a wine, and an ale brewer's strain) revealed that the fuel yeast strains cluster together, and are more similar to the baker's yeast strain than to the brewing or wine strains (Fig. 1B).

The significant CNVs identified by the CGH-Miner program were further analyzed, revealing 133 genes (~2% of the genome) that were common to the fuel ethanol yeast strains, defined as those loci where at least half of the fuel ethanol strains contained a significant consensus CNV (see Methods). To determine which of these CNVs might be specific to fuel ethanol yeasts, we then performed aCGH on an expanded group of representative industrial non-fuel yeast strains (four wine, four baker's, and an ale brewing strain; see Supplemental Table S1 for a complete list of strains used). Within this group of nine non-fuel industrial yeast strains, we identified 279 genetic elements showing significant CNV (relative to S288C) with a consensus FDR of <0.05 (Supplemental Fig. S1); of these, 127 are present in at least half of these nine non-fuel industrial strains. Comparison of the CNVs between the two groups of industrial yeast (fuel vs. non-fuel) resulted in the identification of 72 genes that show altered copy number exclusively in the fuel ethanol yeasts, 61 genes with CNVs in both the fuel ethanol and other industrial yeasts, and 66 genes with CNVs exclusive to the other (non-fuel) industrial strains (Fig. 1C; Supplemental Fig. S2). While a number of gene amplifications were found among the fuel-yeast-specific CNVs, the CNVs exclusive to non-fuel industrial strains were mostly either deletions (or copy number depletions) or highly polymorphic genetic elements (i.e., slightly increased or decreased in hybridization, perhaps representing regions of sequence divergence among the strains; Fig. 1B; Supplemental Fig. S2B,C). Of the significant CNV loci common to both groups of industrial strains, most were genes with reduced copy number; for example, there are a large number of Ty transposons present in the S288C genome that are apparently absent from the genomes of all the industrial yeast strains (Supplemental Fig. S2B). Other genes significantly reduced in copy number in both fuel and non-fuel industrial yeasts include several gene families known to be repeated in S288C and that are involved in the transport of hexoses (*HXT6*, *HXT9*, *HXT15*, and *HXT16*), α-glucosides (*MAL11*, *MPH2*, and *MPH3*), or sodium (*ENA1*, *ENA2*, and *ENA5*), as well as asparaginases (*ASP3-1* to *ASP3-4*) and polymorphic aryl-alcohol dehydrogenases (*AAD3*, *AAD4*, *AAD10*, *AAD15*, and *AAD16*) (Fig. 1B; Supplemental Fig. S2). Most of these gene families, as well as Ty transposons, have been previously reported to be missing and/or decreased in copy number in wine and other non-S288C *S. cerevisiae* strains (Dunn et al. 2005; Edwards-Ingram et al. 2007; Carreto et al. 2008). Thus, these gene families probably reflect amplifications that occurred solely in the S288C genome, which was mostly derived from a presumably
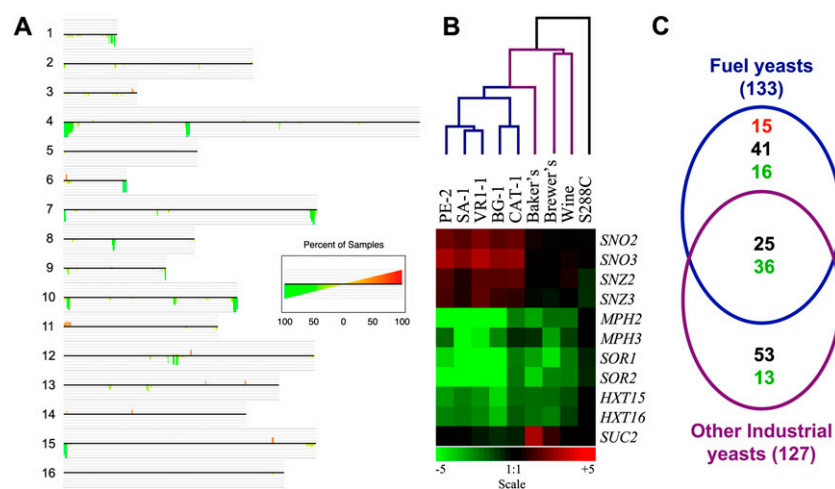

Figure 1. Changes in gene copy number shared by the fuel ethanol yeast strains. (*A*) A consensus karyoscope plot pooling the results of all the fuel ethanol strains relative to the lab strain S288C was generated by the program CGH-Miner (Wang et al. 2005). The plot displays statistically significantly altered copy number variations along each chromosome, and the chromosomes are shown in numerical order from *top* to *bottom*. (Red bars *above* the chromosome) significantly amplified regions, (green bars *below* the chromosome) significantly depleted regions; the height of the bars indicates in how many of the strains the significant observation was made (see reference scale within the figure). (*B*) Clustering dendogram of the entire genome microarray CGH data for the selected fuel yeast strains, together with the same data obtained with a baker's, brewer's, wine, and the lab strain S288C (self–self hybridization) (*top*), and relative gene copy number of selected yeast genes among the different strains (*bottom*). A scale of relative gene copy number is shown at the *bottom*. (*C*) Venn diagram of number of altered gene copy number variations of the selected fuel-ethanol yeast strains that were shared (or not) with a group of nine other industrial yeast strains (four wine, four baker's, and one ale brewing strain). (Red) Amplified genes, (black) polymorphic loci, (green) depleted genes (see Methods).

non-industrial strain isolated from a rotting fig in the 1950s (Mortimer and Johnston 1986).

In contrast to the gene depletions that comprised the majority of CNVs shared by all the industrial strains, the 72 CNVs exclusive to the five fuel ethanol yeasts included many amplified genome regions. Among the most significant gene amplifications identified by CGH-Miner were those of the telomeric *SNO2/SNZ2* and *SNO3/SNZ3* gene pairs; these gene amplifications can be distinctly seen in our aCGH results (Fig. 1; Supplemental Fig. S2A).

Analysis of the *SNO* and *SNZ* loci in the industrial fuel ethanol yeast strains

The *SNO* and *SNZ* genes, which occur in *S. cerevisiae* as divergently transcribed *SNO/SNZ* gene pairs, each encode proteins that are part of a heteromeric enzymatic complex involved in the synthesis of the biologically active form of vitamin B6, pyridoxal 5′-phosphate (PLP), with orthologs in several bacteria and eukaryotes (Fitzpatrick et al. 2007; Neuwirth et al. 2009). In the genome of the sequenced laboratory S288C yeast strain, the *SNO/SNZ* gene family (Rodriguez-Navarro et al. 2002) is composed of three gene pairs: the *SNO1/SNZ1* gene pair (Dong et al. 2004), located in the middle (right arm) of chromosome XIII, and two telomeric gene pairs, *SNO2/SNZ2* (left arm, chromosome XIV) and *SNO3/SNZ3* (left arm, chromosome VI); it is the telomeric versions of the *SNO* and *SNZ* genes that are amplified in the ethanol fuel yeast strains (Fig. 1). These telomeric gene pairs are 99% identical, but exhibit only 72%–80% identity with the *SNO1* and *SNZ1* genes, respectively, indicating that the telomeric gene pairs are the result of a relatively recent gene duplication event. It should be noted that within the S288C genome, the *SNO2/SNZ2* and *SNO3/SNZ3* gene pairs are located just distal to two other genes required for thiamin biosynthesis, *THI5* and *THI12*, respectively. These two *THI* genes are part of a gene family containing yet two other subtelomeric members, *THI11* (right arm, chromosome X) and *THI13* (left arm, chromosome IV), neither of which is associated with a *SNO/SNZ* gene pair; instead, the gene groups *HXT16/SOR1/MPH3* and *HXT15/SOR2/MPH2* are immediately distal to *THI11* and *THI13*, respectively.

Figure 2 shows a chromosomal blot of the telomeric *SNO* loci found in the laboratory S288C strain, as well as in the five fuel ethanol yeast strains. Note that identical results were obtained when a *SNZ* probe (see Methods) was used; we thus presume that the *SNO* genes detected are paired with a corresponding *SNZ* gene. Note also that neither probe detected the more divergent *SNO1/SNZ1* gene pair. While the S288C strain showed the expected *SNO/SNZ* hybridization to only chromosomes VI and XIV, all the industrial fuel ethanol yeasts lacked *SNO/SNZ* genes on chromosome VI (corresponding to the *SNO3/SNZ3* loci in strain S288C), and instead showed *SNO/SNZ* genes located on several other chromosomes, including chromosome XIV (corresponding to the *SNO2/SNZ2* loci in S288C), and new loci probably on chromosomes IX, X, and II (as shown previously for some yeast strains; Padilla et al. 1998). Since most industrial yeast strains lack the two blocks of telomeric *HXT/SOR/MPH* genes found in S288C (Fig. 1B), it is likely that the extra *SNO/SNZ* gene copies arose by replacing the *HXT/SOR/MPH* genes, putting new *SNO/SNZ* gene pairs distal to *THI11* and/or *THI13*. Indeed, we confirmed through PCR that for two of the fuel ethanol strains (strains CAT-1 and PE-2; Supplemental Table S1) a *SNZ* gene is distal to *THI11* on chromosome X (data not shown). Note that there are at least two groups of differing locations of the *SNO/SNZ* gene pairs among the different fuel yeast
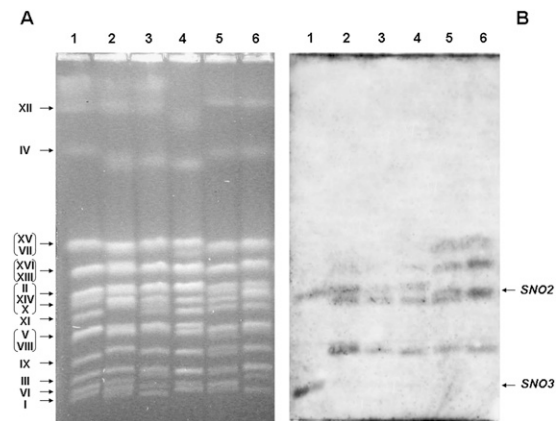

Figure 2. Detection of *SNO2/3* genes in the industrial fuel ethanol yeasts. (*A*) Separation of chromosomes of yeast strains by PFGE; gel stained with ethidium bromide. (*B*) Southern blot of gel shown in panel *A*, hybridized with a probe for *SNO2/3* to detect which chromosomes carry *SNO2* or *SNO3* genes. (Lane *1*) Reference strain S288C, which contains *SNO2* on chromosome XIV and *SNO3* on chromosome VI (also indicated to the *right* of panel *B*); (lanes *2–6*) strains CAT-1, PE-2, VR-1, BG-1, and SA-1, respectively. Roman numerals to the *left* of panel *A* are the chromosome numbers corresponding to the S288C chromosome bands (note that many are seen as doublets).

strains, suggesting that the *SNO/SNZ* amplifications have arisen independently at least twice.

Phenotypic consequence of the amplified *SNO/SNZ* genes

Since genome duplications and gene amplifications are believed to have played a major role in the evolution and success of *S. cerevisiae* as an efficient sugar fermenter (Thomson et al. 2005; Piskur et al. 2006; Conant and Wolfe 2007), we analyzed in more detail the impact of the telomeric *SNO/SNZ* gene amplifications on the physiology of *S. cerevisiae*. Aside from the importance of vitamin B6 itself for amino acid metabolism and many other biochemical pathways, the biologically active form of vitamin B6 (pyridoxal 5′-phosphate, or PLP) serves another role in *S. cerevisiae* as the precursor of thiamin (vitamin B1) (Rodriguez-Navarro et al. 2002; Nosaka 2006). Vitamin B1 forms the cofactor thiamin pyrophosphate, which is essential for sugar utilization by *S. cerevisiae*, in particular for sugar fermentation (Bataillon et al. 1996; Hohmann and Meacock 1998; Mojzita and Hohmann 2006). Indeed, several reports have shown significant up-regulation of thiamin metabolism genes in fermenting yeast cells in various industrial settings including wine, sake, and bread dough (Rossignol et al. 2003; Tanaka et al. 2006; Wu et al. 2006), and in yeast cells cultivated in sugarcane molasses (Shima et al. 2005), indicating that there is a high demand for thiamin (and its precursors) under these industrial conditions.

However, it is well known that for *S. cerevisiae*, as well as for other members of the *Saccharomyces* sensu stricto complex, the presence of thiamin in the medium has a negative effect on the initial growth phase of the yeast cells, although it does not affect the final cell density (Minami et al. 1982; Nakamura et al. 1982; Kamihara and Nakamura 1984). This thiamin-induced growth inhibition has been extensively studied in yeast, and is a consequence of the repressing effect of this vitamin on PLP biosynthesis (Minami et al. 1982; Nakamura et al. 1982; Kamihara and Nakamura 1984; Hohmann and Meacock 1998; Rodriguez-Navarro et al. 2002; Mojzita and Hohmann 2006; Nosaka 2006).

Due to the resultant marked decrease in cellular vitamin B6 (e.g., pyridoxine) content, many metabolic pathways are repressed, including a pronounced lowering of δ-aminolevulinate synthase activity and decreased synthesis of heme, with consequent lowering of respiration and a significant alteration in membrane lipid composition due to inhibition of sterol and unsaturated fatty acids biosynthesis (Minami et al. 1982; Nakamura et al. 1982; Kamihara and Nakamura 1984).

Given the interplay between vitamins B1 (thiamin) and B6 (pyridoxine) in yeast, and their effects on growth and sugar utilization, we hypothesized that the observed *SNO/SNZ* amplifications in the fuel strains would enable them to grow more robustly than strains with fewer *SNO/SNZ* genes under conditions that would normally lead to thiamin-mediated growth repression. We thus compared the growth properties of the laboratory S288C strain and the fuel ethanol yeast strains in defined synthetic medium lacking both vitamins, or in media containing thiamin and lacking pyridoxine (Fig. 3A,B; Supplemental Fig. S3). As had been previously observed, thiamin-induced growth inhibition is clearly evident in the case of the laboratory strain, in which the presence of thiamin in medium lacking pyridoxine causes a severe repression of the specific growth rate (Figs. 3A, 4). This effect is observed even at a very low (~0.15 μM) thiamin concentration (Fig. 4A), a value eight to 20 times lower than the concentration normally found in the most common synthetic defined yeast media. Such growth inhibition at low thiamin concentrations (in the absence of pyridoxine) likely reflects the very high affinity of the yeast *THI7* active thiamin transporter (Enjo et al. 1997), probably leading to increasing intracellular thiamin levels with resultant repression of cellular PLP biosynthesis.

In contrast to the laboratory strain, none of the five fuel ethanol yeast strains analyzed showed the thiamin-repressible growth phenotype; instead, these industrial strains grow at their maximum specific growth rate, independent of the presence thiamin (Fig. 3B; Supplemental Fig. S3). In fact, thiamin had no negative effect on the growth of strain CAT-1 (Fig. 4A), even at a significantly higher (~30 μM) concentration (data not shown). Furthermore, a detailed analysis of the relationship between vitamin supply and sugar utilization by yeast cells revealed that the negative effect of thiamin on the growth rate of the S288C laboratory strain (in media lacking pyridoxine) is even more accentuated at high sugar concentrations (20% glucose), such as those found in molasses and cane juice (Fig. 4B,C; Supplemental Fig. S4). In contrast, the growth performance of the industrial fuel ethanol strain CAT-1 was largely unaffected by the presence or absence of either of these vitamins in the medium, even at a 20% sugar concentration. Competition assays between a CAT-1-derived strain (EDH04, see Methods) and the laboratory diploid S288C strain (GSY157) indeed revealed that the industrial strain quickly overtakes the laboratory strain in high-sugar media containing thiamin and lacking pyridoxine (Fig. 4D); conversely, the laboratory strain out-competes the fuel strain CAT-1 in 2–4 d of growth competition in all other thiamin/pyridoxine combinations (data not shown).

Increased copy number of the *SNO/SNZ* genes confers improved growth under thiamin repressing conditions

To determine whether the additional copies of the *SNO/SNZ* genes in the fuel ethanol yeast strains are directly responsible for their higher specific growth rate in thiamin/sugar conditions that normally repress growth of laboratory and other non-fuel strains, we sought to determine whether the introduction of additional *SNO/SNZ* gene copies into the laboratory strain would confer a similar phenotype. We thus PCR amplified *SNO/SNZ* loci from the two industrial fuel ethanol yeasts strains, CAT-1 and PE-2, as well as from the laboratory S288C yeast strain and cloned the loci into a *URA3*-marked high-copy 2μ-based plasmid. Each of the resulting plasmids was shown to contain a bona fide *SNO–SNZ* divergent gene pair with both full-length ORFs represented; the inserts were sequenced, and a small number of sequence differences were seen among the three strains (Supplemental Fig. S4). A diploid *ura3/ura3* S288C-derived yeast strain was transformed with the plasmids, and three independent transformants were chosen for growth assays in high-sugar medium with or without added thiamin (and lacking pyridoxine). After transformation with the high-copy *SNO/SNZ* plasmids, derived from any of the three strains, the laboratory strain S288C grows as well as the CAT-1 fuel strain under conditions of high sugar concentration in the presence of thiamin and absence of pyridoxine (Fig. 3B,D), indicating that an increased copy number of the telomeric *SNO/SNZ*
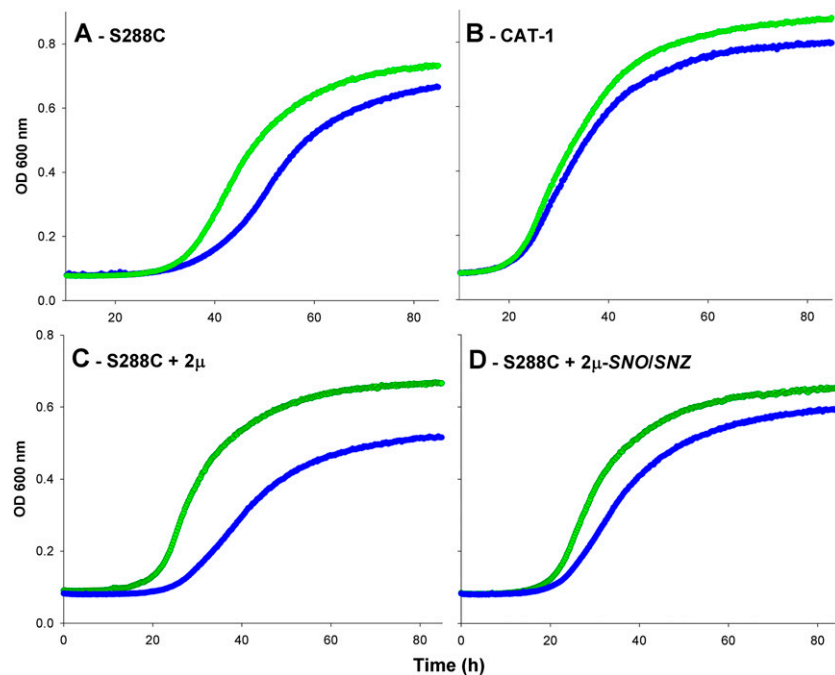

Figure 3. Typical growth curves of the diploid S288C-derived laboratory yeast strain GSY157 (*A*), the industrial fuel ethanol CAT-1 strain (*B*), as well as the diploid laboratory *ura3/ura3* yeast strain FY1679 transformed with a multi-copy 2μ *URA3*+ control (no-insert) plasmid (*C*), or with the same plasmid containing a cloned *SNO/SNZ* gene pair from strain S288C (pBD1001) (*D*), in synthetic defined media (see Methods) containing 200 g/L glucose and lacking both pyridoxine and thiamin (green), or the same media supplemented with 1 mg/L of thiamin (blue).
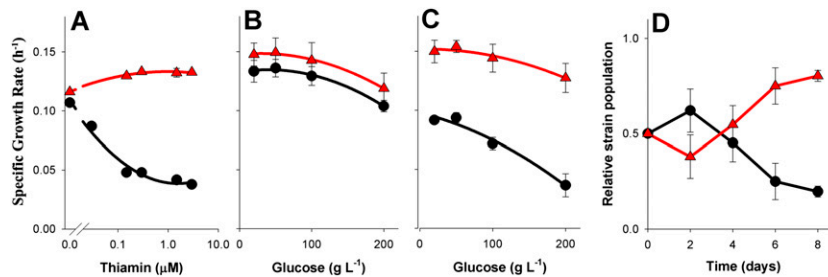
Figure 4. Influence of thiamin (*A*) and sugar concentration (*B,C*) on the specific growth rate (see Methods) of the diploid S288C strain GSY157 (black circles) or the industrial strain CAT-1 (red triangles) in synthetic defined media lacking pyridoxine and thiamin (*B*), or the same media supplemented with thiamin (*A,C*). (*D*) Competition between the diploid S288C strain GSY157 (black circles) and the CAT-1-derived industrial strain EHD04 (red triangles) during sequential growth (see Methods) in synthetic defined media lacking pyridoxine and supplemented with thiamin. Unless otherwise indicated, the concentrations of glucose and thiamin in the medium were 200 g/L and 3 μM (1 mg/L), respectively.

genes is sufficient to confer this phenotype. In contrast, in the control experiment, when transformed with the empty 2μ-based plasmid, no such improvement is seen in S288C (Fig. 3, cf. C and D). Note, although novel alleles of the *SNO* and *SNZ* genes (with one to three amino acid differences from the S288C strain) were found in the genes cloned from the two fuel ethanol strains (Supplemental Figs. S4, S5), transformants with any of the tested telomeric *SNO/SNZ* plasmids showed the same enhanced growth results (Fig. 3D; Supplemental Fig. S3). These results indicate that indeed it is the amplification of these genes, and not novel *SNO/SNZ* alleles, that confers the enhanced growth phenotype.

Discussion

There is increasing evidence that CNVs, and in particular gene amplifications, are an important type of genomic variability that allow adaptation to the selective pressures of changing environments and new ecological niches (Hastings 2007). Examples include increased copy number of the salivary amylase gene as an evolutionary adaptation of humans to starch-rich diets (Perry et al. 2007), duplication of pepsin genes in Antarctic fish as an adaptation for digestion at low temperatures (Carginale et al. 2004), and amplification of genes allowing *Plasmodium* parasites to acquire resistance to antimalarial drugs (Nair et al. 2008; Dharia et al. 2009). In the case of the yeast *S. cerevisiae*, there is evidence for several rounds of evolutionary duplications, starting with the whole-genome duplication (WGD) estimated to have occurred in the ancestor of the *Saccharomyces* sensu stricto species complex ~100 million yr ago (Mya), and leading to the evolution of an efficient fermentative lifestyle in this lineage (Conant and Wolfe 2007; Merico et al. 2007). A subsequent round of more localized amplifications, including regions containing genes specifically involved in hexose transport and thiamin uptake from the medium, have been identified as allowing the successful fermentation of fruit-derived sugars, an evolutionary event that occurred probably some 80 Mya in the Cretaceous age when fleshy fruits first appeared (Thomson et al. 2005).

Significant gene CNVs and other chromosomal rearrangements, such as intra- or interchromosomal translocations and changes in chromosome copy number or ploidy level, have been observed not only in natural *S. cerevisiae* isolates (Dunn et al. 2005; Carreto et al. 2008), but also under controlled laboratory conditions (e.g., Brown et al. 1998; Dunham et al. 2002; Kao and Sherlock 2008). In general, these studies have revealed that geno-

typic and phenotypic evolutionary adaptations depend on the specific selective pressures imposed on the cells. Examples include the increased copy number of the high-affinity *HXT6* and *HXT7* hexose transporters in cells grown in glucose-limited chemostats (Brown et al. 1998; Dunham et al. 2002; Kao and Sherlock 2008), or amplification of the high-affinity *SUL1* sulfate transporter in cells grown under sulfate limitation (Gresham et al. 2008).

Despite the relatively large number of amplified genetic elements found exclusively in the fuel ethanol yeasts (Fig. 1; Supplemental Fig. S2A), unexpectedly, none of these five strains showed amplification of the *SUC2* gene, which encodes the extracellular β-fructosidase (invertase) that is responsible for sucrose hydrolysis and fermentation. Widespread presence of *SUC* genes at multiple telomeric positions has been shown to be a common feature of both baker's and distillers' yeast strains, and is postulated to be an adaptation to sucrose-rich broths (Benitez et al. 1996; Codon et al. 1998). Note that amplification of the *SUC2* gene was indeed observed in the brewing and some baking yeast strains (Fig. 1B; data not shown), but not in the wine strains, and thus this gene was not identified by the CGH-Miner program as significantly altered among the group of non-fuel industrial yeasts. Thus, despite sucrose being the predominant sugar found in sugarcane juice and molasses, these results indicate that invertase activity probably does not limit sucrose fermentation during industrial ethanol production by the fuel yeast strains we studied, and no significant differences in the fermentation of sucrose or glucose were observed between the fuel strains and the lab strains (data not shown).

Our analysis of the genomes of the fuel ethanol yeast strains revealed a significant gene amplification of the telomeric *SNO2/SNZ2* and *SNO3/SNZ3* gene pairs involved in the sequential synthesis of two important vitamins for yeasts, pyridoxine (vitamin B6) and thiamin (vitamin B1), indicating that the fuel ethanol industrial environment imposes a high demand for these cofactors. These results suggest that there must be problems with the bioavailability of these vitamins in the industrial process. In addition, recent data indicate that, contrary to the general view of cofactor recycling during catalysis, the demand for these vitamins might be greater than expected in high-sugar environments because the active form of thiamin (thiamin diphosphate) undergoes a slow but substantial destruction during acetaldehyde (and consequently ethanol) production (McCourt et al. 2006).

Taken together, our data reveal an important set of adaptive gene amplifications present in the fuel ethanol yeast strains. In sugar-rich environments where highly variable substrates (especially in vitamin content) are used—a situation frequently found in industrial fuel ethanol producing units—yeast cells that carry the amplifications of the telomeric *SNO/SNZ* genes would have a clear competitive advantage during the fermentation itself and when being reused for several successive fermentations. Because Brazilian sugarcane biorefineries produce both sugar (sucrose) and ethanol, most factories will first produce sucrose crystals, formed after cane juice concentration and centrifugation, leaving a molasses syrup that is used as the fermentation substrate, usually diluted with water or new cane juice. Besides significant differences

in the raw material (i.e., the type of cane, and processing routines that can vary during the crop season), all treatments used to produce the sugar crystals (heating/boiling, sulfite/acid addition, etc.) are known to affect the stability of vitamins B1 and B6 (Saidi and Warthesen 1983; Mauri et al. 1989). Furthermore, the bio-availability of vitamin B6 for the yeast cells is also of concern, as the form normally found in plants is not readily available for yeast cell usage (Gregory 1997). Not only do the *SNO/SNZ* gene amplifications cause cells to be less sensitive to fluctuations in levels of vitamins B1 and B6 as we have shown here, it should also be emphasized that all of the physiological parameters that have previously been shown to be affected or influenced by the *SNO/SNZ* genes (correct vitamin supply, oxygen content, and membrane lipid and sterol content) have been recently implicated as important for the efficient production of high ethanol concentrations by *S. cerevisiae* (Alfenore et al. 2002, 2004; Cot et al. 2007). While we have observed the *SNO2/SNZ2* and *SNO3/SNZ3* gene amplifications exclusively in the sugarcane industrial fuel ethanol yeasts (of the industrial strains we analyzed), we believe that this genetic adaptation can also be utilized advantageously in the development and engineering of yeast strains for the efficient utilization of biomass-derived sugars from other renewable non-food feedstocks from agriculture and forestry, a prerequisite for the implementation of fuel ethanol on a global scale.

Methods

Yeast strains, media, and growth conditions

Supplemental Table S1 shows a list of the yeast strains analyzed in this study. The five industrial fuel-ethanol yeast strains have been described elsewhere (Amorim 2005; Andrietta et al. 2007; Basso et al. 2008) and have been commercially available since the late 1990s, distributed initially by Lallemand Inc. (http://www.lallemand.com) from Canada, and more recently by LNF Latino Americana Ltda. (http://www.lnf.com.br/) from Brazil. For the competition experiments, strain CAT-1 had one copy of the *SUC2* gene replaced by the kanMX cassette from plasmid pFA6a-kanMX6 as described (Petracek and Longtine 2002) using plasmids SUC2-F1 and SUC2-R1 (Supplemental Table S2). Although deleted in one copy of the invertase gene, the *SUC2/suc2*::kanMX6 strain EHD04 (Supplemental Table S1) had normal invertase activity and fermented sucrose as efficiently as the isogenic CAT-1 strain (data not shown). The diploid laboratory strains FY1679 and GSY157 were *GAL2* and ura⁻ (*ura3-52*) or *URA3* derivatives, respectively, of the widely used laboratory yeast strain S288C (Winston et al. 1995). Strain S288C was the first sequenced *S. cerevisiae* genome, and is also the strain from which the ORFs were obtained through PCR for the microarrays (see below).

Yeast strains were routinely grown on rich YPD medium (1% yeast extract, 2% peptone, and 2% glucose), or synthetic YNB medium (0.67% yeast nitrogen base) containing 2% glucose and supplemented with all auxotrophic requirements (or lacking uracil if uracil drop-out medium was required). To assay the influence of different vitamins on the growth performance of yeast strains, a defined synthetic medium (Verduyn et al. 1992) was used containing per liter of demineralized water: 5 g of $(NH_4)_2SO_4$, 3 g of KH_2PO_4, 0.5 g of $MgSO_4·7H_2O$, 15 mg of EDTA, 4.5 mg of $ZnSO_4·7H_2O$, 0.3 mg of $CoCl_2·6H_2O$, 1 mg of $MnCl_2·4H_2O$, 0.3 mg of $CuSO_4·5H_2O$, 4.5 mg of $CaCl_2·2H_2O$, 3 mg of $FeSO_4·7H_2O$, 0.4 mg of $NaMoO_4·2H_2O$, 1 mg of H_3BO_3, and 0.1 mg of KI. After heat sterilization of the medium for 20 min at 120°C, a filter-sterilized vitamin solution was added in order to achieve a final vitamin concentration of 0.05 mg/L biotin, 1 mg/L calcium panto-

thenate, 1 mg/L nicotinic acid, 25 mg/L inositol, 0.2 mg/L *para*-aminobenzoic acid, and the presence or not of 1 mg/L (4.9 μM) pyridoxine and various concentrations of thiamin as indicated in the main text and figures; glucose was similarly filter sterilized and added to the medium to achieve concentrations of 20–200 g/L. Yeast cells were pre-grown to stationary phase at least two times in each growth medium to be assayed, and 1:100 dilutions of these precultures were used to determine the growth rate at 30°C in 96-well plates in a Tecan GENios microplate reader, each well containing 100 μL of synthetic medium with the indicated vitamins and sugar concentrations. Growth of each culture was monitored by measuring the OD_{600} every 15 min, with high-intensity orbital shaking between measurements. All wells in the plate were tightly sealed with AccuClear Sealing Film for qPCR (E & K Scientific), and thus this growth condition should be considered oxygen-limited (or microaerobic), as no significant growth is observed when fully respiratory substrates (e.g., ethanol or glycerol) were used. All growth curves were performed in triplicate, and the specific growth rate ($μ$, h^{-1}) was determined by the slope of a straight line between the natural log of the OD_{600} and time (h) during the initial exponential phase of growth. For growth competition assays between the laboratory strain GSY157 and EDH04 (CAT-1 *SUC2/suc2*::kanMX6), approximately equal amounts of each strain were inoculated into the synthetic media described above, and a daily 1:100 dilution was reinoculated into new media every 24 h. The percentage of the industrial and the laboratory strain in the media was quantified after appropriate dilutions of the cultures were spread on rich YPD plates either lacking or containing 200 mg/L geneticin (G-418) sulfate.

aCGH protocol

The microarray karyotyping analysis of the industrial yeast strains was performed essentially as described previously (Dunn et al. 2005), using arrays spotted with PCR products corresponding to full-length *S. cerevisiae* strain S288C ORFs (DeRisi et al. 1997). Genomic DNA was isolated with YeaStar columns (Zymo Research) and cut with HaeIII (New England Biolabs); 1 mg was then fluorescently labeled (usually with Cy3-dUTP and Cy5-dUTP [Perkin-Elmer] for the reference and experimental strains, respectively), using the BioPrime random-prime labeling system (Invitrogen). Labeled DNA was purified using Zymo Clean&Concentrate columns (Zymo Research), hybridized to microarrays at 65°C, scanned with an Axon 4000A scanner, and the data were extracted with GenePix (Molecular Devices Corp.). The array data were treated and analyzed as described previously (Dunn et al. 2005); note that all arrays contained duplicated spots for each gene, and all data presented are the averaged values of the duplicate spots. The data for all microarray hybridizations reported in this paper have been deposited in the Stanford Microarray Database (SMD) (Hubble et al. 2009) where they can be retrieved for further analysis, as well as in the Gene Expression Omnibus under accession no. GSE13875.

Array data analysis

Since all array data were normalized by setting the average log ratio of all array elements to a value of zero, differences in hybridization intensity due to ploidy differences among strains (if any) are eliminated. An amplification (three or more gene copies per diploid genome) in the industrial strain is thus expected to have a R/G hybridization ratio of ≥1.5 ($\log_2 ≥ 0.58$) when normalized as described above, while a R/G hybridization ratio of ≤0.5 ($\log_2 ≤ -1.0$) corresponds to less than one gene copy per diploid genome, indicating a gene deletion. To account for technical variation, we did not consider as significant values for a given gene that were within

two standard deviations of the mean calculated from four S288C self–self hybridizations; thus, we only considered data with $\log_2$ values <-0.4 or >0.45. "Polymorphic" genes were defined as those having $\log_2$ R/G values outside this technical variation, and in-between $\log_2$ values of -1.0 and 0.58 (see above); in other words, "polymorphic" genes appear as genes that are "slightly" increased or decreased in copy number, and thus may represent amplified or depleted regions where hybridization is somewhat altered, perhaps to sequence differences among the industrial strains.

In order to identify genes that show a significant consensus CNV relative to the genome of the sequenced reference strain S288C, the CGH-Miner program (Wang et al. 2005) was employed using the default parameters for BAC analysis to improve detection of single probes showing significant CNV (Dunn et al. 2005; Carreto et al. 2008). Four separate S288C self–self hybridizations were used as "normal controls" in the CGH-Miner program, and these were compared with two independent DNA samples (from independent yeast colonies) of each of the five industrial fuel ethanol yeasts (a total of 10 microarrays), or against the microarray data obtained with nine other industrial strains, including brewing (one strain), wine (four strains), and baker's yeasts (four strains). A plot of the genes in industrial strains showing significant altered CNV relative to S288C was generated by the CGH-Miner program (Fig. 1A); this "consensus gene list" was further curated to obtain only CNVs shared by at least three of the five strains in the group of fuel yeasts, or by at least five of the nine strains in the group of other (non-fuel) industrial strains; this curated list was used to generate the Venn diagram in Figure 1C. To determine the relatedness among strains (Fig. 1B), the microarray data were clustered in SMD (Hubble et al. 2009), which uses XCluster, using the Pearson correlation as the metric of similarity.

Cloning telomeric SNO-SNZ genes onto high-copy plasmids, and yeast transformation

Using the S288C sequence from SGD (Cherry et al. 1998), primers SNO/SNZ-For and SNO/SNZ-Rev (Supplemental Table S2) were designed to amplify a 3-kb region spanning the duplicated *SNO-SNZ* divergent gene pairs found close to the left telomeres of either chromosome VI or chromosome XIV. Note that because of the very high sequence identity between these duplicated regions, the sequences are identical at both chromosomal locations, and thus either chromosomal gene pair could be cloned with these primers. Vent polymerase (New England Biolabs) was used to amplify this region from the genomic DNA of three different strains: S288C, CAT-1, and PE-2 (Supplemental Table S1). The fragments obtained from each of the three strains were cloned into the BamHI site of the 2μ-based high-copy plasmid YEplac195 (Gietz and Sugino 1988) containing the selectable *URA3* yeast marker; one plasmid for each of the three strains' inserts was selected for further study. Note that all three plasmids chosen had their inserts in the same orientation, with the PstI site of the *SNO-SNZ* insert further from the endogenous PstI site in the plasmid backbone. The plasmid names (and strains from which their inserts derived) are: pBD1001 (S288C), pBD1002 (PE-2), and pBD1003 (CAT-1). ABI Big-Dye DNA sequencing of each insert was performed using the Seq-SNZ-F*x* and Seq-SNZ-R*x* primers (spaced evenly across the cloned region) (Supplemental Table S2). Using a standard lithium acetate transformation protocol (Gietz and Woods 2006), each plasmid, plus the 2μ vector alone, was introduced into the diploid *ura3-52* S288C-based laboratory strain FY1679 (Supplemental Table S1). Three independent yeast transformants for each plasmid (including the 2μ vector) were chosen for subsequent growth assays (note that all transformants were verified by plasmid loss assays to have been transformed only with plasmids, not by integrative events).

PFGE, chromosome blotting, and hybridization

Yeast chromosomes were prepared as previously described (Guerring et al. 1991) from 1 mL of yeast cells pregrown in YPD medium. The agarose blocks containing the yeast chromosomes were subjected to pulsed-field gel electrophoresis (PFGE) at 10°C using a Gene Navigator pulsed-field system (Pharmacia Biotech) for a total of 27 h at 200 V. The pulse time was stepped from 70 sec after 15 h to 120 sec for 12 h. Following electrophoresis, the chromosomes in the gel were stained with ethidium bromide, photographed, and transferred to a nylon filter (Biodyne A, GIBCO BRL) by capillary blotting. Probes corresponding to nucleotides +54 through +519 relative to the translational start site of the *SNO2/SNO3* genes, or from +495 through +874 relative to the translational start site of the *SNZ2/SNZ3* genes, were generated by PCR using genomic DNA from strain S288C as template and primers SNO2/3-For, SNO2/3-Rev, and SNZ2/3-For and SNO2/3-Rev (Supplemental Table S2), respectively. Labeling of DNA probes and pre-hybridization, hybridization, stringency washes, and chemiluminescent signal generation and detection were performed by using an AlkPhos kit (Amersham Biosciences) as recommended by the manufacturer. After hybridization, an auto-radiography film (Hiperfilm ECL, Kodak) was exposed to the membrane for 2 to 3 h before it was developed. Images were obtained by scanning with an ImageScanner (Amersham Biosciences) and annotated using Microsoft PowerPoint.

References

Alfenore S, Molina-Jouve C, Guillouet S, Uribelarrea J-L, Goma G, Benbadis L. 2002. Improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process. *Appl Microbiol Biotechnol* **60:** 67–72.

Alfenore S, Cameleyre X, Benbadis L, Bideaux C, Uribelarrea J-L, Goma G, Molina-Jouve C, Guillouet SE. 2004. Aeration strategy: A need for very high ethanol performance in *Saccharomyces cerevisiae* fed-batch process. *Appl Microbiol Biotechnol* **63:** 537–540.

Andrietta MGS, Andrietta SR, Steckelbergm C, Stupiellom ENA. 2007. Bioethanol—Brazil, 30 years of Proálcool. *Int Sugar J* **109:** 195–200.

Amorim HV. 2005. *Fermentação alcoólica, ciência & tecnologia*. Fermentec, Piracicaba, São Paulo.

Basso LC, Amorim HV, de Oliveira AJ, Lopes ML. 2008. Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res* **8:** 1155–1163.

Bataillon M, Rico A, Sablayrolles JM, Salmon JM, Barre P. 1996. Early thiamin assimilation by yeasts under enological conditions: Impact on alcoholic fermentation kinetics. *J Ferment Bioeng* **82:** 145–150.

Benitez T, Gasent-Ramirez JM, Castrejon F, Codon AC. 1996. Development of new strains for the food industry. *Biotechnol Prog* **12:** 149–163.

Brown CJ, Todd KM, Rosenzweig RF. 1998. Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Mol Biol Evol* **15:** 931–942.

Carginale V, Trinchella F, Capasso C, Scudiero R, Parisi E. 2004. Gene amplification and cold adaptation of pepsin in Antarctic fish. A possible strategy for food digestion at low temperature. *Gene* **336:** 195–205.

Carreto L, Eiriz MF, Gomes AC, Pereira PM, Schuller D, Santos MA. 2008. Comparative genomics of wild type yeast strains unveils important genome diversity. *BMC Genomics* **9:** 524. doi: 10.1186/1471-2164-9-524.

Cherry JM, Adler C, Ball C, Chervitz SA, Dwight SS, Hester ET, Jia Y, Juvik G, Roe T, Schroeder M, et al. 1998. SGD: *Saccharomyces* Genome Database. *Nucleic Acids Res* **26:** 73–79.

Codon AC, Benitez T, Korhola M. 1998. Chromosomal polymorphism and adaptation to specific industrial environments of *Saccharomyces* strains. *Appl Microbiol Biotechnol* **49:** 154–163.

Conant GC, Wolfe KH. 2007. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol Syst Biol* **3:** 129. doi: 10.1038/msb4100170.

Cot M, Loret MO, François J, Benbadis L. 2007. Physiological behaviour of *Saccharomyces cerevisiae* in aerated fed-batch fermentation for high level production of bioethanol. *FEMS Yeast Res* **7:** 22–32.

DeRisi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278:** 680–686.

Dharia NV, Sidhu AB, Cassera MB, Westenberger SJ, Bopp SE, Eastman RT, Plouffe D, Batalov S, Park DJ, Volkman SK, et al. 2009. Use of high-density tiling microarrays to identify mutations globally and elucidate mechanisms of drug resistance in *Plasmodium falciparum*. *Genome Biol* **10:** R21. doi: 10.1186/gb-2009-10-2-r21.

Dong YX, Sueda S, Nikawa J, Kondo H. 2004. Characterization of the products of the genes *SNO1* and *SNZ1* involved in pyridoxine synthesis in *Saccharomyces cerevisiae*. *Eur J Biochem* **271:** 745–752.

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* **99:** 16144–16149.

Dunn B, Levine RP, Sherlock G. 2005. Microarray karyotyping of commercial wine yeast strains reveals shared, as well as unique, genomic signatures. *BMC Genomics* **6:** 53. doi: 10.1186/1471-2164-6-53.

Edwards-Ingram L, Gitsham P, Burton N, Warhurst G, Clarke I, Hoyle D, Oliver SG, Stateva L. 2007. Genotypic and physiological characterization of *Saccharomyces boulardii*, the probiotic strain of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **73:** 2458–2467.

Enjo F, Nosaka K, Ogata M, Iwashima A, Nishimura H. 1997. Isolation and characterization of a thiamin transport gene, *THI10*, from *Saccharomyces cerevisiae*. *J Biol Chem* **272:** 19165–19170.

Farrell AE, Plevin RJ, Turner BT, Jones AD, O'Hare M, Kammen DM. 2006. Ethanol can contribute to energy and environmental goals. *Science* **311:** 506–508.

Fay JC, Benavides JA. 2005. Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genet* **1:** e5. doi: 10.1371/journal.pgen.0010005.

Fitzpatrick TB, Amrhein N, Kappes B, Macheroux P, Tews I, Raschle T. 2007. Two independent routes of de novo vitamin B6 biosynthesis: Not that different after all. *Biochem J* **407:** 1–13.

Gietz RD, Sugino A. 1988. New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* **74:** 527–534.

Gietz RD, Woods RA. 2006. Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods Mol Biol* **313:** 107–120.

Goldemberg J. 2007. Ethanol for a sustainable energy future. *Science* **315:** 808–810.

Gregory JF 3rd. 1997. Bioavailability of vitamin B-6. *Eur J Clin Nutr* **51:** S43–S48.

Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, DeSevo CG, Botstein D, Dunham MJ. 2008. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet* **4:** e1000303. doi: 10.1371/journal.pgen.1000303.

Guerring SL, Connelly C, Hieter P. 1991. Positional mapping of genes by chromosome blotting and chromosome fragmentation. *Methods Enzymol* **194:** 57–77.

Hastings PJ. 2007. Adaptive amplification. *Crit Rev Biochem Mol Biol* **42:** 271–283.

Hohmann S, Meacock PA. 1998. Thiamin metabolism and thiamin diphosphate-dependent enzymes in the yeast *Saccharomyces cerevisiae*: Genetic regulation. *Biochim Biophys Acta* **1385:** 201–219.

Hubble J, Demeter J, Jin H, Mao M, Nitzberg M, Reddy TB, Wymore F, Zachariah ZK, Sherlock G, Ball CA. 2009. Implementation of GenePattern within the Stanford Microarray Database. *Nucleic Acids Res* **37:** D898–D901.

Kamihara T, Nakamura I. 1984. Regulation of respiration and its related metabolism by vitamin B1 and vitamin B6 in *Saccharomyces* yeasts. *Adv Biochem Eng Biotechnol* **29:** 35–82.

Kao KC, Sherlock G. 2008. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat Genet* **40:** 1499–1504.

Legras JL, Merdinoglu D, Cornuet JM, Karst F. 2007. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* **16:** 2091–2102.

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.

Mauri LM, Alzamora SM, Chirife J, Tomio MJ. 1989. Kinetic parameters for thiamine degradation in foods and model solutions of high water activity. *Int J Food Sci Technol* **24:** 1–9.

McCourt JA, Nixon PF, Duggleby RG. 2006. Thiamin nutrition and catalysis-induced instability of thiamin diphosphate. *Br J Nutr* **96:** 636–638.

Merico A, Sulo P, Piskur J, Compagno C. 2007. Fermentative lifestyle in yeasts belonging to the *Saccharomyces* complex. *FEBS J* **274:** 976–989.

Minami J, Kishi T, Kondo M. 1982. Effects of thiamin on vitamin B6 synthesis in yeasts. *J Gen Microbiol* **128:** 2909–2917.

Mojzita D, Hohmann S. 2006. Pdc2 coordinates expression of the THI regulon in the yeast *Saccharomyces cerevisiae*. *Mol Genet Genomics* **276:** 147–161.

Mortimer RK, Johnston JR. 1986. Genealogy of principal strains of the yeast genetic stock center. *Genetics* **113:** 35–43.

Nair S, Miller B, Barends M, Jaidee A, Patel J, Mayxay M, Newton P, Nosten F, Ferdig MT, Anderson TJ. 2008. Adaptive copy number evolution in malaria parasites. *PLoS Genet* **4:** e1000243. doi: 10.1371/journal.pgen.1000243.

Nakamura I, Ohmura Y, Nagami Y, Kamihara T, Fukui S. 1982. Thiamin accumulation and growth-inhibition in yeasts. *J Gen Microbiol* **128:** 2601–2609.

Neuwirth M, Strohmeier M, Windeisen V, Wallner S, Deller S, Kippe K, Sinning I, Macheroux P, Tews I. 2009. X-ray crystal structure of *Saccharomyces cerevisiae* Pdx1 provides insights into the oligomeric nature of PLP synthases. *FEBS Lett* **583:** 2179–2186.

Nosaka K. 2006. Recent progress in understanding thiamin biosynthesis and its genetic regulation in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **72:** 30–40.

Padilla PA, Fuge EK, Crawford ME, Errett A, Werner-Washburne M. 1998. The highly conserved, coregulated *SNO* and *SNZ* gene families in *Saccharomyces cerevisiae* respond to nutrient limitation. *J Bacteriol* **180:** 5718–5726.

Petracek ME, Longtine MS. 2002. PCR-based engineering of yeast genome. *Methods Enzymol* **350:** 445–469.

Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, et al. 2007. Diet and the evolution of human amylase gene copy number variation. *Nat Genet* **39:** 1256–1260.

Piskur J, Rozpedowska E, Polakova S, Merico A, Compagno C. 2006. How did *Saccharomyces* evolve to become a good brewer? *Trends Genet* **22:** 183–186.

Rodriguez-Navarro S, Llorente B, Rodríguez-Manzaneque MT, Ramne A, Uber G, Marchesan D, Dujon B, Herrero E, Sunnerhagen P, Pérez-Ortín JE. 2002. Functional analysis of yeast gene families involved in metabolism of vitamins B1 and B6. *Yeast* **19:** 1261–1276.

Rossignol T, Dulau L, Julien A, Blondin B. 2003. Genome-wide monitoring of wine yeast gene expression during alcoholic fermentation. *Yeast* **20:** 1369–1385.

Saidi B, Warthesen JJ. 1983. Influence of pH and light on the kinetics of vitamin-B6 degradation. *J Agric Food Chem* **31:** 876–880.

Saldanha AJ. 2004. Java Treeview-extensible visualization of microarray data. *Bioinformatics* **20:** 3246–3248.

Schacherer J, Shapiro JA, Ruderfer DM, Kruglyak L. 2009. Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* **458:** 342–345.

Shima J, Kuwazaki S, Tanaka F, Watanabe H, Yamamoto H, Nakajima R, Tokashiki T, Tamura H. 2005. Identification of genes whose expressions are enhanced or reduced in baker's yeast during fed-batch culture process using molasses medium by DNA microarray analysis. *Int J Food Microbiol* **102:** 63–71.

Silva-Filho EA, dos Santos SKB, Resende AM, de Morais JO, de Morais MA Jr, Simões DA. 2005. Yeast population dynamics of industrial fuel-ethanol fermentation process assessed by PCR-fingerprinting. *Antonie van Leeuwenhoeck* **88:** 13–23.

Tanaka F, Ando A, Nakamura T, Takagi H, Shima J. 2006. Functional genomic analysis of commercial baker's yeast during initial stages of model dough-fermentation. *Food Microbiol* **23:** 717–731.

Thomson JM, Gaucher EA, Burgan MF, De Kee DW, Li T, Aris JP, Benner SA. 2005. Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat Genet* **37:** 630–635.

Verduyn C, Postma E, Scheffers W, van Dijken J. 1992. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8:** 501–517.

Wang P, Kim Y, Pollack J, Narasimhan B, Tibshirani R. 2005. A method for calling gains and losses in array CGH data. *Biostatistics* **6:** 45–58.

Winston F, Dollard C, Ricupero-Hovasse SL. 1995. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. *Yeast* **11:** 53–55.

Wu H, Zheng X, Araki Y, Sahara H, Takagi H, Shimoi H. 2006. Global gene expression analysis of yeast cells during sake brewing. *Appl Environ Microbiol* **72:** 7353–7358.

Industrial fuel ethanol yeasts contain adaptive copy number changes in genes involved in vitamin B1 and B6 biosynthesis

Boris U. Stambuk, Barbara Dunn, Sergio L. Alves, Jr, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2009/10/02/gr.094276.109.DC1
Related Content	**Genome structure of a Saccharomyces cerevisiae strain widely used in bioethanol production** Juan Lucas Argueso, Marcelo F. Carazzolle, Piotr A. Mieczkowski, et al. Genome Res. December , 2009 19: 2258-2270
References	This article cites 63 articles, 10 of which can be accessed free at: http://genome.cshlp.org/content/19/12/2271.full.html#ref-list-1 Articles cited in: http://genome.cshlp.org/content/19/12/2271.full.html#related-urls
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions