

SCIENTIFIC REPORTS



OPEN

Unraveling the genetic basis of xylose consumption in engineered *Saccharomyces cerevisiae* strains

Received: 04 August 2016
Accepted: 11 November 2016
Published: 21 December 2016

Leandro Vieira dos Santos^{1,2}, Marcelo Falsarella Carazzolle¹, Sheila Tiemi Nagamatsu¹, Nádia Maria Vieira Sampaio³, Ludimila Dias Almeida², Renan Augusto Siqueira Pirolla², Guilherme Borelli¹, Thamy Lívia Ribeiro Corrêa¹, Juan Lucas Argueso³ & Gonçalo Amarante Guimarães Pereira^{1,2}

The development of biocatalysts capable of fermenting xylose, a five-carbon sugar abundant in lignocellulosic biomass, is a key step to achieve a viable production of second-generation ethanol. In this work, a robust industrial strain of *Saccharomyces cerevisiae* was modified by the addition of essential genes for pentose metabolism. Subsequently, taken through cycles of adaptive evolution with selection for optimal xylose utilization, strains could efficiently convert xylose to ethanol with a yield of about 0.46 g ethanol/g xylose. Though evolved independently, two strains carried shared mutations: amplification of the xylose isomerase gene and inactivation of *ISU1*, a gene encoding a scaffold protein involved in the assembly of iron-sulfur clusters. In addition, one of evolved strains carried a mutation in *SSK2*, a member of MAPKKK signaling pathway. In validation experiments, mutating *ISU1* or *SSK2* improved the ability to metabolize xylose of yeast cells without adaptive evolution, suggesting that these genes are key players in a regulatory network for xylose fermentation. Furthermore, addition of iron ion to the growth media improved xylose fermentation even by non-evolved cells. Our results provide promising new targets for metabolic engineering of C5-yeasts and point to iron as a potential new additive for improvement of second-generation ethanol production.

The global energy matrix continues to rely heavily on fossil fuels whose combustion releases large amounts of CO₂ into the atmosphere, a potent greenhouse gas and a major contributor to climate change^{1–4}. Bioethanol has been touted as a cleaner and renewable alternative to replace fossil fuels and reduce carbon emissions⁵. In Brazil, first-generation ethanol is produced from the fermentation of sucrose-rich sugarcane extract, using a process in which *Saccharomyces cerevisiae* yeast cells are continuously recycled through successive batches, in some cases for up to 250 days in a row⁶. This prolonged cell recycling regimen combined with the harsh environment found in the bioethanol distilleries require strains that display traits such as high tolerance to abiotic stress and robust growth to outcompete microbial contaminants. One of the most widely adopted strains, Pedra-2 (PE-2), was itself originally isolated as an aggressive wild contaminant⁷. It combines high fermentative capacity and environmental stress tolerance, with the ability to prevent other yeast invaders from getting established in the fermentation tanks, thus providing stability to the distilleries' industrial operations. These desirable traits make PE-2 an ideal biological platform for the delivery of new technologies essential for second-generation (2G) bioethanol production^{8,9}.

A major challenge to the establishment of second-generation operations is the availability of biocatalysts able to metabolize the five-carbon (C5) sugar xylose derived from the hydrolysis of lignocellulosic biomass feedstocks⁵. *S. cerevisiae* is normally unable to metabolize xylose, but two pathways have been commonly used to engineer this activity into yeast cells^{10,11}. One is the oxidative-reductive pathway, catalyzed by the enzymes xylose reductase (XR) and xylitol dehydrogenase (XDH). A more advantageous alternative is the xylose isomerase (*xylA*) pathway, which involves a single-step conversion of xylose to xylulose. The *xylA* enzyme does not require cofactors and therefore does not have the redox bottleneck associated with the XR-XDH pathway. In addition, since it has lower accumulation of by-products, the total amount of ethanol that can be produced is higher^{10,12–14}.

¹Laboratório de Genômica e Expressão, Departamento de Genética e Evolução, UNICAMP, Campinas, São Paulo 13083-970, Brazil. ²GranBio/BioCelere, Campinas, Brazil. ³Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins-CO, 80523-1618, USA. Correspondence and requests for materials should be addressed to G.A.G.P. (email: goncalo@unicamp.br)

The successful expression of a functional *xylA* in *S. cerevisiae* allowed the development of strains able to ferment xylose to ethanol^{15,16}. However, the resulting strains showed low xylose consumption rates under anaerobic conditions, which are incompatible with industrial scale production. To enhance strain performance, some genetic modifications have been shown as beneficial. For example, overexpression of genes from the non-oxidative phase of pentose phosphate pathway (PPP), deletion of the aldose reductase gene *GRE3*¹⁷ and transporters for enhanced xylose uptake rates¹⁸. In addition, adaptive evolution in xylose medium is typically conducted in batch or chemostat cultivation cycles to achieve the best results^{19–21}. Successful examples of strains developed through this approach have been reported^{22–24}, however, the specific mutations responsible for the acquired ability to metabolize xylose have not been well characterized.

In this work, we used PE-2 as the platform for construction of yeast strains able to efficiently ferment xylose into ethanol. Applying a relatively short period of evolutionary engineering (from 3 to 12 batches), we were able to select strains capable of efficiently consuming xylose as a sole carbon source in anaerobic conditions. We report a detailed genomic and physiological characterization of the best-performing evolved strains, and demonstrate that the improved consumption of xylose is associated with a high copy-number of *xylA* and inactivation of the *ISU1* and *SSK2* genes. In addition, with a better comprehension of C5-metabolism, we point out iron as a new additive that can induce improvement on performance of xylose fermenting strains, thus, suggesting new directions to increase the economic viability of second-generation technologies.

Results

Metabolic engineering and adaptive evolution of xylose fermentation. In order to build a robust xylose-fermenting *S. cerevisiae* strain, we dissected a tetrad from the industrial diploid strain PE-2 and chose a *MAT α* haploid spore (LVYA1) for further genetic manipulations. Cassettes containing genes of xylose pathway (Supplementary Fig. S1) were designed and their integration was targeted to regions in the vicinity of centromeres, based on the rationale that such loci might offer higher stability to the integrated constructs due to lower recombination rates relative to other regions of chromosomes²⁵. The first cassette contained the eukaryotic codon-optimized *xylA* gene from *Orpinomyces* sp. was integrated near *CEN5*. Next, we deleted the *GRE3* gene encoding aldose reductase in order to decrease the production of xylitol²⁶, and integrated two additional copies of xylulokinase (*XKS1*) near *CEN2* and *CEN8*. Furthermore, an additional copy of each of the genes encoding the four non-oxidative enzymes from the pentose phosphate pathway, transaldolase (*TAL1*) and ribose-5-phosphate ketol-isomerase (*RK11*) were inserted near *CEN12*; and transketolase (*TKL1*) and ribulose-5-phosphate 3-epimerase (*RPE1*) were inserted near *CEN13*, resulting in the strain LVY27. All these genes were expressed under the control of different strong constitutive promoters of genes from the glycolytic pathway of *S. cerevisiae* (Table 1).

Despite integration of the entire xylose utilization pathway, the resulting strain LVY27 consumed xylose very slowly and growth was observed only under aerobic conditions. To achieve a higher activity of xylose isomerase, we applied an approach to facilitate amplification of the *xylA* gene. Plasmid pOXylATy1 (Supplementary Table 1) was created to incorporate a cassette containing *xylA* flanked by δ LTR sequences. The plasmid was cleaved and the purified fragment, containing the δ -*xylA*- δ fragment was used in a new integrative transformation (Fig. 1a).

Following the premise that cells containing the highest number of δ -*xylA*- δ fragment integrations would have a higher xylose consumption rate, ~300 colonies derived from the transformation of strain LVY27 were selected to undergo adaptive evolution in xylose medium. The different transformants were allowed to compete with each other in YNBX medium in semi-anaerobic conditions. Glucose was added to the medium to allow initial growth and the cells were split into two new sealed bottles and parallel adaptive evolution experiments (A and B) were started (Fig. 1a). After glucose depletion, the cells were maintained in xylose as a sole carbon source for 48 hours before the next transfer. We started to observe cellular growth within three serial batch cultivations in xylose for evolution A and after 12 transfers for the parallel evolution B (approximately 15 and 60 generations, respectively). From each evolution, 12 candidate colonies displaying the largest diameter in solid xylose medium were selected and their growth was screened to identify the best xylose fermenting strains. Based on their faster growth kinetics in xylose medium, strains LVY34.4 and LVY41.5 were selected as the most promising clones derived from evolution A and B, respectively (Supplementary Fig. S2A and S2B).

Initially, each selected clone was compared to the parental strain LVY27 in YP medium supplemented with 30 g/L of xylose. Under these conditions, the parental strain LVY27 was not able to consume xylose or grow, while LVY34.4 and LVY41.5 consumed all xylose in less than 24 hours, with maximum specific growth rate of 0.213 and 0.129 h⁻¹, respectively. Both had final ethanol yields of approximately 0.46 g ethanol/g xylose, but LVY34.4 had higher xylose consumption rate (Fig. 1b, Table 2). Glycerol and xylitol were observed in negligible levels. Taken together, these results show that the evolutionary engineering process allowed the selection of two strains, able to efficiently consume xylose and produce ethanol. In order to determine the genetic mechanisms through which improved xylose utilization was achieved, we sequenced the whole genomes of parent LVY27 and both evolved strains. The sequencing reads were analyzed to detect both structural rearrangements detected through depth coverage and nucleotide polymorphisms in the evolved clones.

Genomic characterization of evolved strains I: *xylA* amplifications. The analysis of copy number variation (CNV) was performed using the cn.MOPS program²⁷, employing the read-depth profile from each strain generated by paired-end reads alignment into the reference genome (parental genome assembly – LVY27). The comparison of read-depth between parental and evolved strains revealed a substantial amplification of a 6,237-bp region containing the genes *xylA* and *LEU2* that was previously inserted into the parental genome with plasmid pOXylATy1. This region was amplified 36-fold and 26-fold during the adaptive evolution process for the LVY34.4 and LVY41.5 strains, respectively. The Supplementary Figure S3 plots read-depth coverage in this region for each evolved strain and demonstrates the large coverage variation in the *xylA* + *LEU2* genes for both clones, that was around 1.34 times higher for LVY34.4 than LVY41.5. To further investigate this amplified region we

Strain	Parent strain	Relevant genotype/features	Reference/source
PE-2		Industrial ethanol production strain PE-2; <i>MATa/α</i>	CBMAI 0959
CAT-1		Industrial ethanol production strain CAT-1; <i>MATa/α</i>	CBMAI 0957
BG-1		Industrial ethanol production strain BG-1; <i>MATa/α</i>	7
SA-1		Industrial ethanol production strain SA-1; <i>MATa/α</i>	CBMAI 1547
Caeté-1		Industrial ethanol production strain Caeté-1; <i>MATa/α</i>	Usina Caeté S.A.
LVYA1	PE-2	<i>MATα</i>	This study
LVYA2	PE-2;	<i>MATa</i>	This study
LVY27	LVYA1	<i>MATα</i> ; <i>CEN5::pTDH1-xylA-tTDH1</i> ; <i>gre3Δ</i> ; <i>CEN2::pADH1-XKS1-tADH1</i> ; <i>CEN8::pADH1-XKS1-tADH1</i> ; <i>CEN12::pTDH1-TAL1-tTDH1-pPGK1-RK11-tPGK1</i> ; <i>CEN13::pTDH1-TKL1-tTDH1-pPGK1-RPE1-tPGK1</i>	This study
LVY34.4	LVY27	<i>MATα</i> ; pOXyATy1 + adaptive evolution and selection	This study
LVY41.5	LVY27	<i>MATa/α</i> ; pOXyATy1 + adaptive evolution and selection	This study
LVY41.5EVx	LVY41.5	<i>MATa/α</i> ; adaptive evolution and selection	This study
LVY54	LVY34.4	<i>MATα</i> ; <i>ura3Δ</i>	This study
LVY55	LVY54	<i>MATα</i> ; <i>iki3Δ::URA3</i>	This study
LVY56	LVY54	<i>MATα</i> ; <i>isu1Δ::URA3</i>	This study
LVY57	LVY55	<i>MATα</i> ; <i>ura3Δ::IKI3</i>	This study
LVY58	LVY55	<i>MATα</i> ; <i>ura3Δ::iki3-Ile398Val</i>	This study
LVY59	LVY56	<i>MATα</i> ; <i>ura3Δ::ISU1</i>	This study
LVY60	LVY56	<i>MATα</i> ; <i>ura3Δ::isu1-Leu132Phe</i>	This study
LVY61	LVY56	<i>MATα</i> ; <i>ura3Δ::isu1-Val89Leu</i>	This study
LVY64	LVY54	<i>MATα</i> ; <i>isu1Δ</i>	This study
LVY65	LVY27	<i>MATα</i> ; <i>ura3Δ</i>	This study
LVY66	LVY65	<i>MATα</i> ; pOXyA2	This study
LVY67	LVY66	<i>MATα</i> ; pOXyA2; <i>isu1-Leu132Phe</i>	This study
LVY68	LVY27	<i>MATα</i> ; <i>isu1-Leu132Phe</i>	This study
LVY69	LVY66	<i>MATα</i> ; pOXyA2; <i>ssk2Δ</i>	This study
LVY70	LVY67	<i>MATα</i> ; pOXyA2; <i>ssk2Δ</i> ; <i>isu1-Leu132Phe</i>	This study
LVY71	LVY34.4 x LVYA2	<i>MATa/α</i> ; <i>isu1-Leu132Phe/ISU1</i>	This study
LVY72	LVY71	<i>MATa/α</i> ; adaptive evolution and selection; <i>isu1-Leu132Phe/isu1-Leu132Phe</i>	This study

Table 1. Yeast strains used in the study.

performed a comparative sequence analysis that identified a specific contig present only in the evolved cells that harbor a combination of *LEU2* terminator, the δ LTR element, and the *TDH1* promoter sequences. This specific combination was generated through the fusion of two cassettes from pOXyATy1, suggesting that a tandem amplification of this site occurred during the adaptive evolution process. Alternatively, the amplification of the cassette could have occurred concomitantly with the initial integration event.

In order to evaluate the stability of the transformants obtained, we retested the presence of each cassette containing the *xylA*, *XKS1* and other genes of the PPP into the genome of the evolved strains through PCR analysis. The presence of all cassettes was confirmed for all strains, except for the *xylA* cassette inserted near *CEN5*, whose presence was not detected in the strains LVY34.4 and 41.5 (Supplementary Fig. S4). By combining primers that anneal at the δ LTR region of plasmid pOXyATy1 with primers of regions flanking the integration site at *CEN5* (Fig. 2b, R2 and R3; Supplementary Table S2), we found that the cassette from pOXyATy1 recombined with the *xylA* already present in the genome of the strain LVY27 and integrated adjacent to that locus (Fig. 2a). In Fig. 2b, PCR reaction R1 used a combination of primers that annealed to the terminal region of each adjacent cassette. This product was excised from the agarose gel and sequenced. The sequencing analysis confirmed that both clones LVY34.4 and LVY41.5 contained tandem segmental amplifications of the *xylA* cassette in chromosome 5 (Chr05; Fig. 2b). Cells containing more copies of *xylA* likely had a growth advantage in xylose and succeeded under selective pressure (Fig. 2a).

To validate the presence of these chromosomal rearrangements we separated full-length chromosomes of the selected clones by Pulse Field Gel Electrophoresis (PFGE), using settings optimized to resolve the whole karyotype (Fig. 3a). We observed that LVY27 had the normal high intensity band at ~580 Kb that corresponds to the position where Chr05 and Chr08 co-migrate. In the LVY34.4 and LVY41.5 clones, we observed significant reductions in the intensities of the ~580 Kb band corresponding only to Chr08, indicating that Chr05 was present elsewhere in the PFGE. In LVY34.4, a new longer chromosome of ~700 Kb appeared, consistent with a ~120 Kb increase in size to Chr05 which would be expected for a tandem segmental amplification containing ~36 copies of the *xylA* cassette estimated from read depth coverage. Likewise, LVY41.5 had a higher intensity band at ~660 Kb, at the position where a single copy of Chr11 is normally found. This was consistent with a ~80 Kb increase in size to Chr05 due the presence of ~26 copies of the *xylA* cassette.

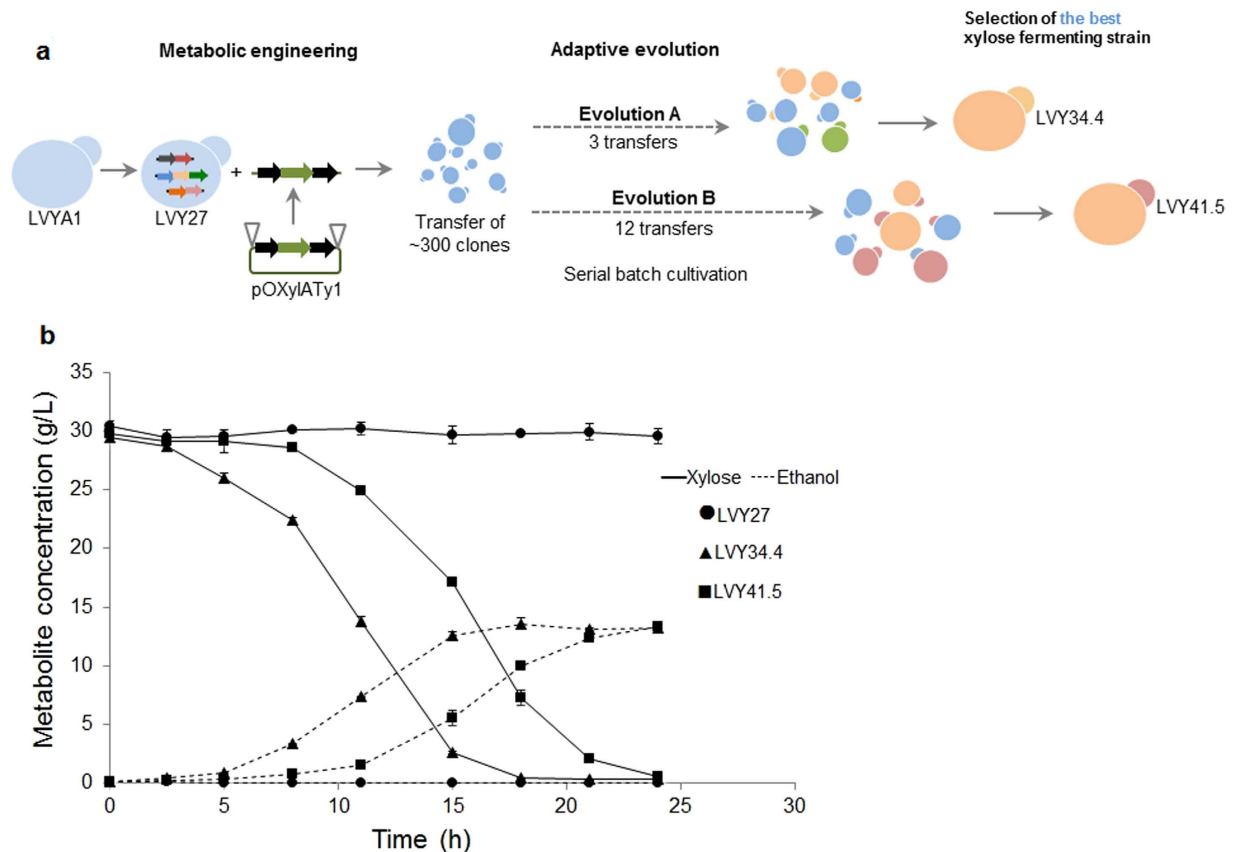


Figure 1. Development and comparative fermentation of evolved strains. (a) Flow diagram of genetic transformation and adaptive evolution steps. LVY27 strain was transformed with a DNA fragment containing the *xylA* gene (green arrow), which was flanked by two δ LTR regions (black arrows). Approximately 300 resulting clones were used in adaptive evolution experiments. After a first batch of cultivation, the cells were split into two parallel evolution experiments (a and b). The colors represent the mutations acquired by the strains during evolution. The fastest xylose fermenting strains from each evolution, LVY34.4 and LVY41.5, were selected for further investigation. (b) Comparative performance of strains in batch fermentation. Strains were cultivated in YP medium supplemented with 30 g/L of xylose with an initial cell density of 0.25 g DW/L of the parental LVY27 (●) and evolved cells LVY34.4 (▲) and LVY41.5 (■). The fermentations were performed in triplicate and error bars represent standard deviation from the average of values.

Strain	Copy number of <i>xylA</i>	Condition	Xylose consumption rate (g/g ⁻¹ h ⁻¹)	Ethanol production rate (g/g ⁻¹ h ⁻¹)	Ethanol yield (g/g ⁻¹)	Xylitol yield (g/g ⁻¹)	Glycerol yield (g/g ⁻¹)
LVY34.4	36	Semi-anaerobic batch, synthetic medium (YPX30)	1.32	0.62	0.46 ± 0.02	0.005 ± 0.00	0.01 ± 0.00
LVY41.5	26	Semi-anaerobic batch, synthetic medium (YPX30)	1.03	0.45	0.45 ± 0.02	0.006 ± 0.00	0.01 ± 0.00

Table 2. Fermentation performance of LVY34.4 and LVY41.5 strains.

To refine our analyses of longer versions of Chr05 at those positions in LVY34.4 and LVY41.5, we changed the PFGE running conditions to maximize the separation of DNA molecules between 500 Kb and 800 Kb (Fig. 3b). Agarose slices were excised from these high resolution gels, and the DNA present in them was used as template for quantitative real time PCR (qRT-PCR; Fig. 3c). This analysis confirmed that Chr05 DNA was enriched relative to normalizing Chr09 and Chr16 sequences in the 580 Kb region of the parent strain LVY27 karyotype, but was not enriched in the two selected strains. Conversely, Chr05 specific DNA was enriched in the 700 Kb region in LVY34.4 and the 660 Kb region in LVY41.5. Taken together, the read depth, PCR, and karyotype analyses above confirmed that each of the two evolved clones carried tandem segmental amplifications of the *xylA* cassette insertion on Chr05.

Genomic characterization of evolved strains II: Point mutations. The paired-end sequencing reads were assembled into large contigs (>1,000 bp) totaling 226 contigs (coverage, ~56×) for LVY27, 213 contigs (coverage, ~140×) for LVY34.4 and 232 contigs (coverage, ~128×) for LVY41.5. The whole genome sequencing

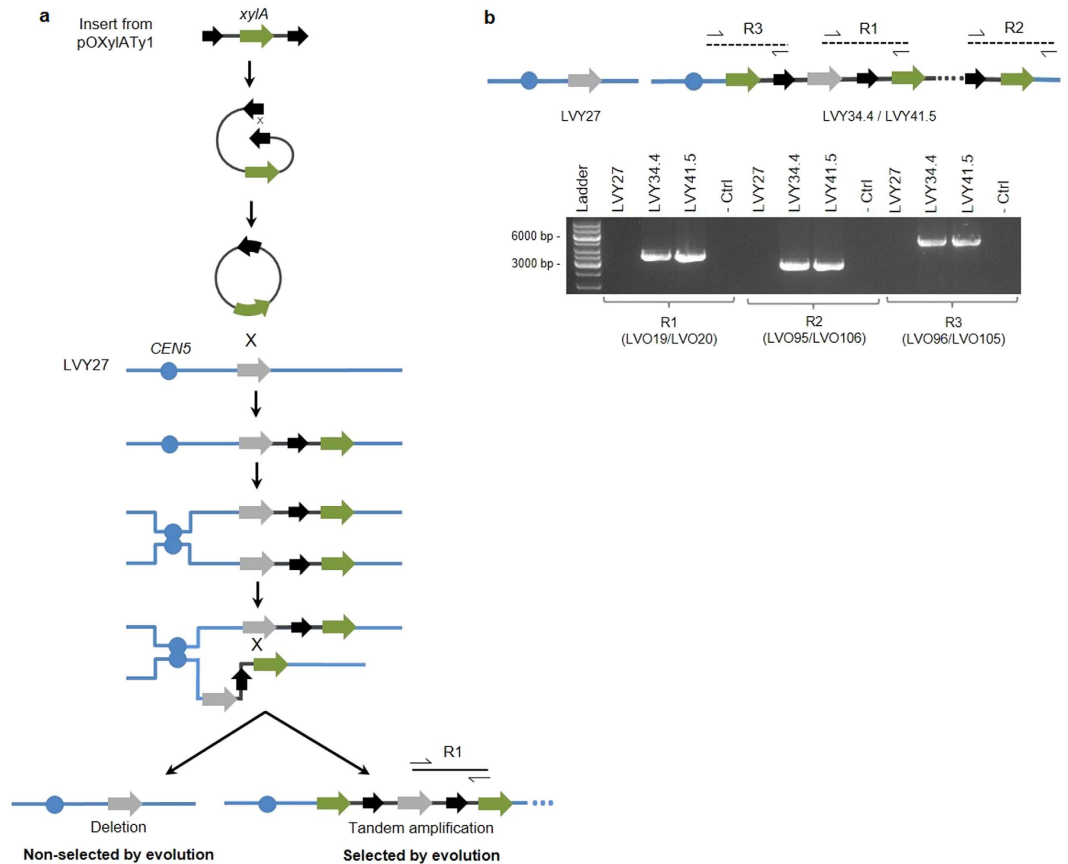


Figure 2. Potential mechanism that led to in tandem amplification of *xylA* in the evolved strains during evolution. (a) The blue strand represents a section of Chr05 in LVY27. The insert contains the *xylA* gene (green arrow) obtained from plasmid pOXylATy1 flanked by two δ LTR sequences (black arrows) in the same orientation. We hypothesize that shortly after the transformation event, the flanking LTR sequences from the DNA fragment recombined with each other creating a circular molecule. Next, a homologous recombination event occurred between the *xylA* sequence from the circle and the *xylA* sequence (represented by a gray arrow) that was previously inserted into the parental genome next *CEN5* region, resulting in the chromosomal integration of the circle. During duplication of the cell carrying this integration, unequal crossovers involving sister chromatids can result in tandem duplications of the insert containing *xylA*, thus leading to the amplification of this gene. Selection for growth in xylose allowed such *xylA* amplification events to increase in frequency in the cell population. (b) PCR reactions showing the tandem multiplication of *xylA* fragment during the evolution. The primers in the R2 and R3 reactions anneal with the Ty1 element (black arrow) and regions outside the integration locus. The R1 reactions amplified regions between two tandem cassettes. As expected, the reactions R1, R2 and R3 did not present positive amplification in LVY27. 1 Kb DNA ladder GeneRuler; (- Ctrl) control reaction without DNA.

results are summarized in Supplementary Table S3. The identification of single nucleotide polymorphisms (SNPs) was performed using GATK package²⁸. We identified a total of 335, 362, 375 variations for parental, LVY34.4 and LVY41.5, respectively. Interestingly, the comparison of allelic frequencies of each SNP between LVY34.4 and LVY41.5 showed that the allelic frequency of SNPs in LVY34.4 was 1 as expected, while for LVY41.5 was always 0.5 (Supplementary Table S4). This result suggested that LVY41.5 potentially underwent a diploidization event during the adaptive evolution.

A customized PERL script was developed to identify polymorphisms that systematically appeared in the evolved strains and that, conversely, were absent in LVY27 (see methods). A total of nine SNPs and one indel were identified using this approach. These SNPs/INDELS were manually annotated using the reference genome of *S. cerevisiae* and filtered by mutations residing only inside coding sequences. This analysis resulted in a short list of potentially significant nucleotide mutations acquired during the adaptive evolution (Supplementary Table S4). Comparing the shared modifications, three non-synonymous SNPs were distributed in two genes: *ISU1* and *IKI3*. In addition, another polymorphism of interest, a non-sense mutation in *SSK2*, was identified in LVY41.5 strain.

Functional analyses of nucleotide mutations. The strain LVY34.4 was initially chosen as the background to individually validate the two identified shared mutations at the *IKI3* and *ISU1* genes. Initially, the *URA3* gene was deleted in LVY34.4 so it could later be used as a selection marker, resulting in the strain LVY54 (Table 1; see methods). Both alleles of the *IKI3* gene were amplified either from the parental LVY27 strain (WT

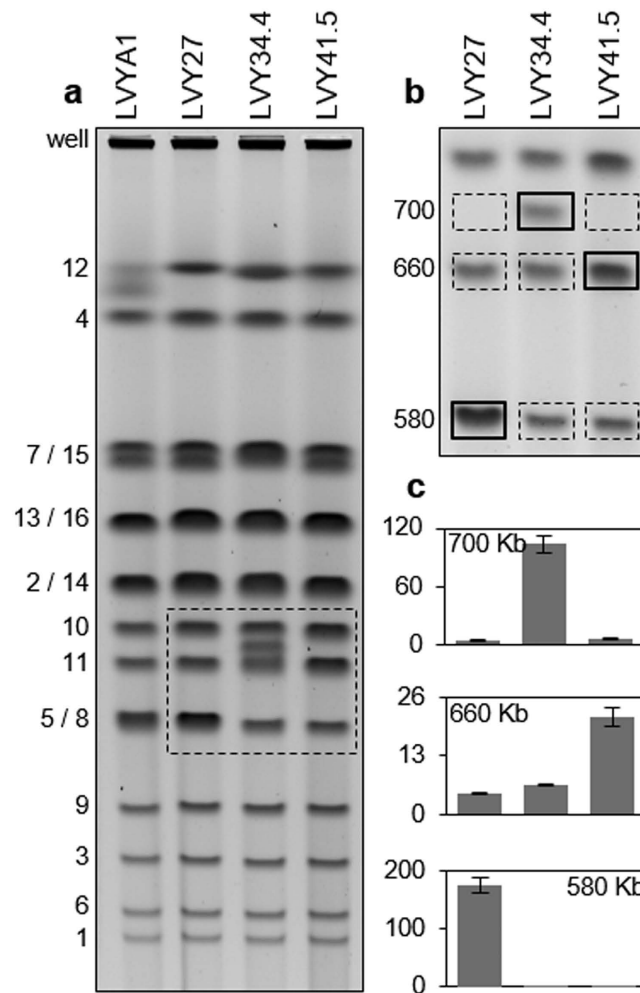


Figure 3. Karyotype analysis of evolved clones. (a) Pulse field gel electrophoresis (PFGE) under running conditions for optimal separation of all chromosomes. The sample loading well and bands corresponding to the wild PE-2 (LVYA1) and the parental strain LVY27 chromosomes are identified to the left. Yeast clones are indicated at the top of each lane. The dashed outline corresponds to the region of the PFGE karyotype analyzed at higher resolution in B. (b) Cropped PFGE under running conditions for maximum separation of molecules between 500 Kb and 800 Kb. Molecules larger than Chr10 ran as a single size-compressed band above the cropping point, and molecules smaller than Chr3 ran off the gel. Rectangles indicate the position of agarose slices excised from the gel, from which DNA was purified for qRT-PCR. Continuous outlines indicate gel slices with enriched Chr5 DNA in each lane, corresponding to the position of the parental size Chr5 in LVY27 (~580 Kb; 1 *xylA* repeat unit), and the rearranged Chr5 molecules containing segmental amplifications of the *xylA* insertion in LVY34.4 and LVY41.5 (700 Kb and 660 Kb, ~36 and ~26 repeat units, respectively). Dashed outlines correspond to gel slices without Chr5 DNA enrichment. (c) Plots of qRT-PCR results, showing the Chr5 DNA relative fold enrichment ratios (Y axes) in the gel slices. The bars corresponding to each clone appear directly under the respective gel lane from B. The size of the DNA fragments in the excised gel slice is indicated inside each plot.

sequence) or the evolved LVY34.4 strain (*Ile398Val*), resulting in the strains LVY57, which contained the gene in its original form, and LVY58 that contained the mutated copy of the gene. The fermentative performance of LVY57 and LVY58 were evaluated and compared to the LVY54 control (LVY34.4 *ura3Δ*, Table 1). All strains displayed a lower rate of xylose consumption due to uracil auxotrophy. Under the tested conditions, we did not observe a significant difference in xylose consumption or ethanol production between LVY57 and LVY58 (Supplementary Fig. S5), suggesting that the SNP identified in *IKI3* was not related to the improved capacity for xylose consumption at the evolved strains.

Next, we sought to investigate if the mutation found in the *ISU1* gene of evolved strains could be associated with the increased rates of xylose fermentation. Similarly to the approach above, *ISU1* was amplified from the genome of parental and evolved cells and used to originate LVY59, that harbored the gene in its original form; LVY60, which contained the allele from LVY34.4 (*Leu132Phe*); and LVY61, which contained allele found in LVY41.5 (*Val89Leu*) (Table 1). The xylose consumption rate in *ISU1* knockout strain LVY64 was also assessed. The LVY54 control strain consumed all the xylose in the medium in 40 hours, similar to those observed for

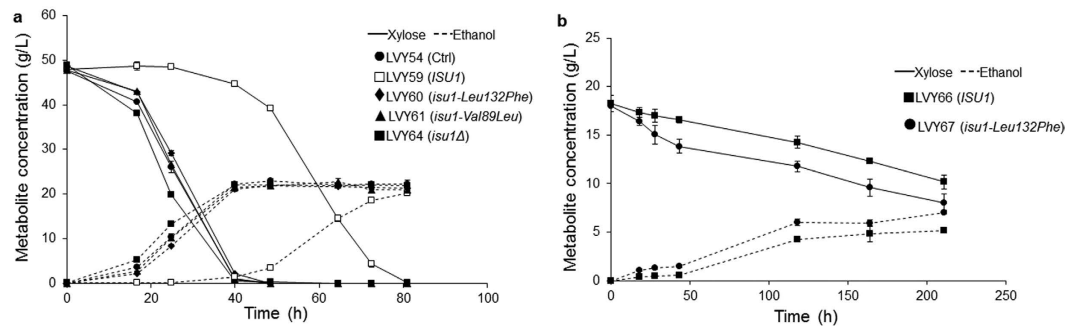


Figure 4. Fermentative performance of the transformants with *ISU1* SNPs in xylose medium. (a) Effect of *ISU1* mutations in the fermentative performance of strains. Strains were cultivated in YPX in batch fermentation with an initial cell density of 0.12 g DW/L. (b) Analysis of the effect of *ISU1* mutation in non-evolved cells. Strains were cultivated in YNBX. All the fermentations were performed in triplicate and error bars represent standard deviation from the average of values.

LUY60 and LUY61. The two mutations had similar effects and no significant difference between the two strains was observed. The deletion of *ISU1* also enabled a similar consumption of xylose compared with the strains that have a mutation that substitutes an amino acid in Isu1p protein. In contrast, LUY59, which had the wild type *ISU1* allele, displayed very slow xylose consumption and sugar depletion was observed only after 80 hours (Fig. 4a). Taken together, these results confirmed that mutations that arose in *ISU1* during the adaptive evolution process contributed significantly to the enhanced fermentation of this sugar.

We also inserted the *isu1-Leu132Phe* allele in a strain that had not been evolved to isolate and test the effect of each mutation. To simulate the tandem amplification of the *xylA* gene, we introduced the multi-copy plasmid pOXyIA2 (*URA3*) into LUY65 (LUY27 *ura3Δ*, Table 1). The fermentation was performed in YNBX medium without uracil. LUY67 consumes xylose slightly better than the cell with the wild type *ISU1* gene, LUY66 (Fig. 4b). However, the consumption rate for both strains in YNBX medium was much slower compared to YPX, even with *xylA* expressed from a high copy plasmid. We also isolated the effect of multiple copies of *xylA* by testing strain LUY68 with a single copy of *xylA* gene. As expected, the growth was very slow, suggesting that high expression of *xylA* gene is a necessary condition for efficient consumption of xylose (data not shown).

As previously mentioned, with the exception of *IKI3*, the SNPs found in the genomic analysis of LUY41.5 appeared at a frequency of 50% (Supplementary Table S4). This heterozygous pattern was not expected for *de novo* mutations that originated in a heterothallic haploid parent strain, thus suggested the occurrence of a diploidization event at some point in the LUY41.5 lineage during evolution. Using PCR²⁹ analysis and flow cytometry³⁰ (data not shown) we confirmed that LUY41.5 was in fact a diploid organism, but the mechanism that led to diploidization remained undetermined. LUY41.5 also contained a heterozygous nonsense mutation in the *SSK2* gene encoding a kinase member of the MAPKKK signaling pathway that regulates responses to osmotic stress conditions³¹. The *SSK2* gene was deleted in the parent strains as a single mutant, and in combination with the *isu1-Leu132Phe* mutation, resulting in LUY69 and LUY70, respectively. The strain containing only the *SSK2* deletion consumed xylose faster than control LUY66, confirming that this modification is also associated with xylose fermentation (Fig. 5a). Furthermore, in comparison to *isu1* mutation, LUY69 consumed xylose more efficiently and the *ssk2Δ isu1-Leu132Phe* double mutant LUY70 had an even better result, demonstrating a synergistic interaction between these two pathways toward enhancement of xylose metabolism (Fig. 5b).

Iron supplementation contributes to increased xylose fermentation. Disruption of mitochondrial iron-sulfur clusters assembly results to increased accumulation of intramitochondrial iron concentrations^{32,33}. Based on the hypothesis that increased levels of iron leads to increased xylose consumption, new fermentation assays were performed in YNBX medium supplemented with iron for strains described above. The evolved LUY34.4 were able to consume xylose and produce ethanol remarkably faster in the presence of increased concentration of iron ions, supporting the connection between the ability to consume xylose and the changes in iron metabolism observed in the evolved strains (Fig. 6a). The possible influence of other cofactors for the xylose isomerase activity such as Mg²⁺ and Mn²⁺ were also tested and did not result in improved growth in xylose (data not shown). The impact of iron ion was also tested on non-evolved cells. In this case, iron supplementation was also able to improve the consumption rate, although the strains harboring the *isu1-Leu132Phe* mutation had a better performance probably due to the higher capacity of accumulate iron intracellularly (Fig. 6b). The same effect was observed for the strains with *SSK2* deletion (Supplementary Fig. S6). Concomitantly, the addition of iron to reaction containing extracts of LUY34.4 strain significantly increased the xylose isomerase activity (Fig. 6c).

Homozygous and heterozygous effect of *ISU1* mutation. LUY41.5 was subjected to a new round of adaptive evolution. After 11 transfers, the derivative strain LUY41.5EVx showed an increase rate of 2.7 times in xylose consumption compared to its immediate parent (Supplementary Fig. S7). As the LUY41.5 is a heterozygous *ISU1/isu1-Val89Leu*, the *ISU1* locus was sequenced in both strains. We observed two peaks in the LUY41.5 chromatogram corresponding to each allele of the parental spore. In LUY41.5EVx, only one peak was observed,

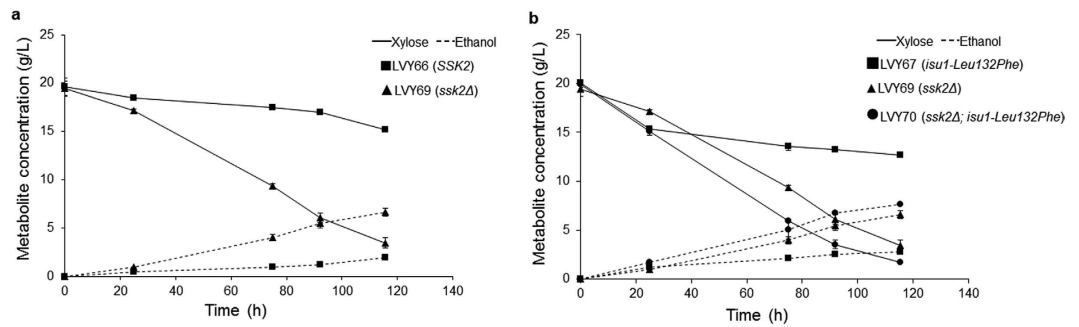


Figure 5. Comparative analysis of *SSK2* and *ISU1* mutations in non-evolved cells. Strains were cultivated in YNBX medium. **(a)** Effect of *ssk2* deletion. **(b)** The *ssk2* deletion confers a better xylose consumption rate than *isu1* mutation. The combinations of mutations give a better performance to double-mutant LVY70. All fermentation was performed in triplicate and error bars represent standard deviation from the average of values.

indicating that this strain had become homozygous *isu1-Val89Leu/isu1-Val89Leu*. Therefore, the increase in the xylose consumption rate was due to the *ISU1* locus becoming in homozygous after second round of evolution. The *SSK2* mutation remained heterozygous.

To further investigate this effect, we crossed our best xylose fermenting strain LVY34.4 (*MAT* α) with a *MAT* α wild type haploid spore from PE-2 (Table 1). The resulting diploid was named LVY71 and its xylose fermenting capacity was compared to the parental spores in YPX. As expected, no growth was observed for the wild type spore, while LVY34.4 consumed all the sugar in 13 hours (Fig. 7a). The diploid was able to consume xylose, however, with very slow growth. LVY71 was heterozygous *ISU1/isu1-Leu132Phe*. With the premise that this is a recessive loss-of-function mutation, LVY71 was subjected to another round of adaptive evolution. The adaptation took twelve transfers and the evolved diploid LVY72 was chosen as the best xylose fermenting strain. LVY72 consumed xylose faster than its non-evolved version, at a rate comparable to the haploid LVY34.4 (Fig. 7a). The *ISU1* locus was sequenced in both strains and a similar sequence result was found. In the evolved strain LVY72, only one peak was observed, indicating that the strain had become *isu1-Leu132Phe/isu1-Leu132Phe*, confirming that homozygous *isu1* mutation ensures a more efficient consumption of xylose (Fig. 7b). We have seen that this was a recurrent phenomenon. We have repeated this procedure by crossing LVY34.4 with 4 haploid cells derived from the industrial strains CAT-1, SA-1, BG-1⁷ and CAETE-1 (isolated from a mill in the northeast of Brazil). In all cases, when the evolution led to an improved xylose consumption, we have identified that the *ISU1* wild type allele was lost, and the strains were homozygous for the *isu1-Leu132Phe* allele.

Discussion

In this work, we have used the industrial strain PE-2 as the genetic background for the construction of a robust xylose-fermenting microorganism. Promoters of three genes that are highly expressed during the industrial 1 G fermentation³⁴ were combined with coding sequences of 6 genes involved in xylose metabolism. Before construction, the codon usage of *Orpinomyces* sp xylose isomerase³⁵ - the main specific enzyme of the xylose metabolism - was optimized for a more efficient expression in *S. cerevisiae*.

All the metabolized xylose is channeled through PPP to the glycolytic pathway for production of ethanol. However, even with the complete xylose consumption pathway integrated and overexpressed, the developed strain LVY27 consumes xylose poorly. Since an efficient expression of xylose isomerase is one of the main metabolic bottlenecks to use this pathway, the cassette containing the *xylA* was reconstructed using the δ LTR element, which could increase the number of integrations as described elsewhere³⁶. Adaptive evolution of the engineered strain in xylose allowed a rapid improvement of the fermentative performance. In a very short time, we were able to select cells that efficiently fermented xylose. We found that in the selected strains that presented the highest fitness, LVY34.4 and LVY41.5, the integration of the second *xylA* expression cassette did not occur by homologous recombination with the retrotransposon Ty1 or solo δ LTRs elements. Instead, the single copy of the *xylA* sequence, already present in the host strain, was the actual site of integration (Fig. 2a and b). The best evolved strain LVY34.4 presents a yield of 0.46 g ethanol/g xylose (0.51 g/g is the *theoretical maximum*), similar or superior to most recently developed xylose-fermenting yeasts described in the literature^{22–24}.

The genome analysis of the evolved cells revealed a large increase in the copy number of the gene encoding xylose isomerase. Recently, it was reported the generation of tandem repeats of *xylA* gene through a self-replicating extrachromosomal circular DNA (eccDNA) during adaptive evolution³⁷. Considering that the *xylA* has a low activity in *S. cerevisiae*, the amplification during evolution seems to be an essential condition to achieve high enzymatic efficiency. The same evidence was observed before, through multiple integration of *xylA* in the evolved strain³⁸. A cell that harbors higher copy number of *xylA* during adaptive evolution in xylose media would convert this sugar faster, being positively selected (Fig. 2a). The difference in the copy number of *xylA* might partly explain the difference of productivity between LVY34.4 and LVY41.5 (Table 2).

In addition to the amplification of *xylA*, we identified some SNPs in evolved cells. As the replacement of the SNP discovered in *IKI3* resulted in no difference in xylose fermentation, this modification was determined not to be related to C5 metabolism (Supplementary Fig. S5). In the case of *ISU1*, the mutation appeared in different regions of the gene in the evolved cells isolated in separate experiments, suggesting a strong selective pressure in this gene for xylose adaptation. *Isu1p* is a conserved protein of mitochondrial matrix that performs a scaffolding

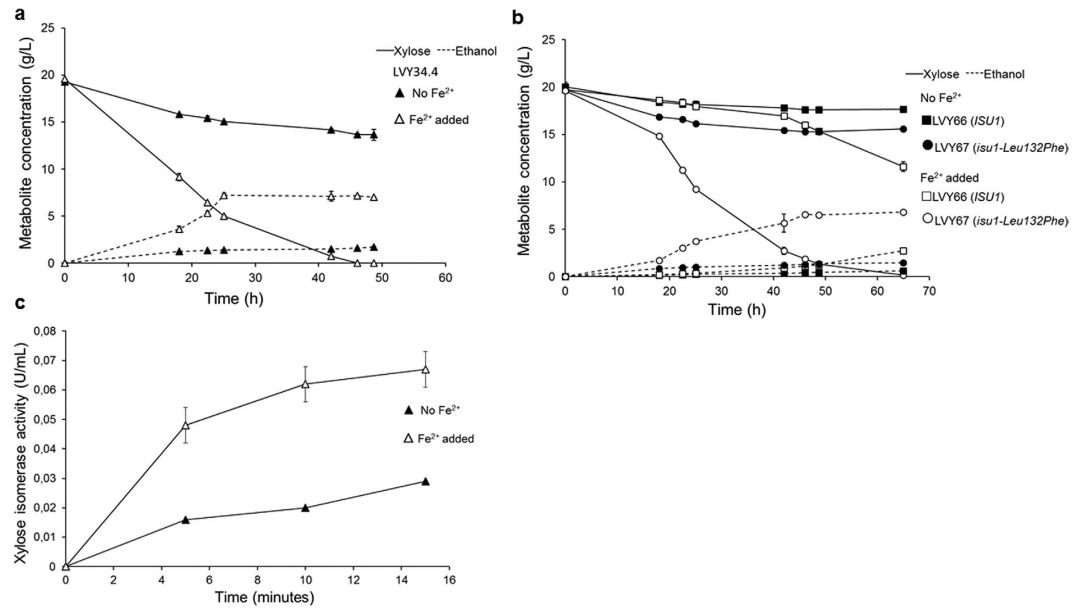


Figure 6. Iron supplementation effect on xylose fermentation. Effect of iron supplementation in YNBX on xylose consumption and ethanol production for LVY34.4 (a) and non-evolved strains LVY66 (*ISU1*) and LVY67 (*isu1-Leu132Phe*) (b). All fermentation was performed in triplicate and error bars represent standard deviation from the average of values. (c) Xylose isomerase activity of LVY34.4. The effect of Fe²⁺ ion was measured at 30 °C. The average of three replicates and standard deviations are presented.

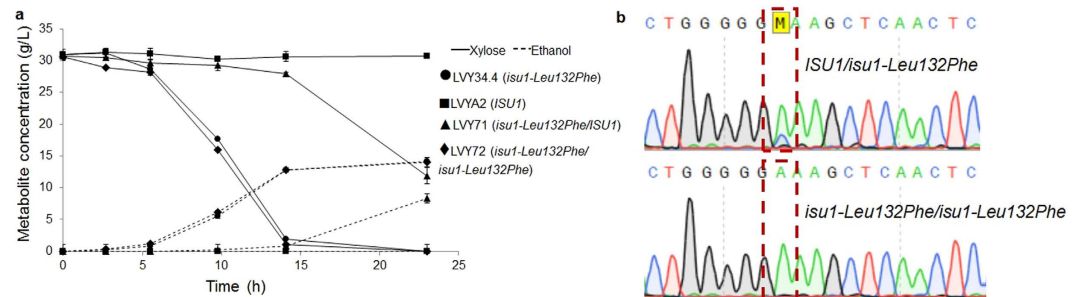


Figure 7. Homozygous and heterozygous effect of *ISU1* mutation in evolved and non-evolved diploid strains. Strains were cultivated in YP medium supplemented with xylose in a comparative batch fermentation (a) The haploids LVYA2 (■) and LVY34.4 (●) were used as controls. The non-evolved diploid LVY71 (▲) was compared to the evolved diploid LVY72 (◆), homozygous for *isu1* SNP. The fermentations were performed in triplicate and error bars represent standard deviation from the average of values. (b) Analysis of the top chromatogram from LVY71 indicated the presence of two peaks originating from both copies of the haploid spores LVYA2 and LVY34.4; while the evolved LVY72 bottom chromatogram from *ISU1* gene presents only one peak indicating the duplicate mutation by exchange of nucleotide at position 396 in both homologous chromosomes occurred during the second round of evolution.

function for the *de novo* synthesis of Fe/S clusters³⁹. Iron/sulfur clusters (ISC) are essential cofactors for proteins involved in electron transport, enzymatic catalysis and regulation of several cellular processes³³. The ISC system consists of a machinery of 17 proteins for Fe/S cluster assembly⁴⁰. The ferrous iron is imported from cytosol to mitochondria through the inner membrane carriers Mrs3p-Mrs4p using proton motive force with the assistance of the glutaredoxins Grx3p-Grx4p as iron donors that facilitate the import⁴¹. The cysteine desulfurase complex Nfs1p-Isd11p donates the sulfur and cooperates with the frataxin Yfh1p that facilitates the delivery of ferrous iron to the scaffold protein Isu1p for [2Fe-2S] biogenesis^{42,43}. In the second step, the co-chaperone Jac1p recruits the holo-form of Isu1p and conducts to the ATP-bound form of the mitochondrial Hsp70 chaperone Ssq1p. The binding of Ssq1p occurs in the LPPVK sequence of Isu1p, which triggers its ATPase activity inducing a conformational change that stabilizes the interaction and induces the dissociation^{40,41,44}. In the evolved strain LVY34.4, the *isu1* mutation occurred at the LPPVK sequence, by changing a leucine for a phenylalanine at position 132. The mutation at this site and at position 89 in strain LVY41.5 probably eliminated the functionality of this protein, since both strains with these mutations showed similar phenotypes in comparison to LVY64 carrying a full *ISU1*

gene deletion (Fig. 4a). The deletion of *ISU1* or the paralog *ISU2* have been reported to increase mitochondrial iron levels as well as reduce activity of [4Fe-4S] aconitase enzyme. Simultaneous deletion of both *ISU1* and *ISU2* is lethal to yeast cells³². In addition, yeast cells with depletion of other proteins of the Fe/S cluster scaffold machinery, such as *YFH1*, *NFS1*, or the chaperones involved also show increased concentrations of iron and mitochondrial oxidative damage³³.

In LVY41.5, we identified a new mutation in the *SSK2* gene that contributes for higher xylose consumption than *isu1* mutations. This phenotype was confirmed through *SSK2* deletion. Remarkably, the combination of mutations in *SSK2* and *ISU1* conferred the best result obtained. The LVY41.5 became diploid during the course of evolution, even PE-2 being a heterothallic strain. The mutation found in LVY41.5 generated a stop codon at *SSK2* and was identified only in heterozygous state. The Sks2p protein interacts with Hog1p by direct phosphorylation to activate HOG pathway, which is a regulatory mechanism involved in the maintenance of cation homeostasis by signaling for the plasma membrane alkali metal cation transporters⁴⁵. Also, there is evidence that HOG pathway leads to the activation of regulatory response under high iron concentration⁴⁶.

Therefore, mutations in genes of the ISC system, especially *ISU1*, significantly contribute to increase the iron ion inside yeast cells. These results could be correlated with high performance in xylose consumption, considering that xylose isomerase is a metalloenzyme that uses divalent cations as cofactor such as Mg²⁺, Mn²⁺, Co²⁺, and iron cations (Fe²⁺), although the latter is not the preferred cation^{47–49}. According to the data displaced above, the SNPs probably led to an increase of cellular iron concentration, boosting the activity of xylose isomerase given that it acts as xylose isomerase cofactor (Fig. 6c). However, other possible consequences of changes in iron metabolism, besides the increased *xylA* activity, cannot be eliminated since Fe/S clusters are essential cofactors for several cellular processes³³. Supporting the initial model, we observed a surprising effect in increasing of xylose consumption by LVY34.4 and other strains when iron was supplemented to YNBX medium (Fig. 6a). Iron supplementation also improves xylose consumption of not evolved strains (Fig. 6b and Supplementary Fig. S6). This result suggests the use of this cofactor as a potential supplement to increase the fermentation efficiency and boost the economic viability of 2G-ethanol production.

In summary, our work proved that specific mutations in the *SSK2* and *ISU1* genes increase xylose fermentation and, consequently, ethanol production by engineered *S. cerevisiae* strains. Surprisingly, both genes converge to pathways that are not related to carbohydrate catabolism. These data shows that the engineering of *S. cerevisiae* strains to consume xylose and enable the production of 2G-ethanol are related to a not obvious complex metabolic network, which provides subsidy for many surprising discoveries in the field of second-generation ethanol.

Methods

Yeast strains and growth conditions. The *S. cerevisiae* strains used in this work are described in Table 1. Yeast sporulation and tetrad dissection were carried out using standard procedures⁵⁰. YP medium (10 g/L yeast extract, 20 g/L peptone) supplemented with 20 g/L D-glucose (YPD) was used for inoculum preparation and yeast cells were cultivated at 30 °C, 200 rpm. Transformants were selected in YNB medium (6.7 g/L yeast nitrogen base without amino acids, Difco), supplemented with 1 g/L drop-out without uracil or leucine (Sigma), 20 g/L glucose or xylose and 20 g/L agar. For counter-selection, the transformants harboring the pSH65⁵¹ plasmid were cultivated in YNB supplemented with 440 mg/L uracil and 1 g/L 5-FOA. For antibiotic resistance, 200 mg/L of geneticin and 200 mg/L of hygromycin was added to YPD for strains expressing the *KanMX4*⁵² or *hph*⁵³ markers, respectively. Likewise, 300 µg/ml of zeocin was added for strains transformed with pSH65⁵¹.

Plasmids. Expression cassettes were assembled in a single reaction⁵⁴, using the pRS304⁵⁵ or pRS426⁵⁶ vector linearized with *Bam*HI. The fragments were amplified using Phusion DNA polymerase (NEB) with the primers described in Table S2. *xylA* gene sequence from *Orpinomyces* sp. was codon optimized for expression in *S. cerevisiae* (Eurofins) (supplementary material). The other genes, promoters and terminators were amplified from *S. cerevisiae* strain LVA1. All cassettes use *URA3* as marker flanked by two *loxP* sites. The main components of the plasmids are summarized in Table S1.

Strains construction. The molecular techniques were performed using standard procedures⁵⁷. Amplicons were purified from agarose gel using the kit Wizard[®] SV Gel and PCR Clean-Up System (Promega). The expression cassettes were amplified by PCR and used to transform yeast cells using the LiAc/SS-DNA/PEG protocol⁵⁸. *URA3* gene was deleted from the LVA1 genome using *Kan-loxP* marker⁵² (G418 resistance), which was subsequently excised with pSH65 (*Cre-loxP* system)⁵¹. *GRE3* was deleted by replacement with *URA3* flanked by *loxP* sites and subsequent excised. The plasmid pOXylATy1 was cleaved with *Bam*HI and the resulting fragment was employed to transform LVY27, using *LEU2* as auxotrophic selection marker.

For functional analysis of mutations, the promoter and terminator sequences of *IKI3* and *ISU1* were fused to both ends of *URA3* and used to replace and delete the native *IKI3* and *ISU1* genes from LVY54's genome (LVY34.4, *ura3*Δ), resulting in the Ura⁺ strains LVY55 and LVY56, respectively (Table 1). The same procedure was used in the non-evolved cell LVY65 (LVY27, *ura3*Δ), to individually test the effect of each mutation. In this case, the plasmid pOXylA2 was inserted to simulate the high copy-number integration of *xylA*. For functional SNP validation, the *IKI3* and *ISU1* genes were amplified from the genomic DNA of LVY27, LVY34.4 and LVY41.5 and the purified PCR product was used to transform the evolved LVY55 and LVY56 and the non-evolved LVY65 strains to restore the genes at the same initial locus. The presence of each allele was confirmed by Sanger sequencing. *SSK2* was inactivated by insertion of hygromycin gene amplified from pAG32⁵³.

Adaptive evolution and selection of xylose-fermenting strains. Approximately 300 colonies derived from LVY27, transformed with the fragment obtained from the plasmid pOXylATy1, were inoculated in

semi-anaerobic conditions into a 100 mL sealed bottle with YNB medium with 40 g/L of xylose, and incubated at 30 °C/200 rpm. Since no growth was observed using xylose as sole carbon source, glucose was added to a concentration of 0.5%. After the first round of evolution, the yeast cells were separated into two parallel experiments (Fig. 1a). If the culture showed growth, glucose was no longer added and a 5 mL aliquot was transferred to a new flask after the culture reached an OD_{600nm} of approximately 3. Three and twelve sequential batch cultivations were carried out for LVY34.4 and LVY41.5, respectively. The same procedure was employed for the diploid LVY71, but without addition of glucose. Culture samples were streaked on plates with solid xylose medium. Colonies showing the largest diameter were evaluated. The fastest growing strains LVY34.4 and LVY41.5 were selected for further characterization.

Sequencing and genome assembly. Whole-genome sequencing of LVY27, LVY34.4 and LVY41.5 was performed using Illumina MiSeq at University of North Carolina (UNC, USA) sequencing facility. For each sample, millions of 2 × 300-bp paired-end reads were generated with an average size of 500 bp. The software PEAR⁵⁹ was used to merge the overlapping paired-end reads in a single sequence. For each strain, overlapped reads were assembled into longer contigs using VELVET assembler⁶⁰ with k-mer parameter optimized to maximize the N50 metric. The complete dataset of DNA-seq reads have been deposited in SRA under accession number SRP078504.

Copy Number Variation (CNV) and Single Nucleotide Polymorphism (SNPs) analysis. The reads from each strain were mapped into the reference genome (LVY27 assembly) using Bowtie2 aligner⁶¹. The resulting BAM files were ordered using Picard package⁶² and used as input for cn.MOPS program²⁷ to perform CNV analysis considering a window size of at least 1 Kb and a read coverage variation >2×. The command DepthOfCoverage available in the GATK program²⁸ was used to extract from BAM files the read-depth coverage for each base pair. These values were normalized by its median to represent the CNV.

The SNP/INDEL detection for each strain was performed by GATK program using default parameters for base quality recalibrator, haplotype caller and variant filtration steps that processed the BAM files and generated three VCF (variant call format) files. Homemade PERL scripts were developed to process and compare the VCF files identifying SNPs/INDELS that appeared in the evolved strains and were absent in LVY27. Resulting SNPs/INDELS were submitted for manual annotation using 150 bp flanking region of each SNP/INDEL, extracted from evolved genome assembly, and BLASTing against *S. cerevisiae* genome using SGD database⁶³.

Molecular karyotyping analysis. PFGE procedure for full karyotype resolution was conducted as previously described⁶⁴. DNA-embedded agarose plugs were loaded into 1% contour-clamped homogeneous electric field gels (CHEF) and gels were run in a BioRad CHEF-DR II system. Running conditions were as follows: 54 hours run time, 5 Volts/cm, initial and final switch times of 47 and 170 seconds, respectively. PFGE optimized for maximum resolution of molecules between 500 Kb and 800 Kb was conducted similarly, except for changes in total run time (84 hours), and initial and final switch times of 60 and 75 seconds, respectively. Slices were excised at sizes 580 Kb, 660 Kb and 700 Kb, purified using Thermo Scientific GeneJET Gel extraction Kit and the eluted DNA was used as template for qRT-PCR. The reactions were performed using PerfeCTa SYBR[®] Green Fast Mix (Quanta BioSciences) on a BioRad CFX Connect Real Time System obeying this conditions: 30 s at 95 °C; 3 s at 95 °C; 15 s at 55 °C; 10 s at 72 °C; optical detection, repeat 34 cycles of steps 2 through 5. The amplification kinetics of three DNA sequence targets were determined: a segment of Chr05 outside of the *xylA* insertion region, and segments used for normalization from Chr09 (450 Kb) and Chr16 (920 Kb). The ΔCt between the Chr05 and Chr09 target was determined for each sample as a measure of Chr05 DNA template enrichment (A); The ΔCt between Chr16 and Chr09 was determined as a measure of enrichment of unspecific DNA templates (B); The A/B enrichment ratio was calculated and used as a measure of the abundance of Chr05 DNA templates in the slices relative to the abundance of unspecific DNA templates. Determinations were made in triplicate. The A/B ratios were averaged and the standard error calculated (Fig. 3c).

Ploidy determination by flow cytometry. Diploidization of LVY41.5 was confirmed by flow cytometry, according described in elsewhere³⁰. PE-2 and LVY34.4 were used as diploid and haploid controls, respectively. Exponentially growing cells were fixed in ethanol 70%. Cells were treated with RNase 2 mg.ml⁻¹ and DNA was stained with propidium iodide 0.5 mg/ml in PBS buffer and incubated overnight. Fluorescence was analyzed with Flow cytometer FACS CANTO II.

Enzyme assay. LVY34.4 cells grown in YPD medium and collected from mid-log phase were washed with sterile ice-cold water, centrifuged and resuspended in Y-PER reagent (Thermo Fisher Scientific). Xylose isomerase activity in LVY34.4 cell extracts was determined as described elsewhere⁶⁵ adapted for microplate assay, using 100 mM Tris-HCl (pH 7.5) buffer, 0.001 mM FeSO₄·7H₂O, 0.15 mM NADH, 2 U sorbitol dehydrogenase (Sigma-Aldrich), and the cell lysate. The reaction started with the addition of 500 mM xylose. The activity (U/L) was determined by monitoring the oxidation of NADH at 340 nm during 15 minutes at 30 °C. One unit equals 1 μmol of substrate converted per minute under the conditions of the assay.

Fermentation conditions. Semi-anaerobic batch fermentations were performed in YP medium supplemented with 30 g/L or 50 g/L of xylose as sole carbon source and the culture was started with a dry cell weight of about 0.25 g DCW/L. For functional analysis of SNPs, fermentations were carried out in YNB medium (without uracil) supplemented with 20 g/L of xylose, with initial cell weight of 0.12 g DCW/L. For iron supplementation tests, FeSO₄ was added at a concentration of 100 μM at YNBX medium. For magnesium and manganese supplementation tests, MgSO₄ and MnSO₄ were added in concentrations of 500 μM and 100 μM, respectively. The fermentations were performed in 100 mL sealed bottles with working volume of 80 mL and incubated at

30 °C/200 rpm. Samples were taken to measure OD and subsequent analysis by HPLC. For cell dry weight determination, cells were collected by centrifugation, washed twice and dried in 65 °C. All experiments were performed in triplicate.

Analytical procedures. Quantification of glucose, xylose, xylitol, glycerol, acetic acid and ethanol, was carried out by high-performance liquid chromatography (HPLC) using chromatograph Alliance (Waters) with refractive index detector (Waters 2414) and photodiode array detector (Waters 2998) at 280 nm. The samples were analyzed by HPLC-RI-PDA using ion exclusion HPX-87H column (300 mm × 7,8 mm, BioRad®), heated in an oven at 35 °C, a 5 mM H₂SO₄ solution as the mobile phase at a flow 0.6 mL/min. A standard curve with known concentrations of compounds of interest was also analyzed using the same procedure.

References

- Smith, P. *et al.* Biophysical and economic limits to negative CO₂ emissions. *Nat. Clim. Chang.* **6**, 42–50 (2015).
- Tavoni, M. *et al.* Post-2020 climate agreements in the major economies assessed in the light of global models. *Nat. Clim. Chang.* **5**, 119–126 (2015).
- Peters, G. P. *et al.* The challenge to keep global warming below 2 °C. *Nat. Clim. Chang.* **3**, 1–3 (2012).
- van Vuuren, D. P. *et al.* The role of negative CO₂ emissions for reaching 2 °C - insights from integrated assessment modelling. *Clim. Change* **118**, 15–27 (2013).
- Santos, L. V. *et al.* Second-Generation Ethanol: The Need is Becoming a Reality. *Ind. Biotechnol.* **12**, 40–57 (2016).
- Della-Bianca, B. E., Basso, T. O., Stambuk, B. U., Basso, L. C. & Gombert, A. K. What do we know about the yeast strains from the Brazilian fuel ethanol industry? *Appl. Microbiol. Biotechnol.* **97**, 979–991 (2013).
- Basso, L. C., De Amorim, H. V., De Oliveira, A. J. & Lopes, M. L. Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res.* **8**, 1155–1163 (2008).
- Argueso, J. L. *et al.* Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Res.* **19**, 2258–2270 (2009).
- Argueso, J. L. & Pereira, G. A. G. Perspective: Indigenous sugarcane yeast strains as ideal biological platforms for the delivery of next generation biorefining technologies. *Int. Sugar J.* **112**, 86–89 (2010).
- Cai, Z., Zhang, B. & Li, Y. Engineering *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: Reflections and perspectives. *Biotechnol. J.* **7**, 34–46 (2012).
- Kim, S. R., Park, Y. C., Jin, Y. S. & Seo, J. H. Strain engineering of *Saccharomyces cerevisiae* for enhanced xylose metabolism. *Biotechnol. Adv.* **31**, 851–861 (2013).
- Van Vleet, J. H. & Jeffries, T. W. Yeast metabolic engineering for hemicellulosic ethanol production. *Curr. Opin. Biotechnol.* **20**, 300–306 (2009).
- Karhumaa, K., Garcia Sanchez, R., Hahn-Hägerdal, B. & Gorwa-Grauslund, M.-F. Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **6**, 5 (2007).
- Li, X. L., Park, A., Estrela, R., Kim, S., Jin, Y. & Cate, J. H. D. Comparison of xylose fermentation by two high-performance engineered strains of *Saccharomyces cerevisiae*. *Biotechnol. Rep.* **9**, 53–56 (2016).
- Walfridsson, M. *et al.* Ethanol fermentation of xylose with *Saccharomyces cerevisiae* harboring the *Thermus thermophilus xylA* gene, which expresses an active xylose (glucose) isomerase. *Appl. Environ. Microbiol.* **62**, 4648–4651 (1996).
- Kuyper, M. *et al.* High-level functional expression of a fungal xylose isomerase: The key to efficient ethanol fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res.* **4**, 69–78 (2003).
- Kuyper, M. *et al.* Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res.* **5**, 399–409 (2005).
- Reider, A. A., Ouellet, M., Szmids-Middleton, H., Keasling, J. D. & Mukhopadhyay, A. Evolved hexose transporter enhances xylose uptake and glucose/xylose co-utilization in *Saccharomyces cerevisiae*. *Sci. Rep.* **6**, 19512 (2016).
- Sonderegger, M. & Sauer, U. Evolutionary engineering of *Saccharomyces cerevisiae* for anaerobic growth on xylose. *Appl. Environ. Microbiol.* **69**, 1990–1998 (2003).
- Çakar, Z. P., Turanli-Yildiz, B., Alkim, C. & Yilmaz, U. Evolutionary engineering of *Saccharomyces cerevisiae* for improved industrially important properties. *FEMS Yeast Res.* **12**, 171–182 (2012).
- Winkler, J. D. & Kao, K. C. Recent advances in the evolutionary engineering of industrial biocatalysts. *Genomics* **104**, 406–411 (2014).
- Demeke, M. M. *et al.* Development of a D-xylose fermenting and inhibitor tolerant industrial *Saccharomyces cerevisiae* strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnol. Biofuels* **6**, 89 (2013).
- Diao, L. *et al.* Construction of fast xylose-fermenting yeast based on industrial ethanol-producing diploid *Saccharomyces cerevisiae* by rational design and adaptive evolution. *BMC Biotechnol.* **13**, 110 (2013).
- Lee, S.-M., Jellison, T. & Alper, H. S. Systematic and evolutionary engineering of a xylose isomerase-based pathway in *Saccharomyces cerevisiae* for efficient conversion yields. *Biotechnol. Biofuels* **7**, 122 (2014).
- Yin, Y. & Petes, T. D. Genome-Wide High-Resolution Mapping of UV-Induced Mitotic Recombination Events in *Saccharomyces cerevisiae*. *PLoS Genet.* **9** (2013).
- Traff, K., Otero Cordero, R., Van Zyl, W. H. & Hahn-Hägerdal, B. Deletion of the *GRE3* aldose reductase gene and its influence on xylose metabolism in recombinant strains of *Saccharomyces cerevisiae* expressing the *xylA* and *XKS1* genes. *Appl. Environ. Microbiol.* **67**, 5668–5674 (2001).
- Klambauer, G. *et al.* Cn.MOPS: Mixture of poisson for discovering copy number variations in next-generation sequencing data with a low false discovery rate. *Nucleic Acids Res.* **40**, 1–14 (2012).
- McKenna, A. *et al.* The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
- Illuxley, C., Green, E. D. & Dunbam, I. Rapid assessment of *S. cerevisiae* mating type by PCR. *Trends Genet.* **6**, 236 (1990).
- Delobel, P. & Tesnière, C. A simple FCM method to avoid misinterpretation in *Saccharomyces cerevisiae* cell cycle assessment between G0 and sub-G1. *PLoS One* **9**, 1–8 (2014).
- Rodríguez-Peña, J. M., García, R., Nombela, C. & Arroyo, J. The high-osmolarity glycerol (HOG) and cell wall integrity (CWI) signalling pathways interplay: a yeast dialogue between MAPK routes. *Yeast* **27**, 495–502 (2010).
- Garland, S. A., Hoff, K., Vickery, L. E. & Culotta, V. C. *Saccharomyces cerevisiae* *ISU1* and *ISU2*: members of a well-conserved gene family for iron-sulfur cluster assembly. *J. Mol. Biol.* **294**, 897–907 (1999).
- Rouault, T. A. & Tong, W.-H. Iron-sulphur cluster biogenesis and mitochondrial iron homeostasis. *Nat. Rev. Mol. Cell Bio.* **6**, 345–351 (2005).
- Carvalho-Netto, O. V. *et al.* *Saccharomyces cerevisiae* transcriptional reprogramming due to bacterial contamination during industrial scale bioethanol production. *Microb. Cell Fact.* **30**, 13:14 (2015).

35. Madhavan, A. *et al.* Xylose isomerase from polycentric fungus *Orpinomyces*: Gene sequencing, cloning, and expression in *Saccharomyces cerevisiae* for bioconversion of xylose to ethanol. *Appl. Microbiol. Biotechnol.* **82**, 1067–1078 (2009).
36. Guerra, O. G. *et al.* A novel system of genetic transformation allows multiple integrations of a desired gene in *Saccharomyces cerevisiae* chromosomes. *J. Microbiol. Methods* **67**, 437–445 (2006).
37. Demeke, M. M., Foulquié-Moreno, M. R., Dumortier, F. & Thevelein, J. M. Rapid evolution of recombinant *Saccharomyces cerevisiae* for xylose fermentation through formation of extra-chromosomal circular DNA. *PLoS Genet.* **11**, 1–21 (2015).
38. Zhou, H., Cheng, J., Wang, B. L., Fink, G. R. & Stephanopoulos, G. Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*. *Metab. Eng.* **14**, 611–622 (2012).
39. Manicki, M. *et al.* Overlapping binding sites of the frataxin homologue assembly factor and the heat shock protein 70 transfer factor on the Isu iron-sulfur cluster scaffold protein. *J. Biol. Chem.* **289**, 30268–30278 (2014).
40. Lill, R. *et al.* The role of mitochondria in cellular iron-sulfur protein biogenesis and iron metabolism. *Biochim. Biophys. Acta* **1823**, 1491–1508 (2012).
41. Rouault, T. A. *Iron-Sulfur Clusters in Chemistry and Biology*. Walter De Gruyter Inc. (2014).
42. Adam, A. C., Bornhövd, C., Prokisch, H., Neupert, W. & Hell, K. The Nfs1 interacting protein Isd11 has an essential role in Fe/S cluster biogenesis in mitochondria. *EMBO J.* **25**, 174–183 (2006).
43. Wiedemann, N. *et al.* Essential role of Isd11 in mitochondrial iron-sulfur cluster synthesis on Isu scaffold proteins. *EMBO J.* **25**, 184–195 (2006).
44. Vickery, L. E. & Cupp-Vickery, J. R. Molecular chaperones HscA/Ssq1 and HscB/Jac1 and their roles in iron-sulfur protein maturation. *Crit. Rev. Biochem. Mol. Biol.* **42**, 95–111 (2007).
45. Arino, J., Ramos, J. & Sychrova, H. Alkali metal cation transport and homeostasis in yeasts. *Microbiol. Mol. Biol. Rev.* **74**, 95–120 (2010).
46. Kaba, H. E. J., Nimtz, M., Müller, P. P. & Bilitewski, U. Involvement of the mitogen activated protein kinase Hog1p in the response of *Candida albicans* to iron availability. *BMC Microbiol.* **24**, 13–16 (2013).
47. Kovalevsky, A. Y. *et al.* Metal ion roles and the movement of hydrogen during reaction catalyzed by D-xylose isomerase: a joint X-Ray and neutron diffraction study. *Structure* **18**, 688–699 (2010).
48. Chen, W.-P. Glucose isomerase (a review). *Process Biochem.* **15**, 36–41 (1980).
49. Danno, G. I. Studies on D-glucose isomerizing enzyme from *Bacillus coagulans*, strain HN-68. *Agric. Biol. Chem.* **34**, 1805–1814 (1970).
50. Ausubel, F. M. *et al.* *Current Protocols in Molecular Biology*. John Wiley & Sons Inc; ringbou edition (2003).
51. Gueldener, U., Heinisch, J., Koehler, G. J., Voss, D. & Hegemann, J. H. A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.* **30**, e23 (2002).
52. Wach, A., Brachat, A., Pohlmann, R. & Philippsen, P. New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* **10**, 1793–1808 (1994).
53. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
54. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
55. Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989).
56. Christianson, T. W., Sikorski, R. S., Dante, M., Shero, J. H. & Hieter, P. Multifunctional yeast high-copy-number shuttle vectors. *Gene* **110**, 119–122 (1992).
57. Sambrook, J., Fritsch, E. & Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring harbor Laboratory Press, Cold Spring Harbor. (1989).
58. Gietz, R. D. & Schiestl, R. H. Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 38–41 (2007).
59. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620 (2014).
60. Zerbino, D. R. & Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
61. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
62. Broad Institute – Picard. <https://broadinstitute.github.io/picard/>. Release 1.123 (2014).
63. Engel, S. R. *et al.* The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3 (Bethesda)* **4**, 389–398 (2014).
64. Zhang, H. *et al.* Gene copy-number variation in haploid and diploid strains of the yeast *Saccharomyces cerevisiae*. *Genetics* **193**, 785–801 (2013).
65. Hector, R. E., Dien, B. S., Cotta, M. A. & Mertens, J. A. Growth and fermentation of D-xylose by *Saccharomyces cerevisiae* expressing a novel D-xylose isomerase originating from the bacterium *Prevotella ruminicola* TC2-24. *Biotechnol. Biofuels* **6**, 84 (2013).

Acknowledgements

We thank BNDES and FINEP for financial support. NMVS received a pre-doctoral fellowship from Brazil's CAPES; Research reported here was supported in part by the National Institute of General Medical Sciences of the USA National Institutes of Health (NIGMS-NIH) under Maximizing Investigator's Research Award (MIRA) number R35GM119788 to JLA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author Contributions

L.V.S., J.L.A. and G.A.G.P. designed experiments. L.V.S., L.D.A., G.B. and T.L.R.C. performed experiments. M.F.C. and S.T.N. performed bioinformatics analysis. R.A.S.P. performed analytical procedures. N.M.V.S. performed PFGE analysis. L.V.S., M.F.C., N.M.V.S., T.L.R.C., J.L.A. and G.A.G.P. wrote and revised the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Santos, L. V. *et al.* Unraveling the genetic basis of xylose consumption in engineered *Saccharomyces cerevisiae* strains. *Sci. Rep.* **6**, 38676; doi: 10.1038/srep38676 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016