FEMS Yeast Research, 21, 2021, foaa071

doi: 10.1093/femsyr/foaa071
Advance Access Publication Date: 8 January 2021
Research Article

RESEARCH ARTICLE

Ethanol production process driving changes on industrial strains

Sheila Tiemi Nagamatsu[1,2,†], Natalia Coutouné[3], Juliana José[2,‡], Mateus Bernabe Fiamenghi[2], Gonçalo Amarante Guimarães Pereira[2,*,§], Juliana Velasco de Castro Oliveira[3] and Marcelo Falsarella Carazzolle[2]

[1]Division of Human Genetics, Department of Psychiatry, Yale School of Medicine, 333 Cedar St, New Haven, CT, 06510, USA, [2]Laboratório de Genômica e BioEnergia, Departamento de Genética, Evolução, Microbiologia e Imunologia, Universidade Estadual de Campinas (UNICAMP), Cidade Universitária Zeferino Vaz, Campinas, São Paulo, 13083-970, Brazil and [3]Laboratório Nacional de Biorrenováveis (LNBR), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), CEP 13083-970, Campinas, São Paulo, Brazil

*Corresponding author: R. Carl Von Linaeus s/n, Cidade Universitária, Zeferino Vaz, 13083-970, Campinas, SP, Brazil. Tel: +55 19 3521-6237; E-mail: goncalo@unicamp.br

One sentence summary: Understanding how the environment can drive changes in industrial yeasts is important for detecting genes that are under selection and may impact the adaptive process.

Editor: Zongbao Zhao

†Sheila Tiemi Nagamatsu, http://orcid.org/0000-0002-4986-5094
‡Juliana José, http://orcid.org/0000-0002-8259-6519
§Gonçalo Amarante Guimarães Pereira, http://orcid.org/0000-0003-4140-3482

ABSTRACT

Ethanol production has key differences between the two largest producing countries of this biofuel, Brazil and the USA, such as feedstock source, sugar concentration and ethanol titers in industrial fermentation. Therefore, it is highly probable that these specificities have led to genome adaptation of the *Saccharomyces cerevisiae* strains employed in each process to tolerate different environments. In order to identify particular adaptations, in this work, we have compared the genomes of industrial yeast strains widely used to produce ethanol from sugarcane, corn and sweet sorghum, and also two laboratory strains as reference. The genes were predicted and then 4524 single-copy orthologous were selected to build the phylogenetic tree. We found that the geographic location and industrial process were shown as the main evolutionary drivers: for sugarcane fermentation, positive selection was identified for metal homeostasis and stress response genes, whereas genes involved in membrane modeling have been connected with corn fermentation. In addition, the corn specialized strain Ethanol Red showed an increased number of copies of MAL31, a gene encoding a maltose transporter. In summary, our work can help to guide new strain chassis selection for engineering strategies, to produce more robust strains for biofuel production and other industrial applications.

Keywords: industrial strain; first-generation ethanol production; sugarcane; corn; comparative genomics; *Saccharomyces cerevisiae*

Received: 1 December 2020; **Accepted:** 4 January 2021

INTRODUCTION

Current changes in the world energy matrix have been repeatedly discussed in United Nations meetings, implicating global warming as the main cause of climate change. In this context, renewable fuel production has been expanded globally to replace petroleum-derived fuels (Lopes *et al.* 2016). Ethanol production from renewable sources is a well-established process widely implemented around the world, which can help in the fight against global warming.

Ethanol production is a diverse field, in which countries have taken advantage of different crops depending on their domestic producing capabilities and available technologies (Santos *et al.* 2016). Ethanol production can be divided into different categories: first-generation (1G) uses feedstock with high starch or sugar content, second-generation (2G) uses lignocellulosic material and solid wastes as a feedstock, and finally, third-generation (3G) uses algae as biomass (Santos *et al.* 2016). The 1G is the process most widely used around the world and the feedstocks are diverse, including sugar-beet, sweet sorghum, cassava, wheat, rice, corn and sugarcane, which shape the features of global processes and specific traits in fermenting strains.

Currently, Brazil and the USA are the largest bioethanol producers in the world, with American production reaching around 15,800 millions of gallons and Brazilian around 8,620 millions of gallons in 2019, while the European Union and China maintain the third and fourth positions with around 1,440 millions of gallons and 900 million, respectively (Renewable Fuels Association 2019). Due to the diversity of feedstock and yeasts' genetic background, the processes are adapted according to the conditions of each country. The main attributes that differentiate Brazilian and American production are (i) raw material: the USA uses starch feedstock from corn, while Brazil mainly uses a sucrose-containing material from sugarcane (Balat and Balat 2009); (ii) the usage of corn starch in the USA requires previous hydrolysis to release soluble sugars combined with simultaneous saccharification and fermentation (SSF) process, while the Brazilian process uses free hexoses in sugarcane juice (Balat and Balat 2009), which exposes the yeast to a high concentration of sugar at the beginning of fermentation; (iii) the Brazilian ethanol production operates with unwashed feedstock, having thus frequent contamination issues (Carvalho-Netto *et al.* 2015); (iv) in Brazil, yeasts employed in fermentation are recycled during all harvest season, wherein yeast cells are separated from wine and treated with sulfuric acid in order to reduce contamination (Lopes *et al.* 2016). The recycling process at the end of a fermentation cycle occurs every 6–15 hours, and the production season may reach a maximum of 240 days a year in Brazil, while in the USA, each cycle lasts more than 48 hours and power plants work all year (Carvalho-Netto *et al.* 2015; Lopes *et al.* 2016). (v) In the US process, due to SSF and prolonged fermentation, the yield of ethanol reaches higher concentration than Brazilian production at the end of the fermentation (Cripwell *et al.* 2019).

As mentioned before, the difference between fermentation processes imposes several challenges and stresses on industrial yeasts. The American process imposes a high concentration of ethanol, while the Brazilian process imposes acid treatment, cell recycling, high temperatures, competition with indigenous yeasts and bacteria, and also osmotic stress, regarding the high concentrations of sugar at the beginning (Basso *et al.* 2008). Due to this stressful process, it was proposed in the literature that industrial fermenters can act as an evolutionary island (Lopes *et al.* 2015), reinforcing the fact that industrial yeasts are more resistant to stress and have a better fermentative performance in relation to laboratory strains. For example, in the Brazilian process, several native Brazilian strains of *Saccharomyces cerevisiae* are widely used in the mills for ethanol production. Many strains were isolated from Brazilian distilleries during the 1990s by different groups (Basso *et al.* 2008), and the strains are known as Barra Grande 1 (BG-1), Santa Adélia 1 (SA-1), Catanduva 1 (CAT-1), Pedra 2 (PE-2) and Vale do Rosário 1 (VR-1), which showed the best features for survival in industrial environments. Different groups have shown that these strains have a remarkable ability to compete with wild yeasts, including other non-*Saccharomyces* genera, surviving and dominating the industrial fermentation process. Other advantages of these strains are: (i) low glycerol and foam formation; (ii) no flocculation; (iii) high yield in ethanol production; (iv) complete fermentation; and (v) ability to maintain viability in cell recycles (Basso *et al.* 2008). For the American process, the industrial strain Ethanol Red (Kumar and Singh 2016) is the most widely employed yeast and is able to reach a final ethanol concentration of up to 20% (v/v) (Mukherjee *et al.* 2017). In summary, the strain phenotype is an essential feature since it directly influences ethanol productivity. This is why exploring the existing natural diversity of strains searching for yeasts with traits that can contribute to a phenotype of tolerance to specific processes during production is essential.

The inevitable change in the world energy matrix makes necessary an increase in the production of renewable fuels with less environmental impact. The development of efficient strains for bioethanol production is directly linked to the strain chassis and its ability to tolerate stress, which can be linked to other strategies such as metabolic engineering and adaptive evolution (Costa *et al.* 2017). Due to exploring and understanding the natural diversity, physiological behavior and genetic basis of ethanol-producing strains are a fundamental part of the process (Basso *et al.* 2008; Della-Bianca *et al.* 2013; Lopes *et al.* 2015). Thus, this work intends to expand our knowledge on the genomics of fermenting strains, comparing them with non-native ethanol strains, in order to identify industrial and specific traits of sugarcane-fermenting and corn ethanol-producing strains, which will bring relevant information for future strain development for even more efficient ethanol production.

MATERIAL AND METHODS

Public genomes of Saccharomyces

The genomes of Ethanol Red (NCBI id: ASM107810v1), CAT-1 (NCBI id: ASM173870v1), VR-1 (NCBI id: ASM173859v1), BG-1 (NCBI id: CTBEBG1), NCIM3186 (NCBI id: ASM102907v1) and SA-1 (NCBI id: SAMN10755908) were downloaded from the NCBI database, while the genomes of S288C, PE-2 (JAY291), CEN.PK2-1C and ZTW1 were downloaded from the Saccharomyces Genome Database (https://downloads.yeastgenome.org/sequence/strains/).

Saccharomyces cerevisiae strains gene prediction

The strains Ethanol Red, BG-1 and NCIM3186 genomic sequences had their coding loci predicted through AUGUSTUS v3.2.3 (Stanke *et al.* 2006) using saccharomyces_cerevisiae_rm11–1a_1 as a model dataset and configured as '–strand = both –gff3 = on'. The quality of the predicted proteins was evaluated using a BUSCO pipeline v2.0 (Simao *et al.* 2015) with the Fungi dataset (*Saccharomycetales.odb9*).

Homology assignment

All proteins from *S. cerevisiae* strains and *Saccharomyces paradoxus* strain UFRJ50816 (NCBI id: ASM207914v1) were assigned to homolog gene groups (orthogroups or gene families) using the Markov clustering algorithm implemented in Orthofinder software v2.2.1 (Emms and Kelly 2015).

Phylogenetic analysis

Two kinds of phylogenetic analyses were performed: phylogenomics to define the strains' relationships and a separate gene tree inference for each gene family with paralogs. Single copy ortholog (SCO) orthogroups were used for the phylogenomics of strains. SCO genes were aligned individually with multiple sequence alignment algorithms implemented in T-coffee v11.00.8cbe486 (Notredame, Higgins and Heringa 2000), setting the parameter '-mode mcoffee'. All SCO alignments were concatenated in a supermatrix for maximum likelihood phylogenetic inference using RAxML v847 with the GTR + GAMMA model of substitutions, and 1000 bootstrap replicates for branch support.

Gene trees for each orthogroup with paralog copies were aligned using T-coffee v11.00.8cbe486 as described above. However, the IQtree software v1.6.6 (Nguyen *et al.* 2015) was used to infer the best-fit substitution model and maximum likelihood phylogeny for these orthogroups.

Analysis of non-synonymous/synonymous rate ratios (dN/dS)

For dN/dS ratio analysis, the coding deoxyribonucleic acid (DNA) sequence for each orthogroup was aligned using MACSE software v1.01 (Ranwez *et al.* 2018) with default parameters in order to maintain codon integrity. Each orthogroup alignment and phylogeny was used to estimate the dN/dS ratio and test evolutionary models in CODEML software from the PAML package v4.3 (Yang 2007). Site models (model 0) and branch-site models (model 2) were tested for all families. Proteins and sites obtained with positive selection were considered as equally selected for all strains. The branch-site analysis used two different sets of strains as foreground: sugarcane-fermenting strains and corn ethanol-producing strains. Proteins and sites with positive selection evidence only for one of the branch-site model foregrounds and not for the site model were considered interesting as putative adaptations for the foreground phenotype. All results were manually curated to select the most promising genes and then they were classified according to ω values and the presence of significant positively selected sites located inside the protein domain.

Copy number variation analysis

Copy number variation (CNV) analysis was performed by applying the method cited by Peter (Peter *et al.* 2018), which uses a ratio between ORF MEDIAN COVERAGE/CHROMOSOME MEDIAN COVERAGE as a measure to estimate copy number. These two metrics were estimated using Picard v1.8 (Broad Institute 2020). We chose 3.0 as a threshold to consider an increase in CNV.

Protein–protein interaction analysis

The protein–protein interaction (PPI) analysis of orthologs was conducted using the STRING database v11.0 (Mering *et al.* 2003). It was set experiments, databases, co-occurrence and co-expression interaction sources with a minimum of 0.9 confidence rate. The Gene Ontology (GO) enrichment was automatically calculated for each network in STRING v11.0.

RESULTS AND DISCUSSION

Feedstock drives evolutionary divergence

A set of publicly available *S. cerevisiae* genomes (industrial strains) was chosen based on their fermentation capability: five sugarcane-fermenting industrial strains from Brazil, one sweet sorghum-fermenting strain from India (NCIM3186), two corn ethanol-producing strains, ZTW1 (China) and Ethanol Red (USA), in addition to two representatives of laboratory strains: S288C and CEN.PK2-1C. The proteins from these 10 *S. cerevisiae* (Table 1) and *S. paradoxus* genomes were grouped into gene families using Orthofinder software v2.2.1 (Emms and Kelly 2015), which identified 6501 orthogroups, 4524 of them being single-copy for all strains. All single copy orthogroups were used to infer a maximum likelihood phylogeny of strains with *S. paradoxus* as an outgroup (Fig. 1).

Our results support that geographic location and niche are the most important features shaping *S. cerevisiae* strains' evolution (Gallone *et al.* 2016; Duan *et al.* 2018; Peter *et al.* 2018), separating the American strains from the European/Asiatic strains. Our phylogenomics analysis reveals a monophyletic clade for sugarcane-fermenting strains, having laboratory strains as a sister clade, and another monophyletic clade composed of the sweet sorghum-fermenting strain and the two corn ethanol-producing strains. Considering geographic location, ZTW1 should be closer to NCIM3186; however, Ethanol Red and ZTW1 were grouped together. This result is suggesting two possibilities: (i) the origin of the Ethanol Red strain is in China; or (ii) in addition to a geographical influence in strains' evolution, there is also an impact of the feedstock from which ethanol is produced.

To identify gene families with positive selection, dN/dS analysis with CODEML (PAML package v4.3) (Yang 2007) was used on single-copy orthogroups. For gene families involving paralogs, a gene tree for each family was inferred and used for further analysis. Branch-site tests involved two different hypotheses reflected in branches selected as foreground: sugarcane-fermenting (BG-1, CAT-1, PE-2, VR-1, and SA-1) vs all others and corn ethanol-producing (ZTW1 and Ethanol Red) vs all others.

Branch-site tests showed 480 families with evidence of positive selection in sugarcane-fermenting yeasts as foreground and 266 in corn ethanol-producing yeasts as foreground, with 31 families detected in both analyses (Table S1, Supporting Information). Gene families with positive selection, for both sugarcane and corn fermenting yeasts, might suggest general adaptations for industrial processes independent of feedstock. Among positively selected gene families, we analyzed carefully those with foreground omega >10 (Table 2), finding metal homeostasis genes in sugarcane-fermenting yeasts, membrane maintenance genes in corn ethanol-producing yeasts and different pathways of stress responses for both groups.

The PPI analysis showed that proteins from genes under positive selection detected in the sugarcane-fermenting analysis (Fig. S1, Supporting Information) have more interactions than the proteins from corn ethanol-producing analysis (Fig. S2, Supporting Information). The functional annotation (Fig. 2) reflected

Table 1. Description of strains and their sequenced genomes. The table shows the main characteristics of the genomes used in the analysis.

Strains	Location	N50 (Kbp)	Total length (Mbp)	Contigs	Predicted genes (>50aa)	Public reads	Reference
PE-2	Brazil	64.3	11.5	453	6.011	Yes	(Argueso *et al.* 2009)
SA-1	Brazil	125.8	11.8	681	5.472	Yes	(Nagamatsu *et al.* 2019)
BG-1	Brazil	337.0	11.7	213	5.597	Yes	(Coutoune *et al.* 2017)
CAT-1	Brazil	911.0	12.1	414	5.616	Yes	(Babrzadeh *et al.* 2012)
VR-1	Brazil	911.9	12.0	299	5.566	Yes	(Gallone *et al.* 2016)
S288C	Laboratory	924.4	12.2	17	6.563	-	SGD
CEN.PK2–1C	Laboratory	115.2	11.6	389	5.359	Yes	(Nijkamp *et al.* 2012)
Ethanol Red	Not defined	95.1	11.5	218	5.386	Yes	(Wallace-Salinas *et al.* 2015)
ZTW1	China	556.9	11.4	33	6.159	-	Unpublished
NCIM3186	India	924.3	11.8	17	5.408	-	(Goud and Ulaganathan 2015)
S. paradoxus	Not defined	872.2	12.2	17	5.525	-	(Yue *et al.* 2017)

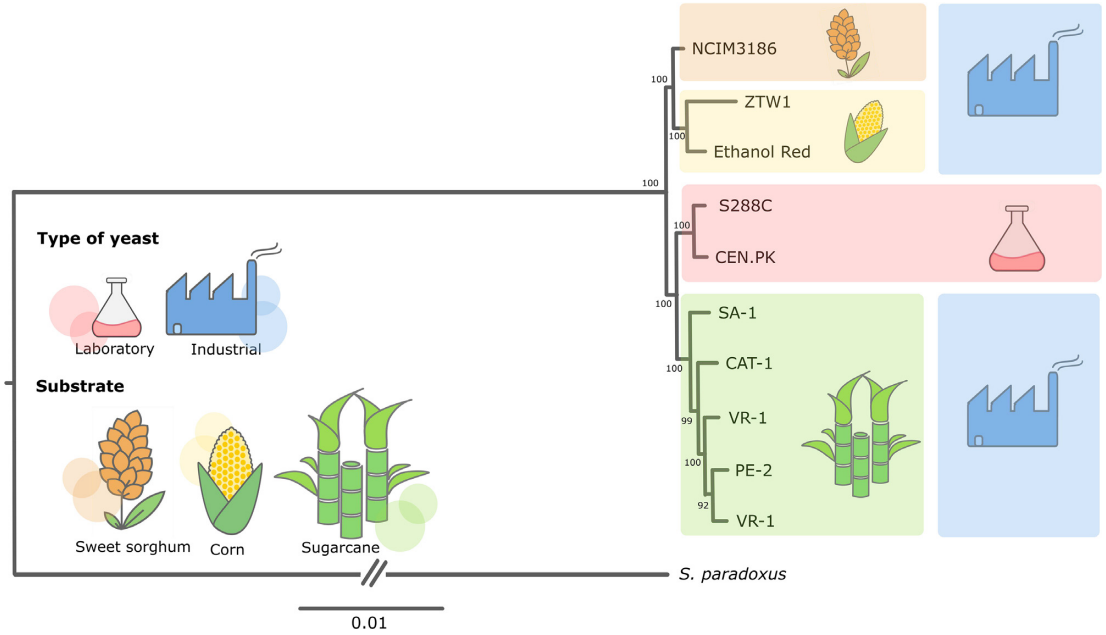

Figure 1. Phylogeny analysis to *S. cerevisiae* strains. The phylogeny shows three clades: **(A)** sugarcane-fermenting strains; **(B)** laboratory strains; and iii) corn ethanol-producing strains and sorghum-fermenting strain.

this result indicating cellular processes and response to stimulus in corn ethanol-producing analysis, while in the sugarcane-fermenting analysis, we detected transcription and messenger ribonucleic acid processing, cellular response to endogenous stimulus, regulation of metabolic processes, biosynthetic processes and other cellular processes. The findings suggest that the sugarcane-fermenting strains are changing their transcriptome and protein pattern. Furthermore, both groups are changing their response to the stimulus.

Sugarcane juice increases selective pressure in genes involved in metal homeostasis

Metals are essential for fermentative metabolism and multiple cell processes since they are used as cofactors of many enzymes. Thus, metal homeostasis must be very well regulated. This process involves mainly two organelles: mitochondria and lysosome-like vacuoles. In general, the excess of metals in the cytoplasm activates transcription factors responsible for regulating genes involved in their capture by both organelles, as well as genes involved in protection against oxidative effects that excess of metals may cause. In contrast, the lack of intracellular metals activates transcription factors that will regulate, usually, genes involved in capturing metals from the extracellular medium.

The sugarcane juice and molasses have been described to have a higher concentration of iron than corn and other plants, reaching around 249 mg/Kg in cane molasses and 138 mg/Kg in stillage from the corn-ethanol process (Wicking and Bian 2015). Interestingly, India uses sugarcane derivatives to produce nutraceuticals to treat iron deficiency anemia (Jain and Venkatasubramanian 2017).

Due to the high iron concentration, we found evidence for positive selection in orthogroups related to metal homeostasis and metal detoxification in sugarcane-fermenting strains. We found a class of metal transporters with evidence of positive selection, including the orthologs/paralogs gene families fam25 (VMR1, YBT1 and NFT1 genes), fam2106 (BPT1 gene) and fam2642 (YPK9 gene), keys in cellular detoxification (Heins-Marroquin *et al.* 2019; Grechko, Podolsky and Cheshchevik 2020).

Table 2. dN/dS to sugarcane- and corn-consumers strains. The table shows interesting orthogroups with foreground omega higher than 10 and detected sites inside domains annotated in the pfam database. The column 'Strains' have two different values: (i) S that means sugarcane-fermenting strains and (ii) C that means corn ethanol-producing strains.

Strains	Orthogroup	Background omega	Foreground omega	Sites detected in domains	Annotation
S	fam16	0.07119	31.64	[503, 504, 506, 507, 550, 675]	Ferric reductases
S	fam25	0.05998	187.79	[386]	Vacuolar membrane proteins
S	fam32	0.09995	288.14	[152, 155, 156]	Heat shock proteins
S	fam50	0.0081	21.10	[547]	Chaperones
S	fam145	0.03729	18.10	[233, 245, 249]	Heat shock proteins
S	fam4698	0.0338	244.09	[429]	Ferric reductase (FRE7)
S	fam1282	0.02201	136.36	[115]	Subunit of trehalose-6-P synthase (TPS2)
S	fam3363	0.03038	15.72	[168]	Copper transport (plasma membrane) (CTR3)
S	fam2642	0.05859	302.14	[409]	Vacuolar protein (YPK9)
S	fam789	0.00208	92.69	[227]	Glutathione oxidoreductase (GLR1)
S	fam4326	0.00929	221.46	[454]	Ferric and cupric reductase (FRE1)
S	fam4621	0.01017	30.37	[316, 317]	Mitochondrial aldehyde dehydrogenase (ALD4)
S	fam2281	0.01729	24.48	[284, 302]	Iron transporter (MRS4)
S	fam789	0.00208	92.69	[227]	Cytosolic and mitochondrial glutathione oxidoreductase (GLR1)
S	fam2106	0.02416	40.18	[437, 1366]	ABC transporter (BPT1)
C	fam3439	0.07114	123.33	[107]	Inositol monophosphatase (INM1)
C	fam4046	0.08823	73.47	[168]	Component of cytosol iron-sulfur assembly (MET18)
C	fam1432	0.08087	137.13	[597]	Chaperone (JEM1)

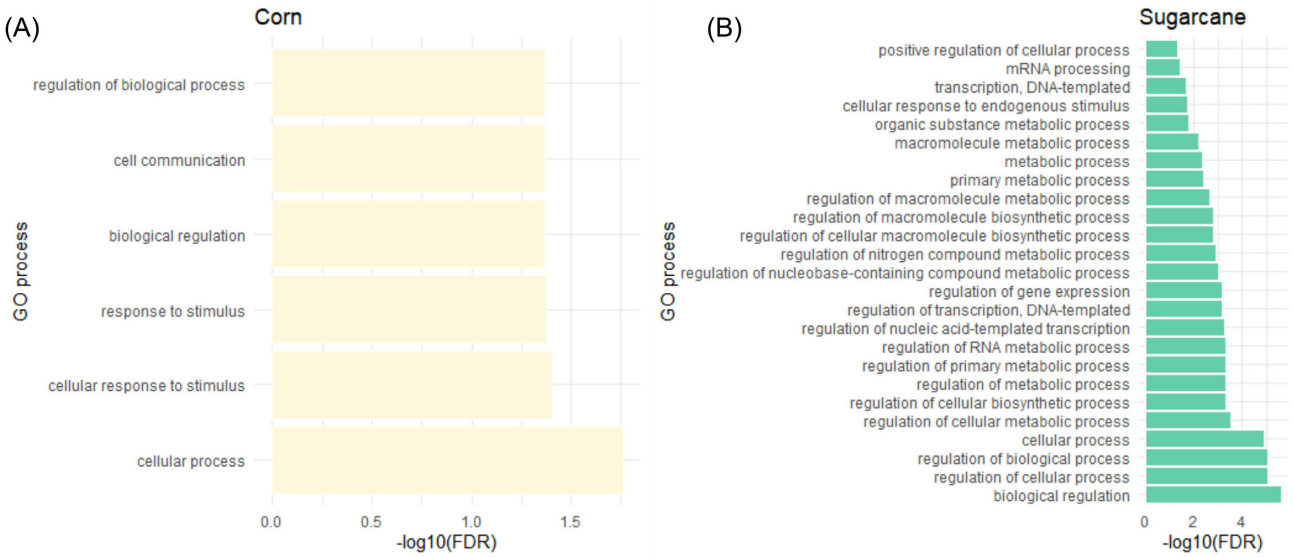

Figure 2. GO enrichment to PPI analysis of sugarcane-fermenting and corn ethanol-producing strains. (A) shows the biological processes (BP) terms enriched in the corn ethanol-producing network, and **(B)** shows the BP terms from the sugarcane-fermenting network.

The genes from fam25 and fam2106 encode proteins involved in vacuolar transport. Both families belong to the ABC superfamily, which is considered an important component of metal homeostasis, acting in the transport of glutathione conjugates with toxic compounds and heavy metals (Sousa, Hanselaer and Soares 2015; Al-Attrache *et al.* 2018; Oestreicher and Morgan 2019). The fam2642 encodes the protein YPK9 (Fig. 3), which is related to the regulation of manganese (Cohen *et al.* 2013), selenium, cadmium and nickel concentrations (Schmidt *et al.* 2009). Moreover, the excess of these elements increases reactive oxygen species (ROS) in yeast cells (Schmidt *et al.* 2009). All these genes with positive selection suggest that the sugarcane strains are under metal exposure, and consequently, homeostasis maintenance genes are under selective pressure.

We also identified positive selection in several genes related to metal starvation. We found evidence of positive selection in the fam2281 composed of the MRS4 gene, an iron transporter placed in the inner membrane of the mitochondria, a

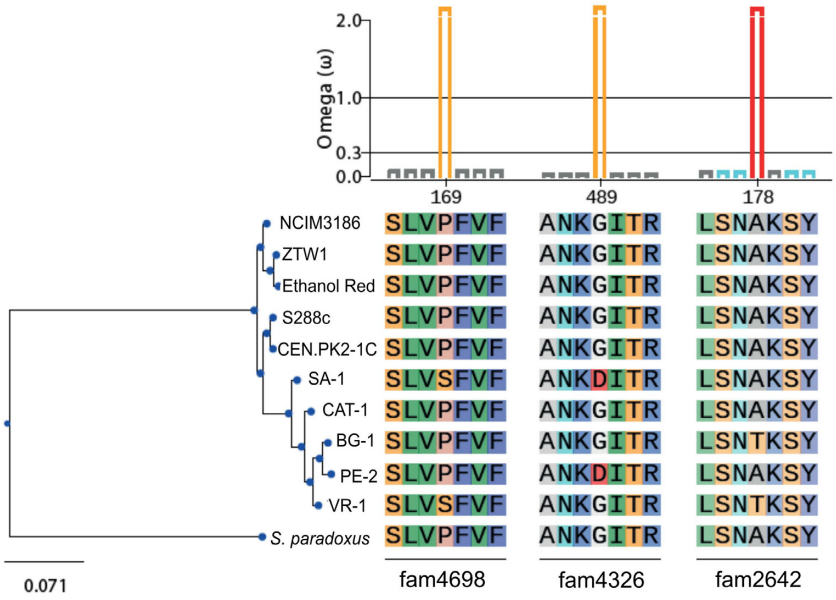

Figure 3. Amino acid positions with significant omega values for three gene families in sugarcane-fermenting strains. The phylogeny obtained for *S. cerevisiae* strains is guiding the alignment order showing the amino acids around positively selected sites. The graph above the alignment shows the omega (dN/dS) values for each amino acid position and the colors refer to *P*-values for each model. Fam4698 represents FRE7, fam4326 represents FRE1 and fam2642 represents YPK9.

fundamental process forming iron-sulfur clusters (Gomez *et al.* 2014). Strikingly, MRS4, and its paralog MRS3, works on iron-limiting conditions, a very different situation that sugarcane-fermenting strains are, which can indicate that these processes are simulating the appearance of new variants (Froschauer, Schweyen and Wiesenberger 2009).

The CTR3 gene (fam3363) is a high-affinity copper transporter located on the cellular membrane. Since the biological importance of the copper acts as a cofactor of many enzymes, it indirectly affects the absorption of iron, mainly through activation of the enzyme SOD1, essential to oxidative stress tolerance (Page *et al.* 2009). Furthermore, we found three orthogroups of metalloreductases, fam16 (FRE2, FRE3 and FRE4), fam4698 (FRE7) and fam4326 (FRE1), with evidence of positive selection. All these proteins are metalloreductases located in the plasma membrane and can reduce ferric iron salts to ferrous form, mostly induced by iron starvation (Yun *et al.* 2001). Also, FRE1 and FRE7 (Fig. 3) can reduce Cu^{2+} to Cu^{+1} at a low intracellular concentration of copper, working coupled with the previously reported CRT1, increasing the SOD expression and consequent antioxidant effect (Berterame *et al.* 2018; Kumar *et al.* 2018). This result suggests a potential coevolution of iron and copper homeostasis genes in order to increase genetic variability in sugarcane-fermenting strains considering that copper is not an abundant element (de Oliveira Lino, Basso and Sommer 2018) in sugarcane juice as iron.

Pathway of stress response mechanism in sugarcane-fermenting strains

The Brazilian fermentation process imposes several challenges and stresses on industrial yeasts, such as osmotic stress and high temperatures. It is known that those stress factors require protein stabilization, folding, and stress response machinery as a cellular defense. Looking at genes involved in those pathways, we found three groups of gene families that encode proteins related to stress response separated by function: chaperones, trehalose biosynthesis and ROS response.

The first group is composed of chaperone families, which play a role in cellular homeostasis preservation during stress events by maintaining protein folding. We found positive selection in three families: (i) the fam32 encodes the DJ-1/ThiJ/PfpI superfamily (HSP31, SNO4, HSP33 and HSP32); (ii) the fam50 encodes the family HSP90 (HSC82 and HSP82); and (iii) fam145 encodes the genes HSP150 and PIR3. The fam32 members are involved in muti-stress cell protection (Natkańska *et al.* 2017); the members of fam50 are chaperones activated for maintenance of folding induced under ethanol and temperature stress in *S. cerevisiae* (Auesukaree 2017); and finally, the fam145 regulates important elements for cell wall stability and integrity, and their expression is triggered by heat shock and oxidative stress (Sanz *et al.* 2018).

The second group of genes is represented by the family fam1282, which encodes a protein related to damage response. The gene TPS2 encodes a trehalose-6-phosphate phosphatase, part of the alpha,alpha-trehalose-phosphate synthase complex. This protein is responsible for the second step of the trehalose synthesis (Singer and Lindquist 1998; Lopes *et al.* 2016). The trehalose is composed of two glucose molecules in a non-reducing carbohydrate. It plays an important role in cell homeostasis against stress, stabilizing proteins and maintaining the integrity of the plasma membrane, triggered by temperature and osmotic stress in *S. cerevisiae* (Magalhães *et al.* 2017).

Finally, the third group of genes is represented by fam4621 (ALD4) and fam789 (GLR1). ALD4 is a mitochondrial aldehyde dehydrogenase isozyme that can use NAD+ and NADP+ as cofactors. It is important to highlight that NADPH is involved in the attenuation of oxidative stress, while NADH has a role in oxidative phosphorylation during energy production (Zhang *et al.* 2018, 2019). Additionally, the GLR1 gene encodes a glutathione oxidoreductase, responsible for regenerating reduced glutathione and keeping cytosolic redox state using NADPH as a cofactor (Kniess and Mayer 2016). Interestingly, these two genes work closely in the cell in order to reduce the stress in the yeast during the fermentation, involved in ROS control, ethanol

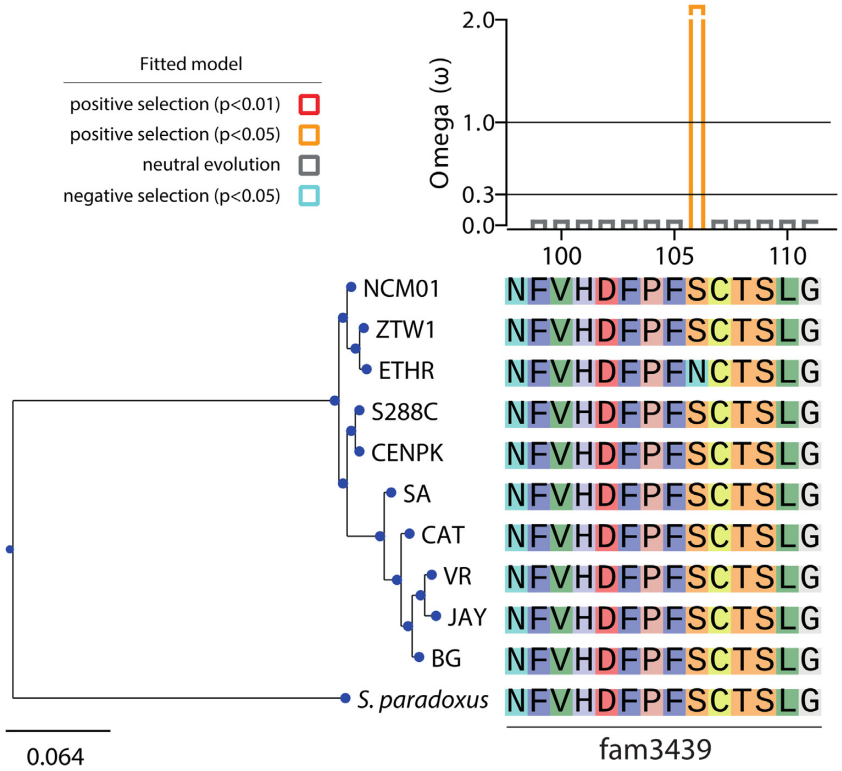

Figure 4. Amino acid positions with significant omega values for two gene families in corn ethanol-producing strains. The phylogeny obtained for *S. cerevisiae* strains is guiding the alignment order showing the amino acids around positively selected sites. The graph above the alignment shows the omega (dN/dS) values for each amino acid position and the colors refer to *P*-values for each model. Fam3439 represents the INM1 gene.

detoxification and osmoregulation (Babazadeh *et al.* 2017; Calabrese, Bruce and Jan 2017).

Corn ethanol-producing process shows positive selection in genes related to membrane maintenance

The corn ethanol production process generates a by-product called dried distiller's grains with solubles (DDGS) used for animal feed. Focusing on increasing productivity of ethanol and DDGS, the process begins with a high initial concentration of milled corn that ferments for more than 48 hours, resulting in a high concentration of ethanol (27.2 g L^{-1}) and DDGS at the end (Mukherjee *et al.* 2017; Cripwell *et al.* 2019). One of the critical mechanisms of tolerance to ethanol is given by alterations in the plasma membrane. Indeed, we found several gene families related to membrane adaptations and vesicular transport with evidence for positive selection in corn ethanol-producing strains. The Ethanol Red strain is well-known for its notable tolerance to high ethanol concentrations, which may cause specific adaptations to this process and the observed resistant phenotype (Mukherjee *et al.* 2017).

For corn ethanol-producing strains, we focused on some families that can be divided into three functional groups according to their annotation: (i) membrane maintenance families, which consists of fam3439 (INM1); (ii) an endoplasmatic reticulum (ER) chaperone: fam1432 (JEM1); and (iii) metal homeostasis-metabolism composed of fam4046 (MET18).

The INM1 (Fig. 4) is an inositol monophosphatase involved in inositol biosynthesis, a precursor of several types of lipids, phosphoinositides and signaling molecules (Murray and Greenberg 2000; Henry, Gaspar and Jesch 2014). The importance of inositol during fermentation can be associated with several crucial functions, such as improving recovery of the membrane fluidity and viability in the ethanol presence (Ishmayana, Kennedy and Learmonth 2015). Also, it is a precursor of phosphatidylinositol 3,5-bisphosphate, working as a biosensor during the osmotic stress response (Malia *et al.* 2018).

The gene JEM1 encodes a homolog of HSP40, a chaperone that belongs to the family DnaJ-like type III and is present in the ER (Nishikawa and Endo 1997; Buck *et al.* 2010). The Jem1p has a major role in nuclear membrane fusion during the mating process. It has been proposed that it is part of a new chaperone complex along with the proteins Kar2p and Kar5p (Brizzio *et al.* 1999; Nishikawa *et al.* 2001; Walsh *et al.* 2004, Berner, Karl-Richard and Dieter 2018). Also, it was shown that the deletion of both ER HSP40 and JEM1 reduces robustness in high temperatures rendering growth at 37°C unfeasible (Nishikawa and Endo 1997).

Finally, MET18 is part of the iron-sulfur assembly machinery in the cytosol and is responsible for transferring iron-sulfur clusters to Fe/S proteins (Chen *et al.* 2017). Several Fe-S proteins play an important role in DNA repair and telomere maintenance and also was described that the MET18 activity is involved in catalase production, helping into the cellular response to oxidant agents, such as H_2O_2 (Puig *et al.* 2017; Chen *et al.* 2017).

Increase in copy number of a maltose transporter gene in Ethanol Red strain

The CNV analysis was made by mapping the available reads against the S288C reference genome and compared the results between sugarcane fermenters, CEN.PK-1C and Ethanol Red. We identified 36 genes with CNV higher than the threshold in at least one strain (Table S2, Supporting Information) and chose some interesting genes to focus our discussion (Table 3).

In some sugarcane-fermenting strains, we detected three genes in tandem in chromosome XI, YKL222C, MCH2 and FRE2,

Table 3 CNV of industrial strains. The table shows the most interesting genes with higher CNV for at least two industrial strains.

Standard Name	Chr	Annotation	BG-1	CAT-1	VR-1	CEN.PK2-1C	Ethanol Red	PE-2	SA1
MAL31	II	Maltose permease	1	1	1	1	6	3	1
YKL222C	XI	YKL222C	2	1	2	1	1	1	2
MCH2	XI	Monocarboxylate permease	3	2	3	1	1	1	2
FRE2	XI	Ferric reductase	2	1	2	1	1	1	2
SAM3	XVI	High-affinity Sadenosylmethionine permease	2	2	2	1	1	2	1
SAM4	XVI	S-adenosylmethioninehomocysteine methyltransferase	2	2	2	1	1	1	1

and other two in chromosome XVI, SAM3 and SAM4, indicating that a duplication of a region larger than a single-copy gene has occurred. Our findings suggest that the region between YKL22C and FRE2 are duplicated and corresponds to 7.7 kb; while the region between SAM3 and SAM4 corresponds to 3.1 kb. The MCH2 gene is annotated as a monocarboxylate permease and was previously described for its importance in the final stages of fermentation where there is an accumulation of ethanol (Legras *et al.* 2016). Also, the FRE2 gene is metalloreductase, activated in low-iron concentration, as in the case of corn stover fermentation (Yun *et al.* 2001). These events reinforce previous works where it has been widely described that stress in yeast can modify the length of telomeres on yeasts, and genes associated with adaptive advantages, over specific pressures, can increase the copy number (Adams *et al.* 1992; Dunn *et al.* 2012; Romano *et al.* 2013). The genes MCH2, SAM3 and SAM4 were already identified with higher copies than S288c in the industrial yeast CAT-1 (Babrzadeh *et al.* 2012).

Strikingly, we also detected in Ethanol Red strain an increase of CNV of the MAL31 gene (six copies) in contrast with PE-2 (three copies) and remaining strains (one copy). The MAL31 gene encodes a maltose permease that allows the uptake of maltose in the cells (Orikasa, Mikumo and Ohwada 2018). This gene is physically located at chromosome II (Fig. S3, Supporting Information), and the high number of copies may increase its expression and improve the ability of Ethanol Red strain to ferment maltose. In fact, Ethanol Red has a high capacity to consume maltose compared with Brazilian industrial strains SA-1 and PE-2 (Junior *et al.* 2012). Overexpression of the MAL31 gene has already been shown to improve the maltose fermentation ratio in *S. cerevisiae* (Hatanaka, Mitsunaga and Fukusaki 2018; Orikasa, Mikumo and Ohwada 2018).

CONCLUSION

This work involved a set of comparative genomics analyses applied to public genomes of industrial *S. cerevisiae* strains widely used to produce ethanol in the world. Until now, this work pioneers in comparing these strains with a focus on different industrial processes and the feedstocks used during fermentation. The Ethanol Red strain, one of the most used corn ethanol-producing yeasts, has never had its genome analyzed before. Our main goal was to verify the similarities and differences at a genomic level, common to the sugarcane-consuming strains compared to strains used in the corn ethanol- and sweet sorghum ethanol-producing mills, and have a deeper understanding of yeast physiology during the ethanol fermentation process. For each gene family, the number of genes grouped for each strain was very similar, suggesting that the divergence is still very recent; therefore, the strains have not yet had time to differentiate between them. Additionally, our phylogenomic analysis using SCO genes corroborates that the geographical location influences strain evolution. However, strain evolution can also be an impacted from the feedstock from which ethanol is being produced, since Ethanol Red (origin not defined) and ZTW1 (China) are grouped together, revealing a possible anthropogenic intervention and artificial selection. The other possibility is that the origin of Ethanol Red is in China. The anthropogenic intervention and artificial selection driven by ethanol production processes are reinforced by the evidence of positive selection in gene families that revealed distinct evolutive pathways between sugarcane-fermenting strains and corn ethanol-producing strains. Among those gene families, we found: (i) metal homeostasis genes in sugarcane-fermenting that were correlated with the well-known high iron concentration in the sugarcane juice; (ii) membrane maintenance genes in corn ethanol-producing yeasts related to the exposure of these strains to high ethanol concentration at the end of fermentation; and (iii) different pathways of stress responses for both. Those results can be employed in future strain selection strategies and genetic engineering studies for supporting the development of new *S. cerevisiae* strains for industrial applications.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSYR* online.

FUNDING

This work was financed by the Center for Computational Engineering and Sciences - FAPESP/Cepid (2013/08293–7); the São Paulo Research Foundation (FAPESP) through grants 2014/26905–2, 2017/00578–3, 2017/02124–0, 2018/06254–8 and 2019/10256–6; the National Council for Scientific and Technological Development (CNPq); and the Coordination for the Improvement of Higher Education Personnel (CAPES)–001.

Conflicts of Interest. None declared.

REFERENCES

Adams J, Puskas-Rozsa S, Simlar J *et al.* Adaptation and major chromosomal changes in populations of *Saccharomyces cerevisiae*. *Curr Genet* 1992;**22**:13–19.

Al-Attrache H, Chamieh H, Hamzé *et al.* N-acetylcysteine potentiates diclofenac toxicity in *Saccharomyces cerevisiae*: stronger potentiation in ABC transporter mutant strains. *Drug Chem Toxicol* 2018;**17**:89–94.

Argueso JL, Carazzolle MF, Mieczkowski PA *et al.* Genome structure of a *Saccharomyces cerevisiae* strain widely sed in bioethanol production. *Genome Res* 2009;**19**:2258–70.

Auesukaree C. Molecular mechanisms of the yeast adaptive response and tolerance to stresses encountered during ethanol fermentation. *J Biosci Bioeng* 2017;**124**:133–42.

Babazadeh R, Lahtvee P, Adiels CB *et al.* The yeast osmostress response is carbon source dependent. *Sci Rep* 2017;**7**:1–11.

Babrzadeh F, Jalili R, Wang C *et al.* Whole-genome sequencing of the efficient industrial fuel-ethanol fermentative *Saccharomyces cerevisiae* strain CAT-1. *Mol Genet Genomics* 2012;**287**:485–94.

Balat M, Balat H. Recent trends in global production and utilization of bio-ethanol fuel. *Appl Energy* 2009;**86**:2273–82.

Basso LC, de Amorim HV, de Oliveira AJ *et al.* Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res* 2008;**8**:1155–63.

Berner N, Karl-Richard R, Dieter HW. Protein quality control of the endoplasmic reticulum and ubiquitin–proteasome-triggered degradation of aberrant proteins: yeast pioneers the path. *Annu Rev Biochem* 2018;**87**:751–82.

Berterame NM, Martania F, Porroa D *et al.* Copper homeostasis as a target to improve *Saccharomyces cerevisiae* tolerance to oxidative stress. *Metab Eng* 2018;**46**:8.

Brizzio V, Khalfan W, Huddler D *et al.* Genetic interactions between KAR7/SEC71, KAR8/JEM1, KAR5, and KAR2 during nuclear fusion in *Saccharomyces cerevisiae*. *Mol Biol Cell* 1999;**10**:609–26.

Broad Institute. *Picard*. https ://broadinstitute.github.io/picard/. (12 September 2020, date last accessed).

Buck TM, Kolb AR, Boyd CR *et al.* The endoplasmic reticulum-associated degradation of the epithelial sodium channel requires a unique complement of molecular chaperones. *Mol Biol Cell* 2010;**21**:1047–58.

Calabrese G, Bruce M, Jan R. Mitochondrial glutathione: regulation and functions. *Antioxid Redox Signal* 2017;**27**:1162–77.

Carvalho-Netto OV, Carazzolle MF, Mofatto LS *et al.* *Saccharomyces cerevisiae* transcriptional reprograming due to bacterial contamination during industrial scale bioethanol production. *Microb Cell Fact* 2015;**14**:13.

Chen Y-q, Liu X-g, Zhao W *et al.* MET18 deficiency increases the sensitivity of yeast to oxidative stress and shortens replicative lifespan by inhibiting catalase activity. *Biomed Res Int* 2017;**2017**:8.

Cohen Y, Megyeri M, Chen OC *et al.* The yeast p5 type ATPase, spf1, regulates manganese transport into the endoplasmic reticulum. *PLoS One* 2013;**8**:e85519.

Costa CE, Romaní A, Cunha JT, et al. Integrated approach for selecting efficient Saccharomyces cerevisiae for industrial lignocellulosic fermentations: importance of yeast chassis linked to process conditions. *Bioresource technology* 2017;**227**:24–34.

Coutoune N, Mulato AT, Riano-Pachon DM *et al.* Draft genome sequence of *Saccharomyces cerevisiae* Barra Grande (BG-1), a Brazilian industrial bioethanol-producing strain. *Genome Announc* 2017;**5**:e00111–17.

Cripwell RA, Rose SH, Favaro L *et al.* Construction of industrial *Saccharomyces cerevisiae* strains for the efficient consolidated bioprocessing of raw starch. *Biotechnol Biofuels* 2019;**12**:201.

Della-Bianca BE, Basso TO, Stambuk BU *et al.* What do we know about the yeast strains from the Brazilian fuel ethanol industry? *Appl Microbiol Biotechnol* 2013;**97**:979–91.

de Oliveira Lino FS, Basso TO, Sommer MOA. A synthetic medium to simulate sugarcane molasses. *Biotechnol Biofuels* 2018;**11**:1–12.

Duan SF, Han PJ, Wang QM *et al.* The origin and adaptive evolution of domesticated populations of yeast from Far East Asia. *Nat Commun* 2018;**9**:2690.

Dunn B, Richter C, Kvitek DJ *et al.* 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* 2012;**22**:908–24.

Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 2015;**16**:157.

Froschauer EM, Schweyen RJ, Wiesenberger G. The yeast mitochondrial carrier proteins Mrs3p/Mrs4p mediate iron transport across the inner mitochondrial membrane. *Biochim Biophys Acta* 2009;**1788**:1044–50.

Gallone B, Steensels J, Prahl T *et al.* Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. *Cell* 2016;**166**:1397–410.

Gomez M, Perez-Gallardo RV, Sanchez LA *et al.* Malfunctioning of the iron-sulfur cluster assembly machinery in*Saccharomyces cerevisiae* produces oxidative stress via an iron-dependent mechanism, causing dysfunction in respiratory complexes. *PLoS One* 2014;**9**:e111585.

Goud BS, Ulaganathan K. Draft genome sequence of *Saccharomyces cerevisiae* strain NCIM3186 used in the production of bioethanol from sweet sorghum. *Genome Announc* 2015;**3**, e00813–15.

Grechko VM, Podolsky DE, Cheshchevik VT. Identification new potential multidrug resistance proteins of *Saccharomyces cerevisiae*. *J Microbiol Methods* 2020;**176**:106029.

Hatanaka H, Mitsunaga H, Fukusaki E. Inhibition of*Saccharomyces cerevisiae* growth by simultaneous uptake of glucose and maltose. *J Biosci Bioeng* 2018;**125**:52–58.

Heins-Marroquin U, Jung PP, Cordero-Maldonado ML *et al.* Phenotypic assays in yeast and zebrafish reveal drugs that rescue ATP13A2 deficiency. *Brain Commun* 2019;**1**:fcz019.

Henry SA, Gaspar ML, Jesch SA. The response to inositol: regulation of glycerolipid metabolism and stress response signaling in yeast. *Chem Phys Lipids* 2014;**180**:23–43.

Ishmayana S, Kennedy UJ, Learmonth RP. Preliminary evidence of inositol supplementation effect on cell growth, viability and plasma membrane fluidity of the yeast *Saccharomyces cerevisiae*. *Procedia Chemistry* 2015;**17**:162–9.

Jain R, Venkatasubramanian P. Sugarcane molasses - a potential dietary supplement in the management of iron deficiency anemia. *J Diet Suppl* 2017;**14**:589–98.

Junior MM, Oliveira JEd, Batistote M *et al.* Evaluation of Brazilian ethanol production yeasts for maltose fermentation in media containing structurally complex nitrogen sources. *J Inst Brew* 2012;**118**:82–88.

Kniess RA, Mayer MP. The oxidation state of the cytoplasmic glutathione redox system does not correlate with replicative lifespan in yeast. *NPJ Aging Mech Dis* 2016;**2**:16028.

Kumar D, Singh V. Dry-grind processing using amylase corn and superior yeast to reduce the exogenous enzyme requirements in bioethanol production. *Biotechnol Biofuels* 2016;**9**:228.

Kumar P, Awasthi A, Nain V *et al.* Novel insights into TOR signalling in *Saccharomyces cerevisiae* through Torin2. *Gene* 2018;**669**:12.

Legras J, Moreno-Garcia J, Zara S *et al.* Flor yeast: new perspectives beyond wine aging. *Frontiers Microbiol* 2016;**7**:503.

Lopes ML, Paulillo SCdL, Cherubin RA *et al.* *Tailored Yeasts Strains for Ethanol Production: Process-Driven Selection*. Fermentec 2015. in https://www.academia.edu/40062698.

Lopes ML, Paulillo SCdL, Godoy A *et al*. Ethanol production in Brazil: a bridge between science and industry. *Braz J Microbiol* 2016;**47**:64–76.

Magalhães RSS, Lima KCD, Almeida DSGd *et al*. Trehalose-6-phosphate as a potential lead candidate for the development of Tps1 inhibitors: insights from the trehalose biosynthesis pathway in diverse yeast species. *Appl Biochem Biotechnol* 2017;**181**:11.

Malia PC, Numrich J, Nishimura T *et al*. Control of vacuole membrane homeostasis by a resident PI-3,5-kinase inhibitor. *Proc Natl Acad Sci USA* 2018;**115**:4684–9.

Mering Cv, Huynen M, Jaeggi D *et al*. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* 2003;**31**:258–61.

Mukherjee V, Radecka D, Aerts G *et al*. Phenotypic landscape of non-conventional yeast species for different stress tolerance traits desirable in bioethanol fermentation. *Biotechnol Biofuels* 2017;**10**:216.

Murray M, Greenberg ML. Expression of yeast INM1 encoding inositol monophosphatase is regulated by inositol, carbon source and growth stage and is decreased by lithium and valproate. *Mol Microbiol* 2000;**36**:651–61.

Nagamatsu ST, Teixeira GS, de Mello F *et al*. Genome assembly of a highly aldehyde-resistant*Saccharomyces cerevisiae* SA1-derived industrial strain. *Microbiol Resour Announc* 2019;**8**:e00071–19.

Natkańska U, Skoneczna A, Sieńko M *et al*. The budding yeast orthologue of Parkinson's disease-associated DJ-1 is amultistress response protein protecting cells against toxicglycolytic products. *Biochim Biophys Acta* 2017;**1864**:39–50.

Nguyen LT, Schmidt HA, von Haeseler A *et al*. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74.

Nijkamp JF, van den Broek M, Datema E *et al*. De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D, a model for modern industrial biotechnology. *Microb Cell Fact* 2012;**11**:36.

Nishikawa S, Endo T. The yeast JEM1p Is a DnaJ-like protein of the endoplasmic reticulum membrane required for nuclear fusion. *J Biol Chem* 1997;**272**:5.

Nishikawa SI, Fewell SW, Kato Y *et al*. Molecular Chaperones in the Yeast Endoplasmic Reticulum Maintain the Solubility of Proteins for Retrotranslocation and Degradation. *J Cell Biol* 2001;**153**:1061–70.

Notredame C, Higgins DG, Heringa J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 2000;**302**:205–17.

Oestreicher J, Morgan B. Glutathione: subcellular distribution and membrane transport (1). *Biochem Cell Biol* 2019;**97**:270–89.

Orikasa Y, Mikumo D, Ohwada T. A 2-Deoxyglucose-resistant mutant of *Saccharomyces cerevisiae* shows enhanced maltose fermentative ability by the activation of MAL genes. *Foods* 2018;**7**:52.

Page MD, Kropat J, Hamel PP *et al*. Two chlamydomonas CTR copper transporters with a novel cys-met motif are localized to the plasma membrane and function in copper assimilation. *Plant Cell* 2009;**21**:928–43.

Peter J, De Chiara M, Friedrich A *et al*. Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* 2018;**556**:339–44.

Puig S, Ramos-Alonso L, Romero AM *et al*. The elemental role of iron in DNA synthesis and repair. *Metallomics* 2017;**9**:1483–500.

Ranwez V, Douzery EJP, Cambon C *et al*. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Mol Biol Evol* 2018;**35**:2582–84.

Renewable Fuels Association. *Annual World Fuel Ethanol Production* 2019. https://ethanolrfa.org/statistics/annual-ethanol-production/(1 December 2020, date last accessed).

Romano GH, Harari Y, Yehuda T *et al*. Environmental stresses disrupt telomere length homeostasis. *PLos Genet* 2013;**9**:e1003721.

Santos LVd, Grassi MCdB, Gallardo JCM *et al*. Second-generation ethanol: the need is becoming a reality. *Ind Biotechnol* 2016;**12**:18.

Sanz AB, García R, Rodríguez-Peña JM *et al*. The CWI pathway: regulation of the transcriptional adaptive response to cell wall stress in yeast. *J Fungi* 2018;**4**:1.

Schmidt K, Wolfe DM, Stiller B *et al*. Cd2+, Mn2+, Ni2+ and Se2+ toxicity to *Saccharomyces cerevisiae* lacking YPK9p the orthologue of human ATP13A2. *Biochem Biophys Res Commun* 2009;**383**:198–202.

Simao FA, Waterhouse RM, Ioannidis P *et al*. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015;**31**:3210–2.

Singer MA, Lindquist S. Multiple effects of trehalose on protein folding in vitro and in vivo. *Mol Cell* 1998;**1**:639–48.

Sousa CA, Hanselaer S, Soares EV. ABCC subfamily vacuolar transporters are involved in Pb (lead) detoxification in*Saccharomyces cerevisiae*. *Appl Biochem Biotechnol* 2015;**175**:65–74.

Stanke M, Keller O, Gunduz I *et al*. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res* 2006;**34**:W435–439.

Wallace-Salinas V, Brink DP, Ahren D *et al*. Cell periphery-related proteins as major genomic targets behind the adaptive evolution of an industrial*Saccharomyces cerevisiae* strain to combined heat and hydrolysate stress. *BMC Genomics* 2015;**16**:514.

Walsh P, Bursac D, Law YC *et al*. The J-protein family: modulating protein assembly, disassembly and translocation. *EMBO Rep* 2004;**5**:567–71.

Wicking JB, Bian Y. Nutritional Supplement Containing Iron. *United States Patent* 2015, No. 20150250839.

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 2007;**24**:1586–91.

Yue JX, Li J, Aigrain L *et al*. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat Genet* 2017;**49**:913–24.

Yun C, Bauler M, Moore RE *et al*. The role of the FRE family of plasma membrane reductases in the uptake of siderophore-iron in *Saccharomyces cerevisiae*. *J Biol Chem* 2001;**276**:10218–23.

Zhang J, Lan H, Chan Y *et al*. BGSA: a bit-parallel global sequence alignment toolkit for multi-core and many-core architectures. *Bioinformatics* 2019;**35**:2306–2308.

Zhang Y, Lin Z, Wang M *et al*. Selective usage of isozymes for stress response. *ACS Chem Biol* 2018;**13**:3059–64.