



UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE BIOLOGIA

João Gabriel Ribeiro Bueno

**Genetic engineering of *Saccharomyces cerevisiae* for  
co-fermentation of xylose and glucose to second-  
generation ethanol production**

**Engenharia genética de *Saccharomyces cerevisiae*  
para co-fermentação de xilose e glicose e produção de  
etanol de segunda geração**

Campinas

2020

**João Gabriel Ribeiro Bueno**

**Genetic engineering of *Saccharomyces cerevisiae* for co-fermentation of xylose and glucose to second-generation ethanol production**

**Engenharia genética de *Saccharomyces cerevisiae* para co-fermentação de xilose e glicose e produção de etanol de segunda geração**

*Dissertation presented to the Institute of Biology of the University of Campinas in partial fulfillment of the requirements for the degree of MSc in Genetics and Molecular Biology in microorganisms genetics area*

*Dissertação apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do Título Mestre em Genética e Biologia Molecular na área de Genética de Microrganismos.*

*Orientador:* Prof. Dr. Leandro Viera dos Santos

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO JOÃO GABRIEL RIBEIRO BUENO E ORIENTADO PELO PROF. DR. LEANDRO VIEIRA DOS SANTOS.

**CAMPINAS**

**2020**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Biologia  
Mara Janaina de Oliveira - CRB 8/6972

B862g Bueno, João Gabriel Ribeiro, 1995-  
Genetic engineering of *Saccharomyces cerevisiae* for co-fermentation of xylose and glucose to second-generation ethanol production / João Gabriel Ribeiro Bueno. – Campinas, SP : [s.n.], 2020.

Orientador: Leandro Vieira dos Santos.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. *Saccharomyces cerevisiae*. 2. Repressão catabólica. 3. Açúcar - Transporte. 4. Sistema CRISPR/Cas9. I. Santos, Leandro Vieira dos, 1982-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Engenharia genética de *Saccharomyces cerevisiae* para co-fermentação de xilose e glicose e produção de etanol de segunda geração

**Palavras-chave em inglês:**

*Saccharomyces cerevisiae*

Catabolite repression

Sugar - Transportation

CRISPR/Cas9 system

**Área de concentração:** Genética de Microorganismos

**Titulação:** Mestre em Genética e Biologia Molecular

**Banca examinadora:**

Leandro Vieira dos Santos [Orientador]

Elizabeth Bilsland

Jeferson Gross

**Data de defesa:** 17-02-2020

**Programa de Pós-Graduação:** Genética e Biologia Molecular

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0003-4873-4444>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3222572016134389>

Campinas, 17 de fevereiro de 2020.

## **COMISSÃO EXAMINADORA**

Prof. Dr. Leandro Vieira dos Santos

Prof. Dr. Jeferson Gross

Profa. Dra. Elizabeth Bilsland

*Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.*

*A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa de Pós graduação em Genética e Biologia Molecular da Unidade Instituto de Biologia.*

## **Agradecimentos**

Agradeço toda a minha família principalmente meus pais Denise e Herlindo, minha irmã Maria Paula e meu tio Eduardo pelas conversas e apoio ao longo desses anos.

A Rhaissa pelo carinho, companhia, paciência e compreensão nos momentos de maiores dificuldades. E para Thiago, Gustavo, Mateus, Antônio, Guilherme, Diego, Bruna, Luís, Naiara, Elisabete e Ubiratan por ajudarem a fazer muito mais leve os caminhos mais desafiadores.

Aos meus colegas de grupo Gisele, Aline, Natália, Bruno, Paulo, Thiago, Rafael e Davi, a Thamy e aos demais membros do Laboratório Nacional de Biorrenováveis (LNBR) pelas conversas, discussões e imenso aprendizado.

Agradeço ao financiamento das respectivas agências na condução do meu projeto: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) - 2017/08519-6 e 2018/00888-5, Conselho Nacional de Desenvolvimento Científico e Tecnológico - 430291/2018-3) e Instituto Serrapilheira - 1708-16205. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001.

Finalmente, o meu profundo agradecimento ao meu orientador Leandro Vieira dos Santos pela oportunidade e, principalmente, pela confiança demonstrada. Levarei com muito apreço nossas longas conversas e seus ensinamentos no decorrer da minha vida profissional e pessoal.

## Resumo

Nas últimas décadas, a comunidade mundial manifestou uma intensa preocupação relacionada às consequências causadas de compostos, como petróleo, gás e carvão, que estão relacionados com o aumento dos gases causadores do efeito estufa. O etanol de segunda-geração (2G), produzido a partir biomassa lignocelulósica, surge com um grande potencial na produção de energia limpa a partir de uma matriz renovável. O desenvolvimento de linhagens modificadas da levedura *Saccharomyces cerevisiae*, capazes de converter em etanol o mais abundante açúcar de cinco carbonos xilose (C5) da biomassa, é um passo essencial na viabilização da tecnologia 2G. Um grande é a baixa eficiência de assimilação de xilose por transportadores de membrana de *S. cerevisiae* e a repressão catabólica, levando a inibição por glicose (C6). Transportadores capazes de assimilar xilose são inibidos em altas concentrações desse açúcar, além de apresentarem maior afinidade em glicose, levando a uma competição entre os substratos e estendendo o tempo da fermentação. A xilose somente é assimilada pela levedura quando as concentrações de glicose estão baixas no ambiente extracelular. Dessa forma, este projeto teve como objetivo a caracterização de novos transportadores heterólogos de xilose eficientes e o desenvolvimento de uma linhagem capaz de co-fermentar xilose e glicose.

Neste trabalho, identificamos e caracterizamos quatro novos transportadores de xilose Cs186, Cs2608, Cs3894 e Cs4130, avaliados na linhagem EBY.VW4000 com a via de consumo de xilose integrada no genoma. Curiosamente, o transportador Cs4130 demonstrou superior capacidade de assimilação de xilose em altas concentrações (50 g/L) em comparação com o controle *GXF1*, descrito o melhor transportador heterólogo na literatura. Nessa condição, *GXF1* foi severamente inibido, oposto ao observado por Cs4130. O modelo estrutural de Cs4130 em comparação com o transportador procariótico de alta afinidade *XylE* e *GXF1* aponta resíduos em sua arquitetura molecular que possam explicar as diferenças de comportamento observadas entre os transportadores. Dessa forma, o novo transportador Cs4130 é apresentado como um excelente candidato para engenharia genética de linhagens na indústria 2G.

Entretanto, os transportadores de xilose Cs4130 e *GXF1* ainda apresentam uma forte inibição pela glicose na mistura dos dois açúcares. A segunda parte do projeto visa a modificação dessas proteínas de membrana e o desenvolvimento de um processo para o aumento da afinidade a xilose em relação à glicose. Assim, desenvolvemos duas novas plataformas de evolução adaptativa, *EBY\_Xyl1\_hxk<sup>0</sup>* derivada da linhagem *EBY\_Xyl1*, e a linhagem *JBY\_hxk<sup>0</sup>* construída a partir de uma cepa industrial fermentadora de xilose. Estas plataformas não são capazes de metabolizar moléculas de glicose devido as deleções dos genes das hexoquinases *HXK1*, *HXK2* e *GLK1*, responsáveis pela primeira etapa da via glicolítica. As linhagens modificadas irão reconhecer as moléculas de glicose apenas como inibidores ao metabolismo da xilose. Por fim, a identificação de novas mutações presentes nas células evoluídas derivadas das plataformas construídas, capazes de co-fermentar xilose e glicose, visam identificar novos alvos moleculares relacionados a repressão catabólica. As informações desenvolvidas nesse projeto irão auxiliar no desenvolvimento de linhagens de *S. cerevisiae* capazes de fermentar simultaneamente xilose e glicose, reduzindo o tempo final de fermentação na indústria 2G.

## Abstract

The global energy matrix is based on fossil fuels, such as oil, gas and coal. In recent decades, the world community has expressed an intense concern related to the consequences caused by the use of these compounds, such as the increase in greenhouse gas emissions. Second-generation (2G) ethanol, produced from lignocellulosic biomass, emerges as promise sustainable advanced biofuel. The development of engineered *Saccharomyces cerevisiae* yeast strains, capable of converting xylose, the second most abundant sugar on lignocellulosic biomass, into ethanol, is an essential step to a feasible 2G process. A major bottleneck in the conversion of bioproducts from these biomass-derived sugars is the low xylose uptake by *S. cerevisiae* membrane transporters and catabolic repression, leading to glucose inhibition (C6). Membrane proteins capable to transport xylose are strongly inhibited on high concentrations of this sugar, in addition to presenting low affinity to xylose compared to glucose, leading to competition between substrates and extending the fermentation time. Xylose is only assimilated by *S. cerevisiae* cells when glucose concentrations are low in the extracellular environment. Therefore, this project aimed to identify novel and efficient heterologous xylose transporters and to develop a strain capable of co-fermenting xylose and glucose. In this work, we identified and characterized four new xylose transporters Cs186, Cs2608, Cs3894 and Cs4130, evaluated in EBY.VW4000 strain modified with the xylose consumption pathway integrated into the genome. Curiously, the Cs4130 transporter demonstrated a superior capacity of xylose assimilation at high xylose concentrations (50 g/L) compared to the *GXF1* control, considered the best heterologous transporter previously described in the literature. In the condition described, *GXF1* was severely inhibited, demonstrating an opposite behavior compared to that observed by Cs4130. The structural model of Cs4130 compared to the high affinity prokaryotic transporter XylE and *GXF1* points out residues in its molecular architecture that may explain the differences in behavior observed between the transporters. The novel transporter Cs4130 is presented as an excellent candidate for genetic engineering of *S. cerevisiae* strains for 2G biocompounds production.

However, the xylose transporters Cs4130 and *GXF1* show strong glucose inhibition on mixtures of both sugars. Therefore, the second part of the project



aims to engineering these membrane proteins and to develop a process to increase the affinity of xylose compared to glucose. We have developed two new microbial platforms for adaptive evolution, EBY\_Xyl1\_ *hvk*<sup>0</sup>, derived from EBY\_Xyl1 strain, and the JBY\_ *hvk*<sup>0</sup> constructed from an industrial high-yield pentose fermenting strain. These new platforms are unable to metabolize glucose molecules due to the disruption of glycolysis pathway by deletion of hexokinase genes, *HVK1*, *HVK2* and *GLK1*. The *S. cerevisiae* modified strains recognized glucose molecules only as an inhibitor to xylose metabolism. Therefore, new genetic mutations present in these evolved cells, capable to co-consume xylose and glucose, will provide new data to understand the genetic basis of glucose repression. The information produced in this project will help the development of *S. cerevisiae* strains suitable of simultaneously fermenting xylose and glucose, reducing the total fermentation time in the 2G industry.

## Sumário

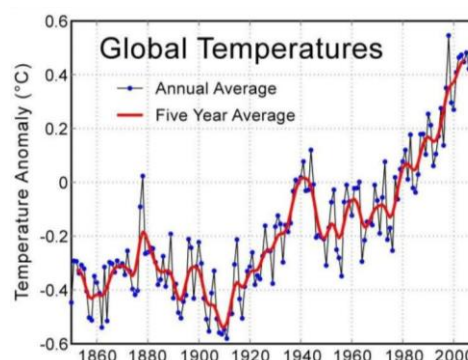
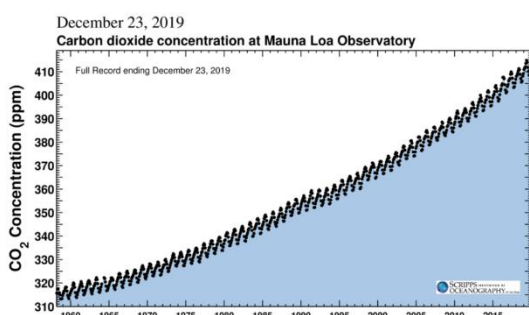
1. Introdução Geral .....	11
1.1. Cenário energético mundial e alternativas .....	11
1.2. Biomassa lignocelulósica .....	12
1.3. Industria 2G.....	14
1.3.1. Pré-tratamento e sacarificação.....	14
1.3.2. Plataformas microbianas.....	17
1.3.2.1. Via de Oxiredução.....	18
1.3.2.2. Via de isomerização.....	19
1.3.2.3. Transporte dos açúcares da biomassa.....	21
1.4. Desafios do processo fermentativo 2G.....	25
Capítulo 1	
Novel xylose transporter Cs4130 expands the range of sugar uptake at high xylose concentrations in recombinant <i>Saccharomyces cerevisiae</i> strains.....	27
Capítulo 2	
Design e construção de linhagens modificadas 2G de <i>Saccharomyces cerevisiae</i> para fermentação simultânea de xilose e glicose.....	67
2. Referências.....	99
Anexo I – Termo de autorização.....	115
Anexo II – Declaração de Bioética/Biossegurança.....	116

## 1. Introdução geral

### 1.1. Cenário energético mundial e alternativas

Desde a Revolução Industrial, a matriz energética mundial é dependente do uso de fontes fósseis como petróleo, carvão e gás. Com o passar dos anos, o aumento da população mundial e as mudanças do padrão de vida moderna intensificaram a necessidade do uso de recursos naturais<sup>1-3</sup>. Visando suprir o aumento dessa demanda energética, as fontes fósseis de energia tem sido utilizadas em quantidades cada vez maiores, intensificando as concentrações de poluentes na atmosfera, como o dióxido de carbono, cuja concentração ultrapassou os 400 ppm na última década<sup>4,5</sup> (Figura 1). Além disso, o alto custo de exploração do petróleo é um grande fator de risco para o mercado internacional, visto que esse recurso é o principal componente da matriz energética atual e qualquer variação significativa pode gerar grandes crises econômicas<sup>6</sup>. Dessa forma, alternativas para a mudança dessa matriz energética poluente e obtenção de energias renováveis e limpas são cada vez mais necessárias para contrapor a alta taxa de emissão desses poluentes.

Latest CO<sub>2</sub> reading: **412.15 ppm**



**Figura 1** | Gráficos da correlação entre o aumento da temperatura global (A) e aumento da concentração de CO<sub>2</sub> na atmosfera (B) <http://www.global-warming-and-heclimate.com/greenhouse-warming-argument.html> e Scripps Institutions of Oceanography 2016.

O primeiro relato de um encontro que abordou de forma global a questão de mudanças climáticas remete a década de 70 com a “*World Climate Conference*”, em 1979<sup>7-9</sup>. Essa conferência inspirou novos encontros e congressos, até que, em 1988, foi criado o IPCC (*Intergovernmental Panel on Climate Change*). O IPCC é uma organização global que regulamenta e promove relatórios científicos sobre as mudanças climáticas, que auxiliam na compreensão de seus impactos e riscos futuros, assim como aponta possíveis

alternativas que revertam tal cenário. Além disso, encontros entre nações, como a Conferência das Partes (COP), buscam um acordo em comum entre países para limitar as consequências no futuro.

Recentemente na COP21, realizada em Paris no ano de 2015, cada nação apresentou uma projeção de ação para mitigar os avanços das emissões de gases do efeito estufa. O Brasil se comprometeu em diminuir em 37% as suas emissões de carbono em 2025 e 43% em 2030. Para atingir essas metas, umas das propostas foi a de aumentar o uso de fontes renováveis na matriz energética brasileira, principalmente na área de biocombustíveis. A meta é de que até 2030, 18% da matriz energética seja composta por biodiesel e biocombustíveis<sup>10</sup>.

O etanol é uma promissora alternativa para atender o mercado mundial<sup>11,12</sup>. No Brasil, sua produção como combustível automotivo começou na década de 70, com o programa governamental denominado Proálcool como alternativa à crise global do petróleo em 1973<sup>13</sup>. A forma tradicional de produção de etanol que esse programa incentivava era composta pela fermentação do caldo de cana-de-açúcar, resultando no etanol de primeira-geração (1G). Em conjunto com um grande incentivo produtivo, a criação de carros movidos exclusivamente com o etanol foi um importante passo para o aumento da produção do combustível<sup>14</sup>.

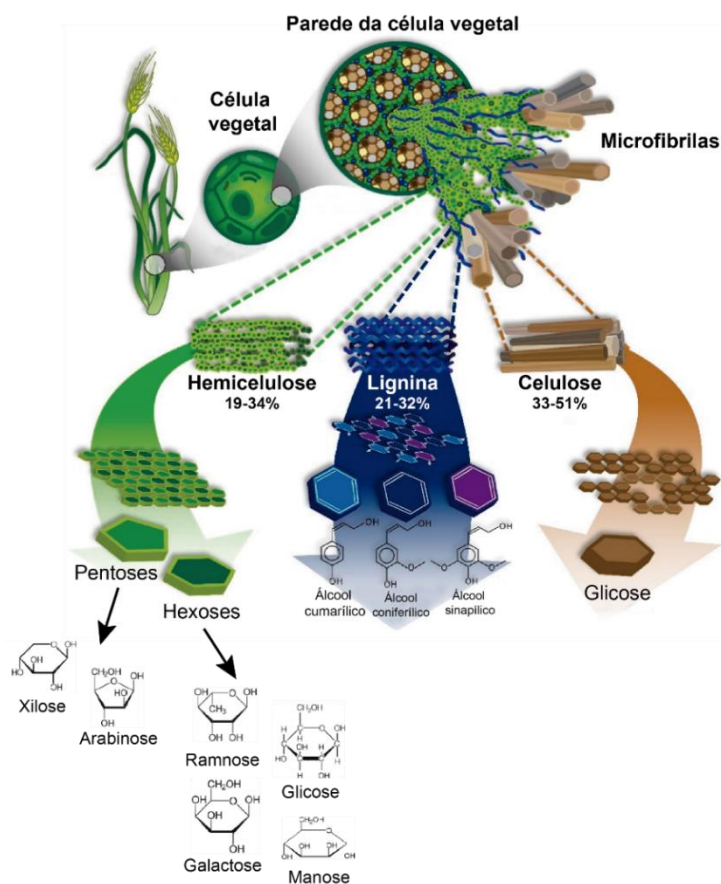
Diversos fatores estão relacionados ao aumento do custo da produção de etanol, como a área de plantio, a variação climática e o preço de insumos agrícolas<sup>15,16</sup>. O destino do melaço fica dividido entre a indústria alimentícia, para a produção de açúcar, e a indústria energética, para a produção de etanol, o que influencia seu custo final<sup>14,17</sup>. Dessa forma, com o intuito de aumentar a produção de etanol por área plantada, um novo processo fermentativo começou a ser desenvolvido através da utilização dos açúcares provenientes da parede celular das células vegetais, sendo a principal matéria-prima desse novo processo a biomassa lignocelulósica (e.g. bagaço e palha de cana)<sup>18</sup>. O etanol produzido através dessa biomassa é denominado de etanol de segunda-geração (2G).

## **1.2. Biomassa lignocelulósica**

A biomassa lignocelulósica apresenta uma complexa arquitetura molecular em sua parede vegetal, sendo composta por polímeros de

carboidratos que compõem a celulose e hemicelulose, e polímeros de monolignóis, representados pela lignina<sup>19,20</sup>. A porcentagem de cada componente desses polímeros da biomassa varia de acordo com espécie vegetal utilizada<sup>21</sup>.

A estrutura da lignina é composta tridimensionalmente por um polímero heterogêneo hidrofóbico composto por três monolignóis, álcool coniferílico, álcool cumarílico e álcool sinapílico<sup>20</sup>. Essa porção da biomassa, além de ser um importante componente estrutural para a planta, auxilia a célula vegetal em situações de stress e contra ação de patógenos<sup>22</sup>. A celulose é um polímero homogêneo, composto por glicose, um açúcar de seis carbonos (C6), que representa a porção mais abundante da biomassa. Por fim, a hemicelulose, um polímero composto por diversos açúcares como hexoses: glicose, manose, galactose e raminose, e açúcares de cinco carbonos (C5), como as pentoses: arabinose e xilose, sendo essa última a pentose mais abundante nos hidrolisados de segunda-geração<sup>19</sup> (figura 2). Assim, são nas porções da celulose e hemicelulose que os açúcares estruturais são utilizados em etapas posteriores para fermentação e conversão em compostos biorrenováveis.



**Figura 2 | Arquitetura dos componentes da biomassa lignocelulósica.** As células da parede vegetal das plantas são ricas em polímeros que compõem as porções da hemicelulose, lignina e a celulose. A lignina é constituída por três compostos, álcool coniferílico, álcool cumarílico e álcool sinapílico. A celulose é constituída por polímeros de glicose e a hemicelulose possui uma diversidade de monômeros de açúcares, sendo as hexoses apresentadas pela glicose, manose, raminose e galactose, e as pentoses representadas pela xilose e arabinose (modificado de Hernández-Beltrán<sup>21</sup>).

## 1.3 Indústria 2G

### 1.3.1 Pré-tratamento e sacarificação

Para que ocorra a fermentação a partir dos açúcares presentes na biomassa, são necessários processos anteriores de pré-tratamento e sacarificação, a fim de disponibilizar esses polímeros em monômeros, para que sejam utilizados por plataformas microbianas na conversão em bioprodutos. Esses processos são de extrema importância para a dinâmica da produção do etanol 2G e outros compostos de valor econômico, e estão diretamente relacionados à eficiência da fermentação.

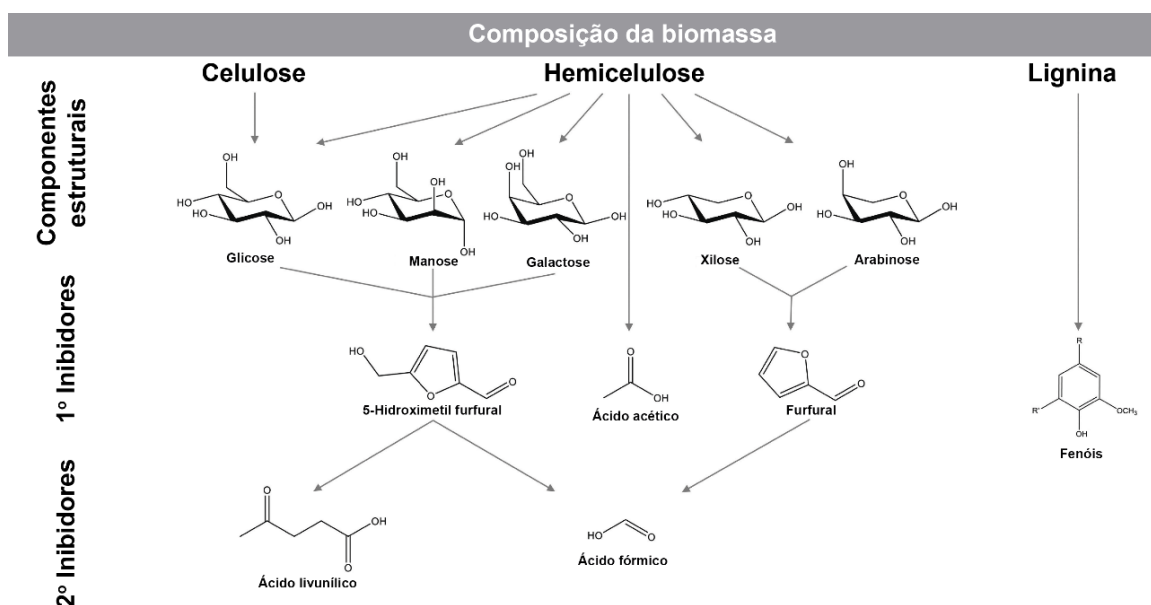
No processo de pré-tratamento, existem algumas estratégias que podem ser utilizadas para auxiliar na desconstrução da biomassa e que podem ser divididas em quatro tipos de classes: tratamento físico, químico, físico-químico e biológico (Tabela 1). Dentre todos os processos de pré-tratamento, a explosão a vapor é um processo físico-químico considerado economicamente mais eficiente<sup>19,23</sup>. Nesse processo a biomassa é submetida à uma condição de alta temperatura e pressão, expondo as fibras de celulose e solubilizando parte da hemicelulose. Contudo, essa solubilização da hemicelulose pode gerar muitos inibidores que afetam o crescimento de microrganismos durante o processo fermentativo<sup>24</sup> (Figura 3). Modificações nesse processo demonstraram menor produção de inibidores e maiores concentrações de açúcares recuperados<sup>19,23,25,26</sup>.

**Tabela 1 | Diferentes tipos de pré-tratamentos utilizados para biomassa lignocelulósica.**

Classe	Tipo de processo	Descrição	Referência
Físico	"Milling"	Tratamento físico de moagem da biomassa para gerar menores frações, aumentando a área de contato na etapa de hidrólise. Em alguns materiais lignocelulósicos pode ser necessário mais ciclos, o que se torna economicamente inviável.	19,27–29
	Pirólise	Procedimento que utiliza altas temperaturas, maiores que 300°C, para decomposição da celulose em compostos voláteis e uma espécie de carvão ( <i>char</i> ). Sendo esse carvão retirado com auxílio de um ácido aquoso gerando uma solução composta basicamente por glicose.	19,30,31

	Micro-ondas	Tratamento que utiliza ondas eletromagnéticas que interagem com material orgânico e geram uma grande quantidade de calor. Esse calor localizado é responsável pela desconstrução da biomassa. Entretanto, é uma tecnologia de alto custo agregado.	19,32
	Tratamento ácido	Processo mais comum para pré-tratamento químico. Consiste na adição em altas temperaturas de ácidos orgânicos como: HCl, H <sub>2</sub> SO <sub>4</sub> , HNO <sub>3</sub> ou H <sub>3</sub> PO <sub>4</sub> . Nesse processo podem ser gerados inibidores como HMF, o que acarreta numa diminuição da eficiência na etapa de fermentação	19,23,32,33
	Tratamento alcalino	Pré-tratamento através da adição de bases como NaOH, Ca(OH) <sub>2</sub> , CaO, KOH ou NH <sub>4</sub> OH. Esse processo retira boa parte da porção hemicelulósica da biomassa	19,23,33
<b>Químico</b>	Deslignificação oxidativa	Processo de desconstrução da biomassa por agentes oxidantes como peróxido de hidrogênio (H <sub>2</sub> O <sub>2</sub> ). Nesse tratamento ocorre a eficiente remoção da lignina, mantendo-se a hemicelulose e celulose. Entretanto, inibidores podem ser gerados.	19,32,33
	<i>“Organosolv”</i>	Utiliza ação de uma mistura de diferentes solventes orgânicos que irão quebrar as ligações entre lignina e a hemicelulose. Esse processo retira quase totalmente a lignina da biomassa deixando a celulose praticamente intacta, além de alguns açúcares da hemicelulose. Entretanto, devido ao uso de solventes, é um processo que proporciona forte inibição no processo de fermentação	19,23,34
	Explosão a vapor	Pré-tratamento onde na presença de vapor de água em alta pressão e temperatura, ocorre a desconstrução da parte celulósica da biomassa. A porção da hemicelulose fica solubilizada.	19,23
<b>Físico-químico</b>	Expansão da Fibra em Amônia (AFEX)	Semelhante ao que ocorre na explosão a vapor com condições de alta temperatura e pressão, mas com adição de amônia. É um processo que possibilita uma melhor eficiência na etapa de hidrólise enzimática. Entretanto, em biomassas com alta teor de lignina, existem limitações para esse tipo de pré-tratamento	19,23

	Explosão por CO <sub>2</sub>	Utiliza o princípio que CO <sub>2</sub> , em alta pressão fica no estado líquido e forma ácido carbônico na presença de água. A ação desse ácido aumenta a taxa de desconstrução da biomassa.	19,23,35
<b>Biológico</b>	Microrganismos (fungos e bactérias)	Utiliza ação de microrganismos para decomposição da biomassa. É um processo que pode levar semanas para ser finalizado. Além disso, os açúcares monoméricos disponibilizados nessa etapa podem ser consumidos pelos microrganismos presentes no processo desconstrução.	19,23,29,32



**Figura 3 | Principais inibidores gerados durante o processo de desconstrução da biomassa e suas rotas de origem de formação.** Adaptado de Dos santos, *et al.*<sup>19</sup>

Após o processo de pré-tratamento, ainda existe uma grande quantidade de polissacarídeos que devem ser fragmentados em suas formas monoméricas, a fim de serem utilizados na etapa posterior de fermentação. A fragmentação desses polissacarídeos normalmente ocorre com auxílio de um coquetel de enzimas hidrolíticas e oxidativas, que atuam de forma mais específica em cada substrato (Tabela 2)<sup>19,36,37</sup>. As enzimas que compõem os coquetéis enzimáticos são obtidas de organismos com capacidade natural de degradar esses substratos, celulose e hemicelulose, como *Trichoderma reesei*, e espécies dos gêneros *Humicola*, *Chrysosporium*, *Penicillium*, *Acremonium* e *Aspergillus*<sup>38,39</sup>.



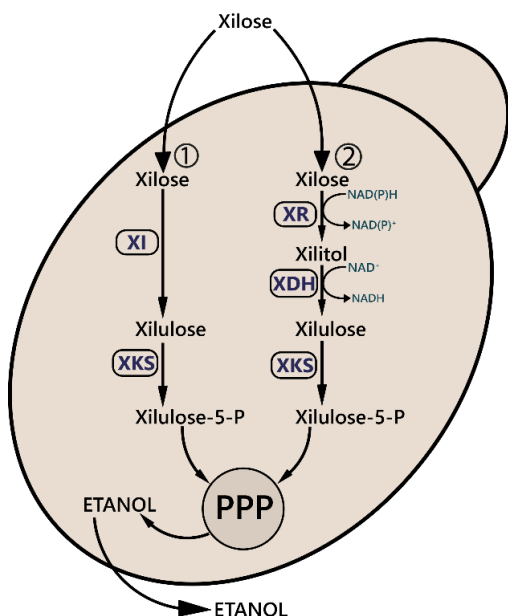
**Tabela 2 | Principais enzimas hidrolíticas utilizadas para disponibilização dos açúcares monoméricos**

Região da biomassa	Classe de enzimas	Descrição	Referência
<b>Celulose</b>	$\beta$ -1,4-endoglucanases (EC 3.2.1.4)	Hidrolisa ligações glicosídicas internas da região amorfa da celulose. Age sinergicamente com as $\beta$ -glicosidases e criam novas entradas para ação dessa classe de enzimas	40–42
	$\beta$ -1,4-exoglucanases/ Celobio-hidrolases (EC 3.2.1.91)	Agem nas extremidades da celulose liberando glicose ou celobiose (duas moléculas de glicose).	40,42
	$\beta$ -1,4-glicosidases (EC 3.2.1.21)	Responsável pela conversão de celobiose e outros oligômeros em glicose.	40,42
<b>Hemicelulose</b>	$\beta$ -1,4-endoxilanases	Hidrolisa entre ligações 1,4 entre moléculas de xilose gerando substratos de vários tamanhos	34,40
	$\beta$ -1,4-xilosidases (EC 3.2.1.37)	Hidrolisa entre as ligações glicosídicas $\beta$ -1,4 em moléculas de xilana e xilobiose	34,40
	$\beta$ -1,4-xilanases (EC 3.2.1.8)	Hidrolisa em ligações glicosídicas $\beta$ -1,4 entre moléculas manose	34,40
	$\beta$ -1,4-endomananases (E.C. 3.2.1.78)	Hidrolisa polissacarídeos de manana com quebra de ligações glicosídicas $\beta$ -1,4 entre moléculas manose	34,40
	$\beta$ -1,4-manosidases (E.C. 3.2.1.25)	Hidrolisa manobiose e manotriose em ligações glicosídicas $\beta$ -1,4 entre moléculas manose	34,40
	$\alpha$ -1,4-galactosidases (E.C. 3.2.1.22)	Liberam moléculas de galactose dentro de xilanas e galactomananas	34,40

### 1.3.2 Plataformas microbianas

Disponibilizados os açúcares monoméricos da biomassa lignocelulósica, ocorre a etapa de conversão em compostos renováveis a partir dos açúcares gerados. A levedura *S. cerevisiae* é um modelo industrial amplamente utilizado na produção de etanol e possui alta resistência a diferentes estresses do processo fermentativo 1G, sendo a linhagem Pedra-2 uma das cepas mais utilizadas em plantas industriais. Entretanto, esse microrganismo não possui a capacidade de consumir a xilose como substrato para fermentação, embora seja capaz de metabolizar lentamente seu isômero xilulose<sup>43</sup>. A xilulose é fosforilada pela ação da xiluloquinase (*XKS1*) em xilulose-5-P, sendo direcionada à via das

pentoses fosfato (PPP) que tem como produto gliceraldeído-3-fosfato e frutose-6-fosfato, posteriormente convertidos em etanol. Dessa forma, estudos vêm sendo realizados visando a inserção de vias metabólicas que possibilitem cepas de *S. cerevisiae* a capacidade de fermentar xilose (Figura 5)<sup>17</sup>.



**Figura 5 | Vias fermentativas utilizadas por *S. cerevisiae*.** Em 1 é representado a via de isomerização. Em 2, a via fermentativa através de reações de oxirredução. XI, xilose isomerase (EC 5.3.1.5.), XR, xilose redutase (EC 1.1.1.21), XDH, xilitol desidrogenase (EC 1.1.1.9) XKS1, xiluloquinase (EC 2.7.1.17). PPP, Vias das Pentose Fosfato.

### 1.3.2.1 Via de oxidação-redução - XR/XDH

As linhagens selvagens de *S. cerevisiae* apresentam o gene *GRE3* que codifica uma aldose redutase e realiza a redução da xilose em xilitol, e o gene *XYL2*, uma xilitol desidrogenase, que realiza a oxidação deste composto em xilulose. Porém, a expressão desses genes é muito baixa e não permite o crescimento das linhagens na presença de xilose como fonte única de carbono. Mesmo diante da superexpressão desses genes endógenos, não foi possível atingir níveis de crescimento satisfatórios<sup>44</sup>. Assim, uma alternativa encontrada para o aumento da eficiência de crescimento e consumo de xilose é a expressão de genes da via de oxidação-redução mais eficientes que os endógenos encontrados em *S. cerevisiae*.

Na natureza, existem organismos com a capacidade de metabolizar naturalmente xilose, como a levedura *Scheffersomyces stipitis*, que utiliza a via metabólica de oxido-redução através dos genes *XYL1* (xilose redutase) e *XYL2* (xilitol desidrogenase) para conversão da xilose<sup>45,46</sup>. Entretanto, quando expresso em *S. cerevisiae*, essa via apresenta como principal limitação a geração de um desbalanço redox de cofatores. A xilose redutase, XR, apresenta uma preferência pelo cofator NADPH, enquanto que a xilitol desidrogenase,

XDH, é NAD<sup>+</sup> dependente<sup>46</sup>. Dessa forma, o cofator produzido na segunda etapa da via de oxirredução não é o mesmo que será utilizado na primeira etapa, gerando um desbalanço entre essas duas etapas da via metabólica. Esse problema de balanço entre cofatores faz com que o composto intermediário xilitol seja acumulado, prejudicando o rendimento total de etanol<sup>17</sup>.

Ao longo dos anos, estudos demonstraram alternativas para solucionar o problema do desbalanço de cofatores e viabilizar o metabolismo de xilose pela via de oxirredução. Uma das alternativas apresentadas é a modificação da afinidade de cofator da xilose redutase para NADH em comparação a NADPH. Através de modificações específicas nesta proteína, foi possível gerar variantes de XR que apresentassem uma menor especificidade por NADPH. Entretanto, a afinidade por NADH não foi ampliada<sup>47-49</sup>. Em paralelo, considerando uma abordagem de modificação na enzima xilitol desidrogenase, foi possível gerar uma variante desta enzima com alta afinidade em NADP<sup>+</sup> em relação a NAD<sup>+</sup><sup>50</sup>. A enzima com as mutações D207A/I208R/F209S apresentou um grande aumento na afinidade por NADP<sup>+</sup> e sua expressão em linhagens de *S. cerevisiae* aumentou a produtividade de etanol e diminuiu a quantidade de subprodutos como xilitol<sup>50,51</sup>. Entretanto, dependendo do *background* da linhagem, a expressão desse mutante de XDH nem sempre reproduz os resultados obtidos com uma maior produtividade de etanol<sup>52</sup>.

### 1.3.2.2 Via de isomerização - XI

Além da via de oxi-redução encontrado em fungos, existe a via de isomerização da xilose em xilulose por meio da enzima xilose isomerase (XI), a qual é mais comumente encontrada bactérias (procariotos). A expressão funcional de enzimas XI bacterianas em linhagens de *S. cerevisiae* apresenta algumas limitações<sup>53</sup>. Muitos fatores podem estar relacionados a baixa funcionalidade dessas enzimas em *S. cerevisiae*, como o enovelamento errôneo da enzima<sup>54</sup>. Recentemente alguns trabalhos mostraram que a expressão de chaperonas, enzimas envolvidas no enovelamento correto das proteínas, permitiu a expressão funcional de uma XI bacteriana e a produção de etanol por linhagens recombinantes<sup>55,56</sup>. Contudo, a expressão do gene *xyIA* de *Clostridium phytofermentans* mostrou ser funcional em *S. cerevisiae* e atingiu altos valores de crescimento em xilose e rendimento na produção de etanol<sup>57</sup>.

Após muitos estudos sobre as XI bacterianas, em 2003 foi descrita uma xilose isomerase proveniente do fungo anaeróbico *Piromyces sp.*, que quando expresso em *S. cerevisiae*, permitia que essas linhagens crescessem em xilose como única fonte de carbono<sup>58,59</sup>. Após a descoberta dessa nova xilose isomerase, outros exemplos de genes que codificam essas enzimas provenientes de hospedeiros eucarióticos começaram a surgir na literatura. O gene *xylA* encontrado no fungo *Orpinomyces sp.* apresenta 94 % de identidade de aminoácidos com a sequência de *Piromyces sp.*, e quando expresso em linhagens de *S. cerevisiae* demonstrou resultados similares a XI de *Piromyces sp.*<sup>60</sup>.

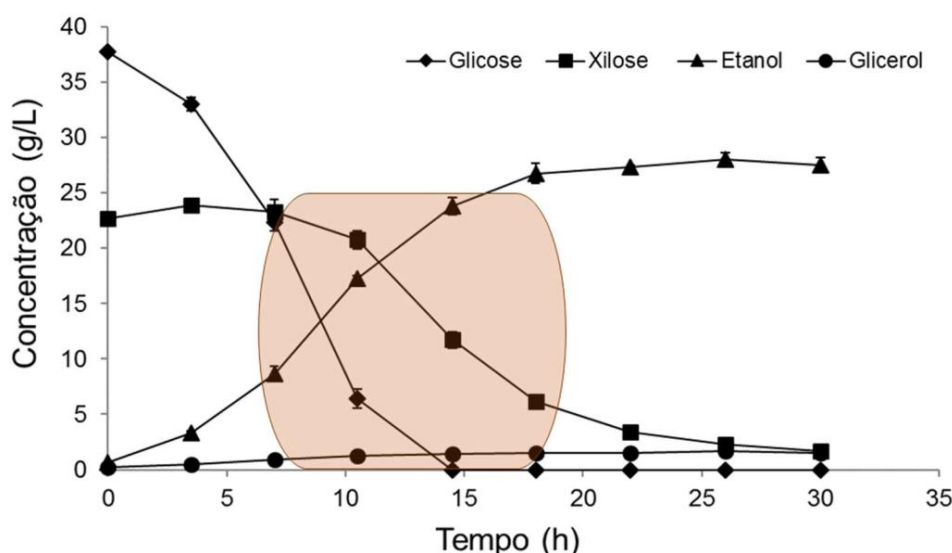
Contudo, mesmo com a presença das vias metabólicas de conversão de xilose, as cepas de *S. cerevisiae* não apresentam inicialmente uma grande eficiência fermentativa<sup>61</sup>. Uma forma de se aumentar essa eficiência é o estabelecimento de um processo de evolução adaptativa<sup>62</sup>, levando a uma melhora no rendimento de produção de etanol e diminuição do tempo total de fermentação. Linhagens de *S. cerevisiae* com via de da xilose isomerase que passaram pelo processo de evolução adaptativa tiveram diferentes mutações relacionadas com o aumento no consumo de xilose. Deleções de genes como *ISU1*, *SSK2*, *PMR1*, *HOG1* e *IRA2* foram descritas em diferentes linhagens evoluídas de experimentos distintos<sup>63-65</sup>. Curiosamente, mutações relacionadas a homeostase de metais em genes como *ISU1* e *PMR1*, proporcionaram maiores aumentos de performance nas linhagens evoluídas<sup>63,64</sup>. Dessa forma, uma das hipóteses descritas até o momento é que o acúmulo de metais divalentes como  $Fe^{2+}$  e  $Mn^{2+}$ , resultam em uma melhora atividade da xilose isomerase, uma vez que a XI é uma metaloenzima que utiliza cátions divalentes como cofatores<sup>66</sup>.

Associando as metodologias de engenharia metabólica e evolução adaptativa é possível identificar novos alvos moleculares e desenvolver linhagens capazes de fermentar xilose eficientemente<sup>63-65</sup>. Entretanto, ainda existem limitações no processo de assimilação da xilose que precisam ser otimizadas visando um eficiente processo na conversão dos açúcares da biomassa.

### 1.3.2.3 Transporte dos açúcares da biomassa

Uma das maiores limitações no metabolismo de xilose é a ineficiente assimilação desse açúcar para o ambiente intracelular, principalmente em meios que apresentem a mistura de substratos C5 e C6<sup>67,68</sup>. Os transportadores de xilose descritos previamente na literatura, tanto endógenos quanto heterólogos, demonstram baixa afinidade a xilose<sup>69-71</sup>. Além disso, alguns transportadores descritos com maior afinidade em xilose apresentam dificuldades para serem expressos de forma funcional em linhagens *S. cerevisiae*<sup>69,72,73</sup>. O desenvolvimento de transportadores que apresentem alta eficiência e afinidade por esse açúcar C5 pode promover um aumento na velocidade da fermentação, melhorando a produtividade do processo de produção do etanol 2G.

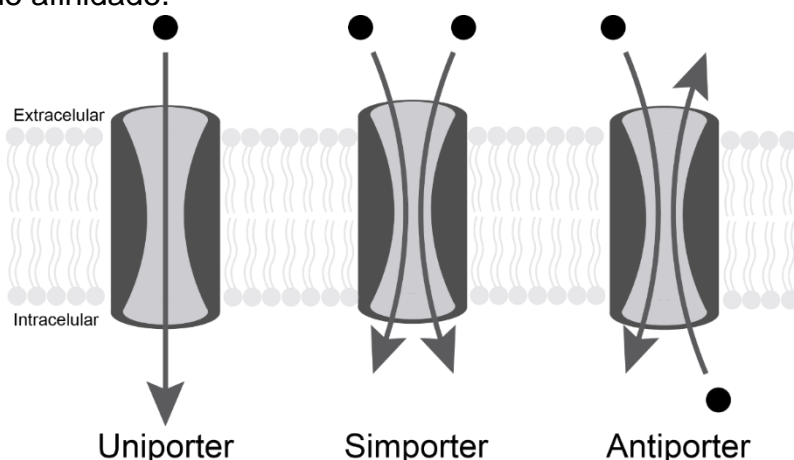
Na mistura dos açúcares C5 e C6, as leveduras irão consumir inicialmente a glicose. O transporte de xilose só se inicia somente quando a glicose está presente em baixas concentrações no meio<sup>67</sup>. Dessa forma não ocorre um consumo simultâneo dos substratos C5 e C6, levando a um processo com alto tempo de fermentação, figura 6. Esse mecanismo de preferência de glicose em detrimento a outro açúcar é comumente conhecido como repressão catabólica<sup>67</sup>. Dessa maneira, essa preferência de consumo ocorre em primeiro momento devido a maior afinidade dos transportadores de levedura por glicose em comparação aos outros açúcares<sup>71,72</sup>.



**Figura 6 | Fermentação de uma linhagem modificada de *S. cerevisiae*.** A linhagem geneticamente modificada só começa a consumir xilose quando a glicose está sendo depletada do meio (região realçada), levando a um extensivo tempo de consumo desses açúcares sendo o composto glicerol um sub produto indesejável da fermentação.

A etapa inicial para metabolismo de qualquer carboidrato é seu transporte do ambiente extracelular para o ambiente intracelular. Moléculas complexas como os carboidratos necessitam de transportadores de membrana que facilitem a entrada desses açúcares na célula<sup>74</sup>. Esses transportadores de membrana são membros principalmente da “*Major Facilitator Superfamily*” (MFS), encontrados em todos os reinos biológicos<sup>75–77</sup>. Os transportadores MFS são divididos em três classes: transporte uniporte onde apenas um substrato é carregado; transporte simporte, onde o substrato é carregado e co-transportado juntamente a um íon<sup>78</sup>, e por último o transporte antiporte, onde dois substratos são carregados simultaneamente em sentidos opostos<sup>79</sup>, figura 7. A superfamília MFS pode ser dividida em 74 subfamílias<sup>80</sup>, onde podemos encontrar os transportadores heterólogos e endógenos de *S. cerevisiae* capazes de carrear xilose.

Os transportadores endógenos de *S. cerevisiae* *GAL2* e *HXT7* já foram descritos como os melhores transportadores capazes de carrear xilose em levedura<sup>81</sup>. Entretanto, eles apresentam uma baixa afinidade por esse açúcar,  $K_m = 225.6 \pm 15.8$  mM e  $K_m = 200.3 \pm 13.2$  mM, respectivamente<sup>82</sup>. Por outro lado, esses mesmos transportadores demonstram uma grande afinidade para glicose,  $K_m = 1.9 \pm 0.2$  mM para *GAL2* e  $K_m = 0.5 \pm 0.1$  mM para *HXT7*<sup>82</sup>. Dessa maneira, em um cenário industrial 2G, onde temos a presença tanto de glicose quanto de xilose em altas concentrações, os transportadores de *S. cerevisiae* irão apresentar uma maior preferência para glicose em relação a xilose devido a diferenças de afinidade.



**Figura 7 | Transporte de moléculas contra concentração de gradiente.** Transporte direto por uma molécula (uniporte), co-transporte através de duas moléculas em uma mesma direção (simporter), e transporte de duas moléculas em direções opostas (antiporte).

Da mesma maneira, alguns transportadores heterólogos já foram analisados e expressos em *S. cerevisiae*, visando otimizar a eficiência de assimilação de xilose pela levedura. Esses transportadores foram isolados de organismos que naturalmente consomem xilose como *Candida intermedia*, *Scheffersomyces (Pichia) stipitis*, *Neurospora crassa*, *Debaryomyces hansenii*, *Aspergillus niger* e *Trichoderma reesei*<sup>72,73,83-87</sup>. Desses organismos, foi possível encontrar transportadores como *XUT1*<sup>81</sup>, *XUT3*<sup>81</sup>, *SUT1*<sup>68,83</sup>, *Xyp29*<sup>85</sup>, *GXF1*<sup>73,84,88,89</sup>, *GXS1*<sup>73,84,90</sup>, *SUT2*<sup>83</sup>, *RGT2*<sup>91</sup> e *SUT3*<sup>83</sup> que demonstraram capacidade de carrear xilose. Além desses genes, muitos outros putativos transportadores de xilose já foram testados. Entretanto, pouco foram funcionalmente expressos em linhagens de *S. cerevisiae*<sup>69</sup>. Além disso, aqueles que foram funcionais, não apresentaram altos desempenhos na assimilação de xilose<sup>81</sup>.

O melhor transportador heterólogo descrito na literatura atualmente é o *GXF1*<sup>81</sup>, isolado da levedura *C. intermedia*, e apresenta uma afinidade por xilose de  $K_m = 48,6 \pm 6,5 \text{ mM}$ <sup>84</sup>, superior aos transportadores endógenos *GAL2* e *HXT7*. Curiosamente, apesar de apresentar maior afinidade em relação aos endógenos, linhagens de *S. cerevisiae* com o *GXF1* apresentaram um crescimento ligeiramente inferior comparado ao transportador nativo *GAL2*<sup>81</sup>. O facilitador *GXF1* apresenta alta afinidade por glicose,  $K_m = 2 \pm 0,6 \text{ mM}$ <sup>84</sup> em relação a xilose,  $K_m = 48,6 \pm 6,5 \text{ mM}$ <sup>84</sup>. Dessa forma, na presença dos principais açúcares do hidrolisado lignocelulósico, xilose e glicose, não teremos um consumo simultâneo desses dois açúcares em linhagens recombinantes de *S. cerevisiae* expressando *GXF1*. Mesmo sendo o melhor transportador já descrito, ele permite a assimilação da xilose apenas em baixas concentrações e sem a presença de glicose, limitando seu uso em condições industriais de produção de etanol 2G<sup>81,88,92</sup>.

Dessa maneira, grupos de pesquisa focaram seus esforços para obter um transportador de xilose mais eficiente e com maior afinidade para esse açúcar em relação a glicose<sup>86,93-99</sup>. Além da competição pelo transporte no ambiente extracelular, a presença de glicose intracelular tem um efeito inibidor no metabolismo C5<sup>100-102</sup>. Os mecanismos para superar essa competição tanto extracelular quanto intracelular devido a repressão catabólica por glicose ainda estão sendo estudados e elucidados.

Uma alternativa encontrada para superar essa competição extracelular no transporte dos açúcares foi através de modificações em sítios específicos nesses transportadores. Dessa forma, a partir de uma população de variantes desses transportadores, foram selecionadas proteínas de membrana com maior afinidade por xilose e baixa inibição por glicose. Utilizando essa estratégia, foram gerados mutantes do transportador endógeno *GAL2* que apresentou aumento na afinidade por xilose e uma menor assimilação de glicose. A expressão desse transportador modificado de *GAL2* em linhagem de *S. cerevisiae* levou ao um aumento na taxa de co-consumo na mistura dos açúcares<sup>103</sup>

Uma nova estratégia encontrada para se estudar os efeitos da inibição por glicose é através do uso de linhagens mutantes com deleções nos genes que codificam hexoquinases da via glicolítica (*GLK1*, *HXK1* e *HXK2*)<sup>95</sup>, interrompendo o primeiro passo da via glicolítica. A linhagem *hvk<sup>0</sup>* obtida perde a capacidade de consumir a glicose, mesmo sendo capaz de transportar esse açúcar, acumulando intracelularmente. Assim, essa linhagem não reconhece a glicose como uma fonte de energia intracelular, mas como um inibidor no crescimento, uma vez que não será mais utilizada para o metabolismo. Dessa forma, através do uso de uma linhagem *hvk<sup>0</sup>*, um processo de evolução adaptativa resultou em mutantes que apresentavam alto crescimento em xilose e sem a inibição por glicose<sup>93,98</sup>. Dois experimentos distintos de evolução adaptativa utilizando leveduras *hvk<sup>0</sup>* tiveram diferentes mutações dentro da região de ligação com glicose nos transportadores *GAL2*<sup>98</sup> e *HXT36*<sup>93</sup>, o que acarretou na perda da capacidade de transportar esse açúcar C6. Contudo, o transporte de xilose se manteve. Isso ocorreu através da modificação de apenas um aminoácido N376F<sup>98</sup> e N367A<sup>93</sup>, em *GAL2* e *HXT36*, respectivamente, que foram responsáveis pelo aumento da especificidade no transporte de xilose. Assim, com a perda da interação da molécula de glicose com o transportador, a molécula de xilose fica livre para ser carregada sem que ocorra uma competição pelo sítio de ligação. Esses novos transportadores obtidos através dessas modificações pontuais ampliaram as perspectivas para a possibilidade de construção de novas linhagens de *S. cerevisiae* capazes de consumir simultaneamente xilose e glicose. Essas variantes de proteínas de membrana devem ser capazes de carrear xilose com alta especificidade e alta velocidade,

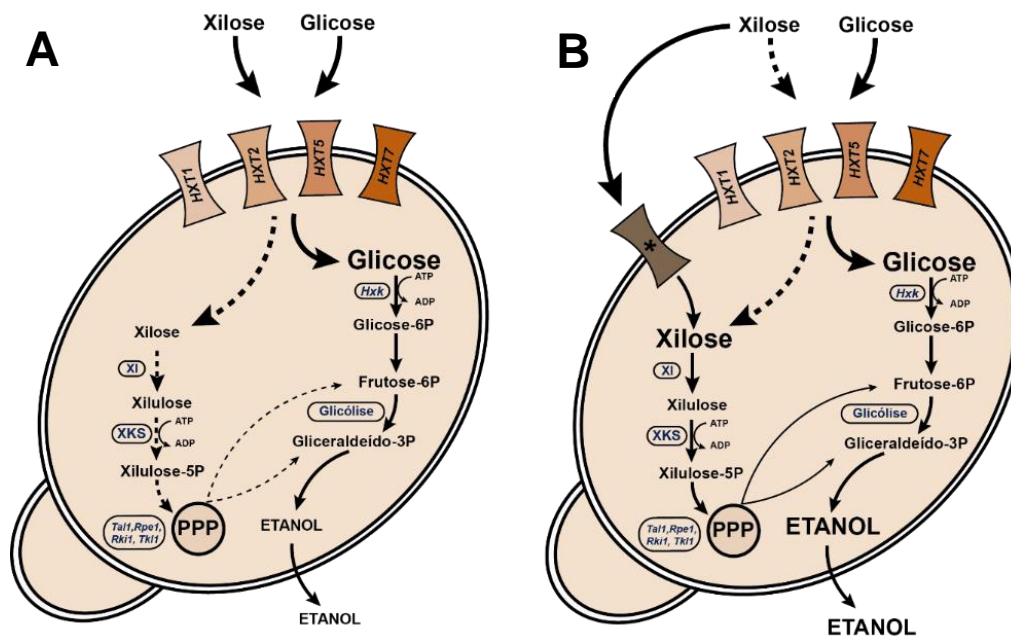


melhorando a eficiência de assimilação. Os mecanismos moleculares envolvidos na repressão catabólica e inibição por glicose e, principalmente, como acontece a repressão do uso de outras fontes de carbono ainda precisam ser completamente elucidados.

#### **1.4 Desafios do processo fermentativo 2G**

Com o aumento da demanda por uma renovação na matriz energética mundial, intensificam-se as pesquisas na área da indústria de Segunda Geração. A construção de linhagens cada vez mais robustas e eficientes capazes de produzir bioquímicos renováveis a partir de material lignocelulósico é necessária para a viabilização do processo 2G. Além disso, um dos principais gargalos para o estabelecimento dessa tecnologia é o design e a construção de metabolismos mais eficientes para a conversão dos açúcares provenientes da biomassa.

Uma linhagem que possua a capacidade de assimilar e fermentar simultaneamente os principais açúcares do material lignocelulósico, glicose e xilose, resultará em um aumento da produtividade do processo de produção. Para atingir esse objetivo, ocorreu ao longo dos anos uma intensa busca por transportadores específicos e de alta eficiência no transporte de xilose. Entre aqueles que foram funcionais em *S. cerevisiae*, nenhum apresentou um aumento significativo no transporte em altas concentrações de xilose e na presença de glicose. Na primeira etapa desse trabalho, identificamos e caracterizamos o transportador Cs4130 do organismo *Candida sojae*, representante da classe dos transportadores uniporte, que demonstrou uma perda natural de inibição em altas concentrações de xilose. Além disso, Cs4130 pode ser utilizado em procedimentos de engenharia metabólica visando o desenvolvimento de cepas para a indústria 2G (Capítulo 1). Contudo, esse novo carreador ainda apresenta inibição promovida pela glicose. Na segunda parte do trabalho, desenvolvemos plataformas microbianas específicas para desenvolver e selecionar variantes de mutantes capazes de consumir simultaneamente substratos C5 e C6 (Capítulo 2), figura 8.



**Figura 8 | Consumo de xilose e glicose em células de *S. cerevisiae*.** **A** – Linhagens selvagens de *S. cerevisiae* com baixa assimilação e xilose pelo transportadores endógenos. **B** – Linhagem capaz de simultaneamente consumir os substratos C5 e C6, devido transportador específico de xilose (\*).

## Capítulo 1

**Novel xylose transporter Cs4130 expands the range of sugar uptake at high xylose concentrations in recombinant *Saccharomyces cerevisiae* strains**

## **Novel xylose transporter Cs4130 expands the range of sugar uptake at high xylose concentrations in recombinant *Saccharomyces cerevisiae* strains**

João Gabriel Ribeiro Bueno<sup>1,2</sup>, Guilherme Borelli<sup>2</sup>, Thamy Livia Ribeiro Corrêa<sup>1</sup>, Mateus Fiamenghi<sup>2</sup>, Juliana José<sup>2</sup>, Leandro Cristante de Oliveira<sup>3</sup>, Gonçalo A. G. Pereira<sup>2</sup>, Leandro Vieira dos Santos<sup>1,2</sup> §

<sup>1</sup> Brazilian Biorenewable National Laboratory (LNBR), Brazilian Center for Research in Energy and Materials (CNPEM), 13083-100, Campinas, São Paulo, Brazil

<sup>2</sup> Department of Genetic and Evolution – Institute of Biology, UNICAMP, Campinas, São Paulo 13083-970, Brazil.

<sup>3</sup> Department of Physics - Institute of Biosciences, Humanities and Exact Sciences, UNESP - São Paulo State University, São José do Rio Preto, SP 15054-000, Brazil.

§To whom correspondence should be addressed. E-mail: leandro.santos@lnbr.cnpem.br / leandrovrst@gmail.com

Keywords: *Saccharomyces cerevisiae*, second-generation ethanol, xylose uptake, sugar transporter

## Abstract

**Background:** The need to restructure the world energy matrix based on fossil fuels and mitigate greenhouse gases emission stimulated the development of new technologies based on renewable energy. One promising and cleaner alternative towards energy restructuration is the second-generation (2G) fuels, produced from lignocellulosic biomass sugars. A major challenge on 2G technologies establishment is the inefficient assimilation of the five-carbon sugar xylose by engineered *Saccharomyces cerevisiae* strains, increasing the fermentation time. The uptake of xylose across plasma membrane is a critical limiting step and the recombinant *S. cerevisiae* strains are not designed with a versatile xylose transport system capable to transport the pentose in a wide range of sugar concentrations present in 2G process.

**Results:** Assessing diverse microbiomes such as the digestive tract of plague insects or several decayed lignocellulosic biomasses, we isolated several yeast species capable of using xylose. Comparative fermentations select the yeast *Candida sojae* as a potential source of high-affinity transporters. Therefore, comparative genomic analysis elects four potential xylose transporters whose properties were evaluated in the transporter-free EBY.VW4000 strain carrying the xylose pathway integrated into the genome. While the traditional xylose transporter *GXF1* is inhibited at concentrations above 10 g/L, the strains containing Cs3894 and Cs4130 show superior xylose uptake, not affected at concentrations up to 50 g/L. An extensive comparison of Cs4130 and *GXF1* showed that Cs4130 is a more promiscuous transporter than *GXF1* given it supports higher concentrations of xylose. In addition, Cs4130 variants would probably result in a diversity of phenotypes not observed in *GXF1*, placing Cs4130 as a model to obtain transporters with improved properties regarding the assimilation of substrates other than glucose.

**Conclusions:** Considering that xylose concentrations in second generation hydrolysates reaches high values in several designed processes, Cs4130 are a profitable candidate for xylose uptake due to natural loss of inhibition. Here we demonstrate a novel eukaryotic molecular transporter protein that is not inhibited in high xylose concentrations and can be used as promising target towards engineering efficient pentose utilization in yeast.

## Introduction

The concern to mitigate the climate change and global warming impacts led to a commitment to restructure the energy matrix based on fossil fuels. One promising and cleaner alternative towards energy restructuration demands the development of sustainable advanced biofuels produced from renewable lignocellulosic biomass, also known as second-generation biofuels (2G)<sup>11,104–106</sup>. The conversion of plant biomass sugars into bioproducts requires the efficient extraction and breakdown of the main macromolecular components from the recalcitrant structure of the plant cell wall. The complex architecture of plant cell wall presents cellulose (40-50%), hemicellulose (25-35%) and lignin (15-20%) as major constituents polymers<sup>19,107</sup>. Cellulose consists on chains of glucose, which are the most abundant sugar monomer present at the biomass. From the hemicellulose fraction, xylose, a five-carbon sugar, is the second most abundant monomer of the lignocellulosic portion.

To be able to convert the totality of biomass-derived sugars in different bioproducts is essential the development of robust and high-yielding microbial platforms through synthetic biology approaches. The budding yeast *Saccharomyces cerevisiae* has been traditionally used as model eukaryotic platform to design new routes and engineer metabolic fluxes toward desired high-value bio-products, including production of fuels, chemicals, foods, feeds, and pharmaceuticals<sup>108,109</sup>. However, the C5 sugar xylose is not naturally metabolized by wild type strains of the yeast *S. cerevisiae*, requiring systems metabolic engineering strategies in order to develop high-performance strains<sup>63,110</sup>. Xylose catabolism in *S. cerevisiae* occurs through integration of heterologous metabolic pathways associated with adaptive evolution strategies to reshape cellular metabolism and improve fermentation fitness<sup>111</sup>. Mutations related to fitness-enhanced phenotypes arises during evolution to alleviate metabolic bottlenecks and accelerate xylose conversion<sup>63,64</sup>. Using these combined approaches, is possible to improve catabolic fluxes and engineer strains to assimilate the C5 sugar xylose rapidly. However, several limitations still need to be addressed to optimize biomass conversion, including the slow xylose transport by *S. cerevisiae* strains.

The transport of sugars occurs through transmembrane proteins, which perform the uptake of a wide array of substrates between extracellular and intracellular environments of the cell. Sugar transporters are mainly members of the major facilitator superfamily (MFS), found on all branches of life<sup>75-77,79</sup>. The MFS members are divided in three classes: uniport transport with a unique substrate, symporter transport where upon a substrate is usually coupled with an ion, and antiport transport with two different substrates moving on opposite directions<sup>75,78,79</sup>. Moreover, these transporters are branched on 74 subfamilies within MFS<sup>80</sup>, in which some of them include well known membrane proteins that can transport xylose such as *GAL2*<sup>71,81,96,112,113</sup>, *HXT*<sup>71,81,96,112</sup>, *GXF1*<sup>81,88,89</sup> and *GXS1*<sup>73,90,114</sup>. The budding yeast *S. cerevisiae* has a set of native hexose transporters (*HXT1*, 2, 4, 5, 7 and *GAL2*) that can import xylose with low affinity and presents strong glucose repression, limiting fermentation of mixed biomass sugars<sup>98,115</sup>. *S. cerevisiae* endogenous transporters have preference to assimilate glucose over xylose in industrial lignocellulosic hydrolysates, increasing the time of fermentation<sup>67</sup>. The expression of heterologous xylose transporters in engineered *S. cerevisiae* strains has already resulted in improved xylose assimilation rates. For example, the most well described xylose facilitator, *GXF1*, isolated from *Candida intermedia*, improved transport kinetics and xylose utilization in recombinant xylose-fermenting *S. cerevisiae* by 3-fold<sup>84,88,89,92,116</sup>. The glucose/xylose facilitator *GXF1* from *C. intermedia* is considered to be the most efficient transporter when compared with other known heterologous xylose transporters<sup>81,89,92</sup>. However, the transport of sugars seems concentration-dependent. Improvements in transport rates and growth were observed only on very low xylose concentration, whereas remained unchanged at higher concentrations<sup>81,88,92</sup>.

The inefficient and unspecific xylose transport across the plasma membrane in *S. cerevisiae* is a major drawback to design microbial platforms for industrial biomass utilization<sup>103</sup>. Identify superior transporter proteins and molecular mechanisms for improved xylose utilization are essential at the metabolic engineering field for 2G biorefineries. In this work, by assessing diverse microbiomes such as the digestive tract of plague insects or several biomasses, we isolated several yeast species capable of using xylose rapidly. Assessing the genome of the best xylose-consuming yeast, *Candida sojae*, we were able to

identify new xylose transporters members of the MFS. Transporter-free yeast cells expressing the symporter Cs4130 show increased growth rates on high concentrations of xylose whereas the traditional well described facilitator *GXF1* is severely inhibited. Considering that xylose concentrations in 2G hydrolysates reaches high values, the new transporter is profitable candidate for metabolic engineering procedures due to natural loss of inhibition. An extensive structural analysis by molecular docking and dynamics were conducted to point differences in residues related do sugar (xylose/glucose) coordination and uptake/trapping. Here we demonstrate a novel eukaryotic transporter protein that are not inhibited in high xylose concentrations and can be used as a promising target towards efficient pentose utilization in engineered yeasts.

## **Methodology**

### **Media and culture methods**

Yeast cells were grown on liquid YP medium (10 g/L yeast extract and 20 g/L peptone) supplemented with 20 g/L D-glucose (YPD) for cell propagation. For xylose cultivation, D-glucose was replaced for 20 g/L D-xylose. For standard transformants selection, cells were grown at 30 °C in complete synthetic media YNB (6.7 g/L yeast nitrogen base without amino acids, Difco) supplemented with 1 g/L drop-out without uracil, 20 g/L glucose and 20 g/L agar<sup>147</sup>. YP were autoclaved at 121 °C for 20 min and YNB media were filter-sterilized using 0.2- $\mu$ m bottle-top filters. Strain EBY.VW4000, kindly provided by Prof. Eckhard Boles from Goethe university <sup>125</sup>, and strain EBY\_Xyl1 were grown in synthetic media YNB plus D-maltose instead of D-glucose, and EBY\_Xyl1\_ *hxxk*<sup>0</sup> were cultivated in 3% ethanol as carbon source. EBY\_Xyl1 expressing transporter genes were grown in solid synthetic media YNB supplemented with different carbon sources (D-galactose, D-mannose and D-fructose) for spot assay test. High cell density fermentation of EBY\_Xyl1 expressing heterologous transporters were performed on YNB supplemented with 5 g/L of casamino acids (Difco), 1 g/L of Tryptophan (Sigma) and the correspondent carbon source, D-maltose, D-xylose or D-glucose. The yeast cells isolation media was adapted from Cadete et al. (2013)<sup>148</sup>: 6.7 g/L yeast nitrogen base without amino acids (Difco) supplemented with 1 g/L of complete drop-out, 7 g/L of D-xylose, 0.2 g/L of chloramphenicol and



for solid media was used 10 g/L D-xylose and 20 g/L agar. The selection media for strains with pSH65 plasmid was supplemented with 300 µg/ml of zeocin (Thermo Scientific) on solid plates. For *URA3/5-FOA* counter-selection procedure, strains expressing pSH65 plasmid were grown in synthetic medium supplemented with 440 mg/ml of uracil and 1g/L of 5-FOA (Zymo Research). For antibiotic resistance, 200 mg/L of geneticin and 200 mg/L of hygromycin was added to YPD for strains expressing the *KanMX4* (G418)<sup>149</sup> or *hph*<sup>150</sup> markers, respectively. Likewise, 300 µg/ml of zeocin was added for strains transformed with pSH65<sup>151</sup>. Permanent stocks of yeasts and *E. coli* cells were mixed with glycerol before storage at – 80 °C.

### **Isolation and screening of D-xylose-consuming yeasts**

The C5-yeasts were isolated from two different sites: São Miguel dos Campos, Alagoas, Brazil, and Cosmópolis, São Paulo, Brazil. Samples from decayed energy cane and sugarcane and insects' pests associated with this material were collected from the first location. Insects pests isolated from corn were isolated from the second location. The decayed energy cane and sugarcane samples were collected closest to an ethanol mill, where the material was exposed to the environment for long periods. The local climate is hot and humid, with an average temperature from 20 to 32 °C and average annual precipitation of 1408 mm. Twenty samples from each material were collected in July 2015, stored in sterile recipients and transported under refrigeration to laboratory in less than 24 hours, according to Cadete et al., 2013<sup>148,152</sup>. The same procedure was conducted with the pest insects associated to plant material. From sugarcane and energy cane materials, 1 gram of each sample was placed separately in 125 mL erlenmeyer flasks containing 50 mL of isolation xylose medium. For insects' samples, in sterile environment, guts were removed and macerated with a micro pistil in a 1,5 mL conical tube containing 400 µL of isolation medium, and then re-suspended separately in 125 mL erlenmeyer flasks containing 50 mL of isolation xylose medium. The flasks were incubated for 2-6 days in orbital shakers at 27 °C, 150 rpm until growth was visually detected. Five consecutive batches were conducted in the same medium for population enrichment and selection of strains capable of assimilating xylose rapidly. After growth, 50 µL of medium were poured

and spread with drigalski spatula at plate dishes containing xylose isolation solid medium. Plates were incubated at 30 °C until colonies were observed. The isolated yeasts were maintained in permanent frozen stocks using 30 % glycerol solution at -80 °C<sup>147</sup>.

### **DNA extraction and molecular identification**

The C5-yeast species were identified using the internal transcribed spacer sequence (ITS1-5.8S-ITS2) and D1/D2 domains. The DNA extraction was performed using the standard *S. cerevisiae* protocol for DNA extraction described in Ausubel, 2003<sup>24</sup>. The ITS1-5.8S-ITS2 and D1/D2 region sequences were amplified by PCR using GoTaq® DNA polymerase (Promega) as described in FELL et al., 2000 and Cadete et al., 2013<sup>152</sup>. PCR amplicons were sequenced in a Sanger platform using an Applied Biosystems® Genetic Analyzer 3500. The amplicons were manually trimmed using a base QV of 20 and grouped in contigs. These contigs were used in BLASTn megablast program at NCBI website (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The ITS and D1/D2 domains were retrieved from type material (CBS - <http://www.cbs.knaw.nl/>) of the top five hits in BLASTn and the sequences were globally aligned against the used query utilizing Clustal W in MEGA version 6<sup>153</sup>. The query that has <1% divergence to material type sequence, was considered from the same species of the type strain<sup>154</sup>.

### **Comparative fermentations of xylose-fermenting yeasts**

In order to select the most suitable source of genes related to xylose metabolism, the isolated yeasts, molecularly identified, were compared in YPX medium. The ability to ferment D-xylose was screened using Durham tubes and the yeast *S. stipitis* NRRL Y- 7124 as positive control. The tubes were incubated at 30 °C on shaker at 150 rpm for 10 days to observe the gas production. The ability to assimilate xylose of the isolated yeast species were also compared in erlenmeyers with 150 mL of YPX (30 g/L) in orbital shakers, incubated at 30 °C and 150 rpm, starting with optical density at 600 nm of approximately 1.0. All the experiments were conducted in triplicates and samples were collected to measure optical density and to quantify metabolites with HPLC. The yeast

*Candida sojae* was also tested in hemicellulosic hydrolysate obtained from sugarcane bagasse in the same conditions described above.

### Analytical procedures

Samples were analyzed using Waters e2795 HPLC with a Hpx 84h column from Biorad and measured for glucose, xylose, ethanol, glycerol, xylitol, succinic acid and acetic acid. A standard curve with known concentrations of compounds of interest was also analyzed using the same procedure. Ethanol, xylitol and glycerol yields ( $Y_p/s$ , g/g) were calculated considering the ratio of the specific product formation to the sugar consumption. Productivity ( $Q_p$ , g/L.h) was calculated as the ratio between product concentration (g/L) and fermentation time (h).

### Phylogenetic methods

The ability to ferment xylose for yeast species used in phylogenetic analysis was assumed using the description of type strains deposited in Centraalbureau voor Schimmelcultures (CBS - <http://www.cbs.knaw.nl/>) yeast collection. Using the Joint Genome Institute (JGI - <http://genome.jgi.doe.gov/mycocosm/species/tree:VyB1A?organism=ascomycota>) phylogenetic tree of Ascomycetous species, we chose and downloaded the deposited genomes from yeasts in near and far branches which our isolates were present. Orthology prediction of a Markov clustering was implemented in OrthoMCL<sup>155</sup> which cluster orthologs and paralogs per similarity in functional groups called hereafter gene families. These resulting families are given in numerical descending order, such that larger families come before smaller families, and Fam0 being the largest. Transporter families were obtained using known xylose transporter proteins as baits for a BLAST search. Using homologous gene families as a starting point, we chose 1255 ortholog single-copy genes shared by all 18 chosen-species for phylogeny reconstruction. The adjustment of nucleotide and amino acids substitution models used Bayesian Information Criteria (BIC) in jModelTest2<sup>156</sup> and ProtTest3<sup>157</sup>. Maximum Likelihood inference was used to reconstruct phylogenies in RAxML v8<sup>158</sup> with 5000 bootstrap tests for branch support. For transporters' trees reconstructions, amino acid and nucleotides were globally aligned using MAFFT and Maximum

Likelihood inference was carried in IQTree using its own adjustment substitution model inference (BIC and AIC criteria implemented) and 1000 bootstraps for branch support.

### **Plasmid assembly**

Expression cassettes and plasmids were assembled in a single reaction using the Gibson Assembly<sup>®</sup> Master Mix (New England Biolabs - NEB)<sup>159</sup>, and the backbone pRS426<sup>160</sup> cleaved with *Bam*HI (Promega). The coding sequences of putative transporters were amplified by PCR using Phusion DNA polymerase (New England BioLabs - NEB) from gDNA of *C. sojae*, and the control transporter *GXF1* was amplified from *Candida intermedia* genome, using specific primers for each gene (Table S1). Sequences of *TDH1* promoter and terminator were amplified from *S. cerevisiae* strain LVA1 genome<sup>63</sup>. Fragments from PCR amplifications were purified from agarose gel using the Kit Wizard<sup>®</sup> SV Gel and PCR Clean-Up System from Promega before the assembly reaction. The cassettes containing the genes *XYL1* and *XYL2* from *Scheffersomyces stipitis* and *XKS1* from *S. cerevisiae* were amplified from the plasmids pSsXRDXH and pScXKS (Santos, unpublished results). Both cassettes use *URA3* as marker flanked by two *loxP* sites. The correct construction of plasmids and expression cassettes were confirmed by Sanger sequencing. The main components of the plasmids are summarized in Table S2.

### ***Saccharomyces cerevisiae* strains construction**

The molecular techniques were performed using standard procedures<sup>R1</sup>. The *S. cerevisiae* strains, primers and plasmids used in this study are listed in table 4 and supplementary tables S1 and S2, respectively. The expression cassettes containing the genes *XYL1*, *XYL2* and *XKS1* were amplified by PCR, purified from agarose gel using the kit Wizard<sup>®</sup> SV Gel and PCR Clean-Up System (Promega) and used to transform the EBY.VW4000 strain with the LiAc/SS-DNA/PEG protocol<sup>27</sup>. The stable integration of the two expression cassettes was molecularly verified. The auxotrophic marker *URA3* was excised using the plasmid pSH65 (Cre-*loxP* system)<sup>151</sup>. The strain was induced on maltose complete medium for a basal expression of the Cre-recombinase

enzyme, restoring the uracil auxotrophy. Counter selection was performed using 5-FOA to select strains which excised *URA3* gene. The transformation of EBY\_Xyl1 with plasmids pCIGXF1, pCS186, pCS2608, pCS3894, pCS4130 and pRS426 was selected in YNB medium lacking uracil and confirmed by PCR. The deletion of hexokinase genes *HXK1*, *HXK2* and *GLK1* was conducted through insertion of antibiotic markers *hph*, *KanMX4* and *ble*, amplified from pAG32<sup>150</sup>, pFA6KANMX4<sup>149</sup> and pSH65<sup>151</sup>, respectively, resulting in EBY\_Xyl1\_ *hxxk*<sup>0</sup> mutant cell.

**Table 1** | *S. cerevisiae* strains used in the study.

Strain	Relevant genotype/feature	Source
EBY.VW4000	<i>CEN.PK2-1C hxt13Δ::loxP hxt15Δ::loxP hxt16Δ::loxP hxt14Δ::loxP hxt12Δ::loxP hxt9Δ::loxP hxt11Δ::loxP hxt10Δ::loxP hxt8Δ::loxP hxt514Δ::loxP hxt2Δ::loxP hxt367Δ::loxP gal2 Δ stl1Δ::loxP agt1Δ::loxP ydl247wΔ::loxP yjr160cΔ::loxP</i>	125
EBY_Xyl1	<i>CEN5::pPGK1-XYL1-tPGK1-pTDH1-XYL2-tTDH1; CEN2::pADH1-XKS1-tADH1</i>	This study
EBY_Xyl1_GXF1	pCIGXF1	This study
EBY_Xyl1_186	pCS186	This study
EBY_Xyl1_2608	pCS2608	This study
EBY_Xyl1_3894	pCS3894	This study
EBY_Xyl1_4130	pCS4130	This study
EBY_Xyl1_pRS	pRS426	This study
EBY_Xyl1_HXK <sup>0</sup>	<i>hphΔhxxk1; kanΔhxxk2; zeoΔglk1</i>	This study
EBY_Xyl1_HXK <sup>0</sup> _pRS	pRS426	This study
EBY_Xyl1_HXK <sup>0</sup> _4130	pCS4130	This study
EBY_Xyl1_HXK <sup>0</sup> _GXF1	pCIGXF1	This study

### GFP- tagged strains and microscopy

The codon-optimized monomeric GFP mUkG1 from soft coral<sup>161</sup> was synthesized by Genscript and fused to the plasmids pCs4130. Strain EBY\_Xyl1 harboring the Cs4130-mUkG1 fusion protein was grown on YNB medium without uracil, supplement with maltose 20 g/L overnight, at 30 °C / 200 rpm. Yeast cells were examined by confocal fluorescence microscopy (Leica TCS SP8 inverted microscope). Fluorescence images were taken using the emission signal range of 503 – 530 nm. Images acquire from fluorescence and phase contrast brightfield were analyzed on LAS X core software (Leica microsystems).

### Spot growth assay

The growth phenotype of the EBY\_Xyl1 mutants were analyzed by spot assay as previously described<sup>98</sup>. The yeast cells were cultured in YNB medium without uracil, supplemented with maltose 20 g/L for 24 h at 30°C/200 rpm. Cells were harvested by centrifugation at 3.000 rpm for 5 min, washed three times and resuspend in water with cell density normalized at OD<sub>600</sub>=1.0. Using an automatic multichannel pipette, 5uL of a ten-fold serial dilution were spotted on plates with synthetic complete medium without uracil and different carbon sources. The plates were supplemented with 20 g/L of maltose (control), glucose, galactose, xylose, mannose or fructose. The dose-dependent xylose growth was complementarily analyzed using 5, 10, 30 and 50 g/L. Plates were incubated at 30°C for 3 days (maltose), 5 days (glucose, galactose, mannose and fructose) and 7 days (xylose). All the experiments were performed in triplicate and repeated at least twice. For EBY\_Xyl1\_ *hvk*<sup>0</sup> drop-assay preparation, the cells were inoculated in 50 ml erlenmeyer with YNB medium without uracil, supplement with 3% ethanol for 48 h at 30 °C / 200 rpm. Plates were supplemented with 10 g/L of xylose and glucose (1:1), 10 g/L of xylose and 20 g/L of glucose (1:2) and 10 g/L of xylose and 50 g/L of glucose (1:5). Control plates were glucose 20 g/L, ethanol 3% and xylose 10 g/L. All experiments were performed in triplicate and plates were incubated for 10 days.

### Batch high cell-density fermentation

To test the *C. sojæ* transporters and compare to the traditional *GXF1*, the yeast strains were pre-grown on synthetic complete medium supplemented with 5 g/L of casamino acids and 50 g/L of D-maltose until reach stationary phase. Cells were collected by centrifugation, washed three times with sterile water and resuspend at initial OD<sub>600</sub> of 20. Comparative fermentation experiments were performed using 60 mL of complete medium supplemented with 5 g/L of casamino acids, 10 g/L and 50 g/L of xylose and for simultaneous glucose and xylose co-fermentation, 10 g/L of both sugars were used. The cells were incubated at 30 °C / 200 rpm. Experiments were performed in triplicates and samples were collected to measure optical density and for HPLC analysis.

## Docking and molecular dynamics analysis

The three-dimensional structural models of Cs4130 and *GXF1* were generated using I-TASSER<sup>162</sup> with the Xyle from *Escherichia coli* (PDB code: 4GBY) as reference. The identification of the docking binding sites was performed using the UCSF Dock 6.6 package<sup>163</sup> and UCSF Chimera 1.7<sup>164</sup>. The procedures were the same as described for T4 lysozyme<sup>165</sup>. The cavity sites were suggested by the docking program and confirmed using the crystallographic structures PDB id: 4GBY as reference. Because a large box was employed for the grid calculation, 300 simulation runs were performed for each substrate. The ligand orientations, ligand, receptor and overlap bins were set to 0.2 Å, and the distance tolerance for matching of ligand atoms to the receptor was set to 0.75 Å. The substrates were obtained from the HIC-Up database (<http://xray.bmc.uu.se/hicup/>) and their charges were calculated using UCSF Chimera (AMBER ff12SB force field). After the molecular docking studies for Cs4130 and *GXF1*, the conformations with the lowest grid energy score were computationally solvated in water (SPC216) and equilibrated (after energy minimization) using the GROMACS 4.5.5 package<sup>166</sup> and a CHARMM22<sup>167</sup> force field over a 5ps time duration.

The Normal Mode Analysis (NMA) was applied to investigate the dynamic behavior of Cs4130 and *GXF1*. The Root Mean Square Fluctuation of 20 conformations of the 10 low-frequency modes was obtained by Nonlinear Rigid Block Normal-Mode Analysis Method (NOLB)<sup>136</sup> and evaluated by GROMACS<sup>166</sup> and Pymol<sup>168</sup>. All assays were done mimetizing a lipid bilayer micro-environment.

## Results

### Isolation and molecular identification of C5 yeasts

Only a small number of heterologous sugar transporters capable of enhancing xylose uptake in *S. cerevisiae* have been identified and characterized to date. Previous reports have been demonstrated that xylose transport in recombinant xylose-fermenting *S. cerevisiae* is facilitated by non-specific endogenous hexose transporters with poor affinity for xylose and can be improved by expression of a heterologous xylose transporters<sup>68,81,90,92,117,118</sup>. In order to explore novel C5 uptake systems, we access the microbiome of potential sources such as decayed energy cane and sugarcane and insect pests isolated

from these materials and also from corn plantations. A total of 300 yeast colonies have been recovered from growth in xylose medium and molecularly identified from D1/D2 domains and ITS region. The consecutive batches using the xylose isolation medium was used to enrich the population and select only strains capable of assimilating xylose more efficiently, resulting in only few species from each material that dominated the population. The presumable wide diversity in starting populations colonizing the initial samples was probably severely reduced in response to a selective xylose-enriched environment. Table 2 summarizes the results of the 13 yeast species identified and the isolation source of each specie, confirming the dominance of few species for each material, resulting from adaptation and competition in xylose as a sole carbon source.

**Table 2 |** Isolated C5 yeast species with respective original source

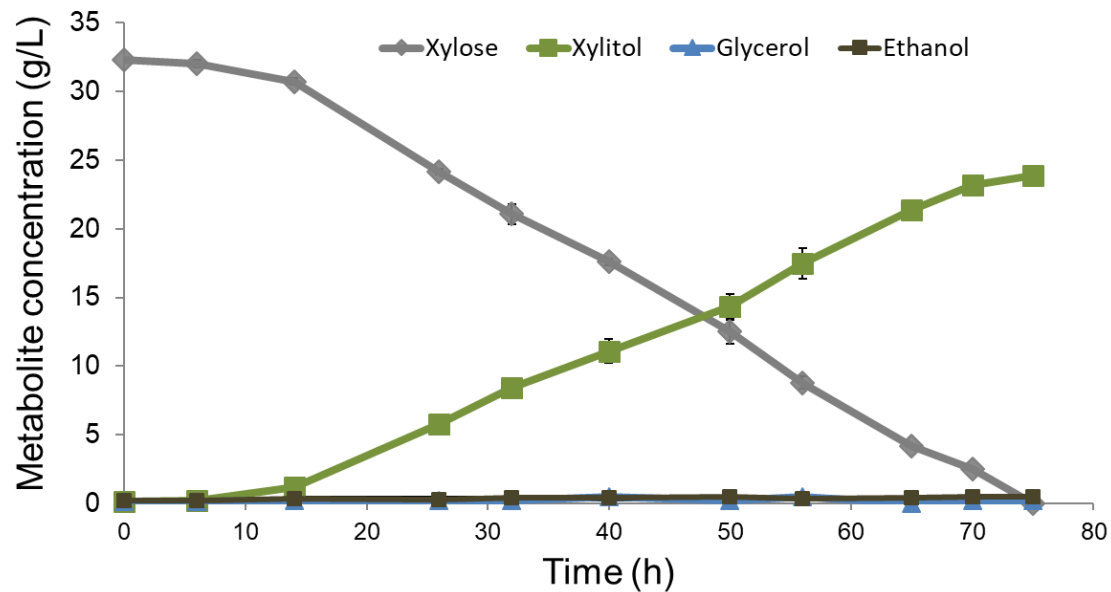
<b>C5 yeast specie</b>	<b>Microbial Isolation Source</b>
<i>Blastobotrys (Arxula) adenivorans</i>	Energy cane
<i>Candida boidinii</i>	Sugarcane straw
<i>Candida pseudointermedia</i>	Corn coleoptera adult
<i>Candida sojae</i>	Energy cane Coleoptera larvae
<i>Candida tropicalis</i>	Energy cane Lepidoptera larvae
<i>Cryptococcus laurentii</i>	Sugarcane colm
<i>Kodameae ohmeri</i>	Corn hemiptera
<i>Meyerozyma caribbica</i>	Energy cane colm
<i>Pichia sp.</i>	Sugarcane Coleoptera adult
<i>Pseudozyma sp.</i>	Corn hemiptera
<i>Trichosporon asahii</i>	Sugarcane root basidia
<i>Trichosporon coremeiformi</i>	Corn coleoptera adult
<i>Wickerhamomyces anomalus</i>	Sugarcane Coleoptera adult

### **Comparative xylose fermentation**

The enriched yeast culture library represents a potential source of novel specific transport systems for xylose. To discover novel potential gene donors' candidates for heterologous expression of C5 transporters in *S. cerevisiae*, we perform comparative fermentations using xylose as carbon source in order to measure fitness and rank higher xylose consumption rates. The results from



Durham tubes indicate that although all strains were able to grow and consume xylose, none of the evaluated species were able to ferment this sugar. The inability to ferment xylose is probably caused by an unbalanced redox cofactor by a xylose reductase with higher affinity to NADPH and a xylitol dehydrogenase NAD<sup>+</sup> dependent, a pattern traditionally present in wild C5 yeasts. All species were able to consume xylose and produce xylitol preferably, with low concentrations of glycerol and ethanol observed as fermentation byproducts. The kinetic fermentative parameters are summarized in table S1. All yeast cells were able to consume xylose with a wide difference in consumption rate. In 54 hours of experiment, starting with 20 g/L of xylose, the differences ranged from 28.5 to 100% of xylose consumed. The four yeasts from *Candida* genus presented the best performance regarding assimilation of xylose with approximately 90 % of the sugar consumed for *C. boidinii* and 100 % for the others in the group. Except for *C. boidinii*, the other *Candida* yeasts present the highest yields of xylitol with a maximum of ~ 0.67 g xylitol/g xylose produced by *C. sojae*. The xylitol concentration increases to 0.8 g/g in hemicellulosic hydrolysate produced from sugarcane using *C. sojae* as a fermentative platform (Figure 1). The *Wickerhamomyces anomalus* isolate also performed well compared to the other yeasts, with a yield of 0.54 g xylitol/g xylose. In YPX medium containing 30 g/L, under aerobic conditions, the *C. sojae* isolate consumed all the sugar in approximately 70 hs, starting with a low cell density, producing practically only xylitol as a fermentation product and very low concentrations of ethanol and glycerol (Figure 1), at yields similar to those reported for other *Candida* yeasts<sup>119-121</sup>. Our group previously sequenced the genomes of two isolates from *Candida* genus - *C. boidinii* and *C. sojae* isolates<sup>122</sup>. We hypothesized that, since *C. sojae* displayed a fast xylose consumption compared to other strains, the isolate is possibly a promising source of novel xylose transporter-encoding genes for heterologous expression in *S. cerevisiae*.

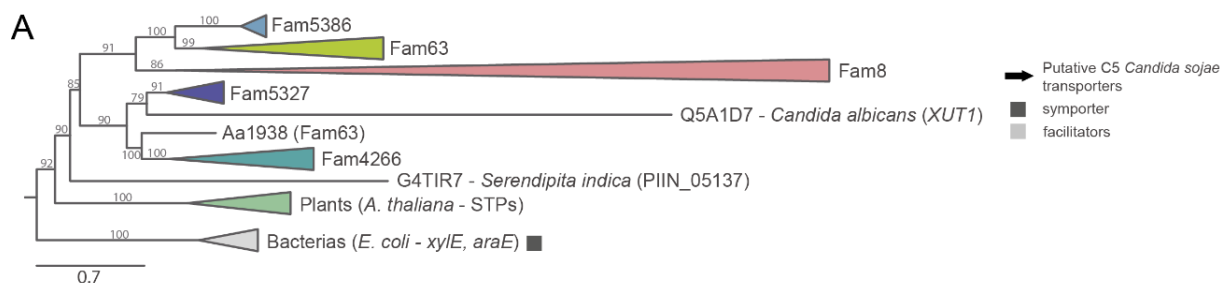


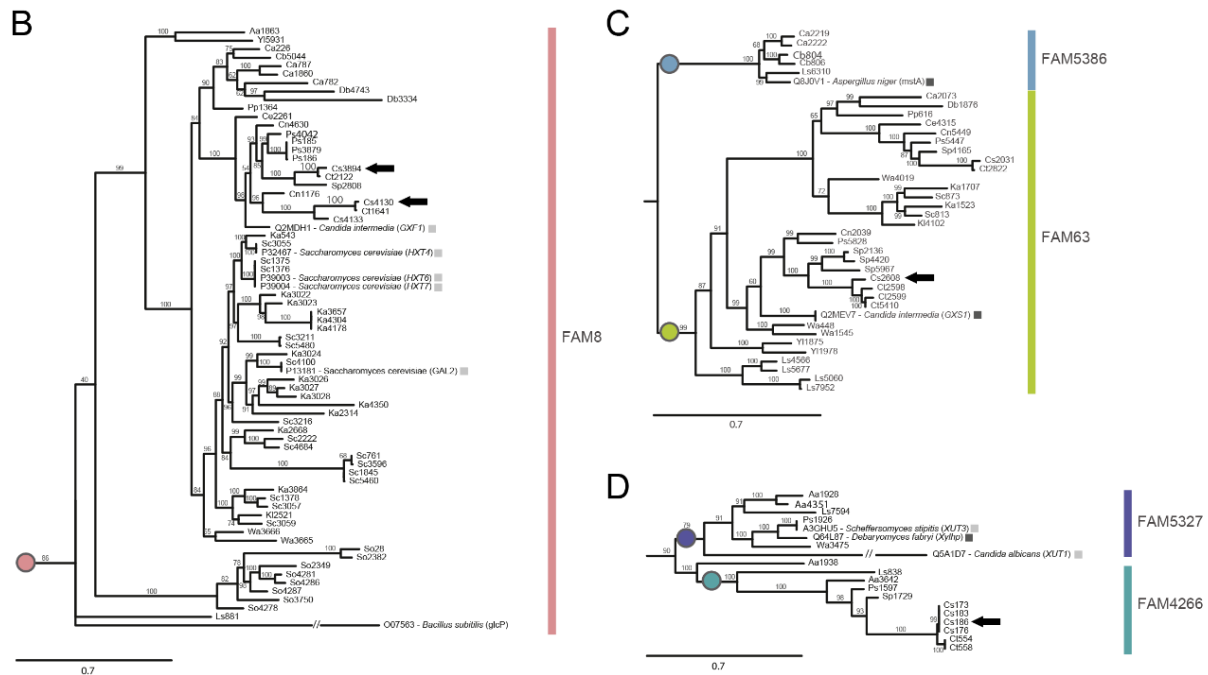
**Figure 1 | *Candida sojae* fermentation of xylose as a sole carbon source.** *C. sojae* were cultivated in YPX (20 g/L) in batch fermentation with a low initial optical density of 0.5. The fermentation was performed in triplicate and error bars represent standard deviation from the average of values.

### Genome mining of xylose transporter-encoding genes from *C. sojae*

A set of known microbial xylose transporters from public databases were used as a bait to screen putative and promising transporter genes in *C. sojae* genome and classify as symporter or facilitator. Transporter gene families had all their genes clustered between themselves in the phylogeny, except Aa1938, which was closer to Fam4266 than to the rest of Fam63, the one it was clustered in OrthoMCL (Figure 2A). This shows that OrthoMCL successfully separated homolog gene families for yeasts transporters. Four of the five gene families had at least one known xylose transporter (bold in figure 2) clustered in their clades allowing some inference on the phenotype of each family genes. Comparing yeast transporter genes with other fungi, the *mstA* transporter from *A. niger* is in the family 5386 tree and grouped with a *Lipomyces starkeii* transporter (Ls6310). This observed relationship is not expected by species phylogeny as *A. niger* is an *Ascomycete* from a different Class. In the same way, the bacterial transporter *glcp* from *B. subtilis* clustered with family 8 yeast transporters instead of clustering with other bacterial transporters as expected. The incongruence among gene trees and the species phylogeny may arise from different factors as horizontal gene transfers, incomplete lineage sorting and also as similar selective pressures

leading to sequence convergence. Gene family expansions are easily recognized throughout yeast transporter phylogeny (Figure 2). In family 8, Cs3894 and Cs4130 appear as copies from a putative gene duplication in the base of *Pichia*, *Spathaspora* and *Candida* clade (Figure 2B), where the first shares a subclade that follow species phylogeny<sup>123</sup> and the second in a subclade exclusive of *C. sojae*, *C. tropicalis* and *C. tanzawanensis*, putatively being lost in *Pichia* and *Spathaspora* lineages. In family 4266, Cs186 is a putative result of many duplication events inside *C. sojae* lineage (Figure 2D). The high duplication rate coupled with high conservation of these genes suggest that these transporters are dosage-dependent as reported for hexose transporters<sup>124</sup>. For dose dependent genes, it is potentially adaptive to increase gene number maintaining sequences characteristics because it implies in higher transcription and a larger number of proteins working on the cell. In this way, transporter genes that undergo recent duplication events might be evolving under a selective process. By using an evolutionary approach for xylose transporters prospection of gene candidates for industrial applications, we hypothesize that transporters from our dataset more closely related to xylose transporters known from literature would have a similar capacity for uptaking this pentose. We then selected as candidates closely grouped genes to these transporters, such as Cs4130 and Cs3894 (close to *GXF1*), as well as some more distantly grouped - Cs186, a whole orthogroup distant from *XUT1*, *XUT3* and *Xylhpl* and Cs2608, close to *GXS1*. Based on alignment, Cs186, Cs3894 and Cs4130 are facilitators and Cs2608 is a symporter.



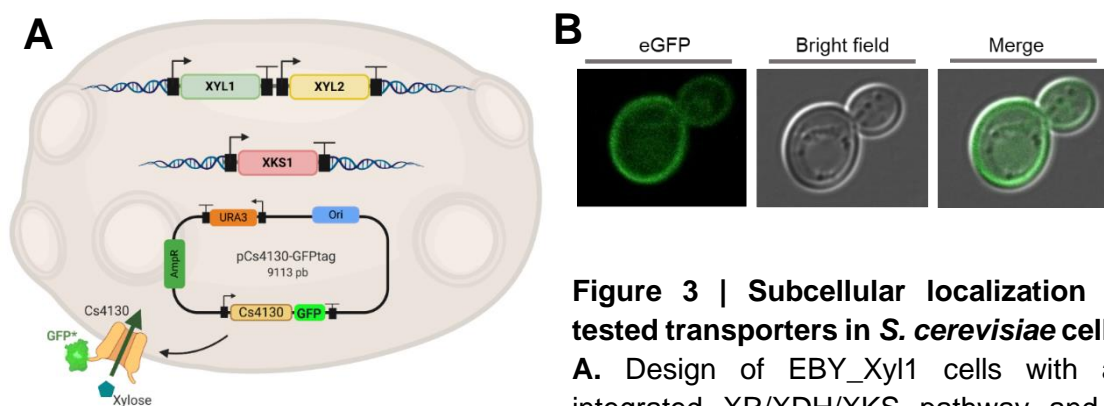


**Figure 2 | Phylogenetic inference for transporter gene families in yeasts.** The gene tree was reconstructed using maximum likelihood method as implemented in IQtree, with branch support estimated using 1000 bootstrap replications. A collapsed version showing relationships among families and the known transporters for other fungi, plants and bacteria is shown in **A**. Expanded versions for each family of yeast transporters are in **B-D**. Already characterized xylose transporters are identified in squares. Light gray squares demonstrated sugar transporters from facilitator class, and dark gray square from symporter class. Arrows indicate putative C5 transporter proteins from *Candida sojae*.

### **C. sojae** transporters mediate uptake of various carbon sources

The substrate spectrum from the selected sugar transporters Cs186, Cs2608, Cs3894 and Cs4130 was tested on the strain EBY\_Xyl1, a modified yeast cell derived from EBY.VW4000, which lack transporters of the families *HXT1-17* and *GAL2*, and is not able to grow on glucose as a sole carbon source<sup>25</sup>. EBY\_Xyl1 was designed with the integration of an expression cassette containing the genes *XYL1* and *XYL2* from *S. stipitis* in the vicinity of centromere five, and an additional copy of xylulokinase (*XKS1*) was stably integrated near *CEN2*. The three genes were expressed under the control of different strong constitutive promoters of genes from the glycolytic pathway of *S. cerevisiae* (Table 4). The candidate transporters were cloned in a multi-copy vector pRS426 using the promoter and terminator regions of *TDH1* from *S. cerevisiae*. As a positive transport-system control, we choose the traditional xylose-facilitator *GXF1* from *C. intermedia*, since this gene is well-known as the best heterologous xylose

transporter already described on literature<sup>81,92</sup>. The plasmids expressing the cloned putative C5 transporters were used to transform the EBY\_Xyl1 yeast cell. The candidate transporters were fused to a codon-optimized GFP-mUkG1 sequence to confirm the correct cellular localization in the yeast cell membrane and observed using a confocal microscope. The yeast cells expressing the GFP-tagged plasmids presented a fluorescence halo in the cell periphery indicating the correct localization in the cell membrane as observed in figure 3. However, the presence of fluorescence in the cytoplasm in few yeast cells was observed, indicating possible degradation through ubiquitination or incorrect folding caused by a multi-copy gene expression.

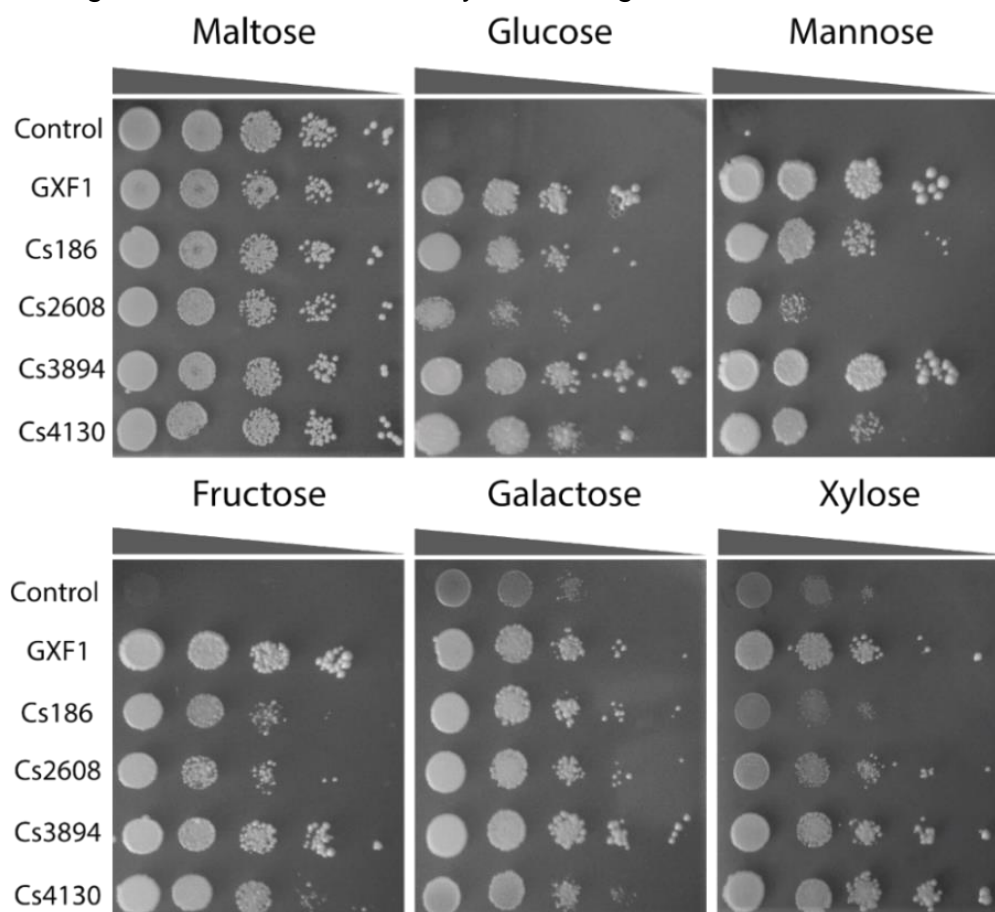


**Figure 3 | Subcellular localization of tested transporters in *S. cerevisiae* cells.**

**A.** Design of EBY\_Xyl1 cells with an integrated XR/XDH/XKS pathway and a multi-copy vector expressing the transporter fused with a codon-optimized eGFP-ymUkG1. **B.** A confocal image confirm the presence of a fluorescence halo in the cell periphery indicating the correct localization in the cell membrane. Bright field and fluorescence images were taken by a confocal microscope and merged. Scale bars = 5  $\mu$ m.

We analyzed the substrate spectrum of EBY\_Xyl1 mutants carrying the sugar transporters using six different carbon sources on solid plate – 2% maltose (control), mannose, fructose, xylose, glucose and galactose. The transformants were grown overnight to the exponential phase and spotted in 10-fold serial dilutions onto solid plates. This analysis shown that *C. sojae* transporters restore growth of EBY\_Xyl1 strain in all sugars tested (Figure 4). Curiously, the control strain EBY\_Xyl1 carrying pRS426 plasmid shown a slight background on galactose and xylose. Parental strain EBY.VW4000 was previously described to show basal growth on galactose<sup>125</sup>. As this background was similar comparing these two sugars, probably, the remaining transporters on the EBY.VW4000 strain which confer growth on galactose could be responsible for this behavior observed on xylose. However, on other sugars tested, except maltose (control),

no growth was demonstrated using the control EBV\_Xyl1\_pRS426 strain. All the tested transporters were functional and able to restore growth in all evaluated carbon sources showing a broader substrate specificity. The only exception was Cs186 which presented a slight background growth similar to pRS426 control on xylose, indicating that this transporter is unable to assimilate this sugar. Cs3894 and *GXF1* were able to support better and similar patterns of growth on glucose, mannose and fructose compared to the other transporters. The strain expressing the Cs2608 present a poor growth on mannose and glucose compared to the other substrates. And Cs3894 and Cs4130 presented a superior growth on 20 g/L of xylose and, interestingly and not expected, the growth of the strain expressing Cs4130 were better on xylose than glucose.



**Figure 4 | Functional profile of sugar affinities of *C. sojae* transporters. A.** Drop-assay analysis on different sugars, 20 g/L, of EBV\_Xyl1 strain expressing the indicated transporters and the control vector pRS426. Plates were incubated at 30°C for 3 days (maltose), 5 days (glucose, mannose, fructose and galactose) or 7 days (xylose). Yeast cells inoculum were grown in complete medium without uracil supplemented with maltose for 24 hours. Initial OD<sub>600</sub> was established at 1, before a serial dilution. Yeast dilution are represented by the triangle above each spot. All experiments were performed in triplicate.

## Growth on different range of xylose concentrations

The budding yeast *S. cerevisiae* successfully evolved and adapted to survive by expanding the repertoire of high and low-affinity glucose transporters and by adjusting the expression regulation levels to be able to capture sugar in different concentrations<sup>126</sup>. The extracellular sugar concentration on growth media reflects the transporters expression levels and stability in cellular membrane and this ability allows the survival against adversity in different environments and conditions<sup>127</sup>. The expression of genes from HXT family in *S. cerevisiae* is mostly regulated by differences in glucose levels<sup>128</sup>. For example, *HXT7* is a high-affinity transporter, upregulated in low glucose levels and degraded in high concentrations<sup>127,129,130</sup>. Similarly, *HXT3* is a low-affinity transporter, expressed on high glucose concentration, but are inhibited on low glucose environments<sup>127,131</sup>.

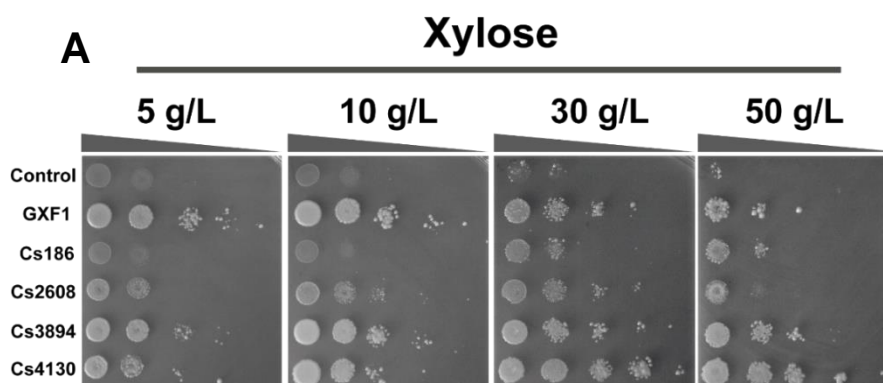
*S. cerevisiae* was not originally designed to assimilate xylose and does not have a vast repertoire of transporters capable to assimilate this pentose in different concentrations, and neither present a fine-tuned expression regulation adjusted for this purpose. The heterologous xylose transporters previously described in the literature were able to allow growth of engineered *S. cerevisiae* strains only in low concentrations. *GXF1* is an example and allows growth of *S. cerevisiae* only in very low xylose concentrations. The design of microbial cell factories to efficiently convert biomass sugars in biorefineries requires the introduction of transporters and adjustment of the regulation levels to create strains that are able to incorporate xylose in different concentrations available in lignocellulosic hydrolysates. In this context, in our experiments, we introduce a broader screening range in order to identify and characterize putative xylose transporters with different properties.

The growth of EBY\_Xyl1 mutants carrying *C. sojae* transporters and *GXF1* as positive control was compared in solid medium with 0.5, 1, 2, 3 and 5% of xylose as a sole carbon source. The facilitator *GXF1* restores growth in all xylose concentrations. Analyzing the results of *GXF1*, it is possible to notice a decrease in growth with increasing concentrations of xylose. Although this facilitator allows a higher growth of EBY\_Xyl1, compared to all other *C. sojae* transporters on low xylose concentrations, 5 and 10 g/L, on intermediary and high concentrations, a

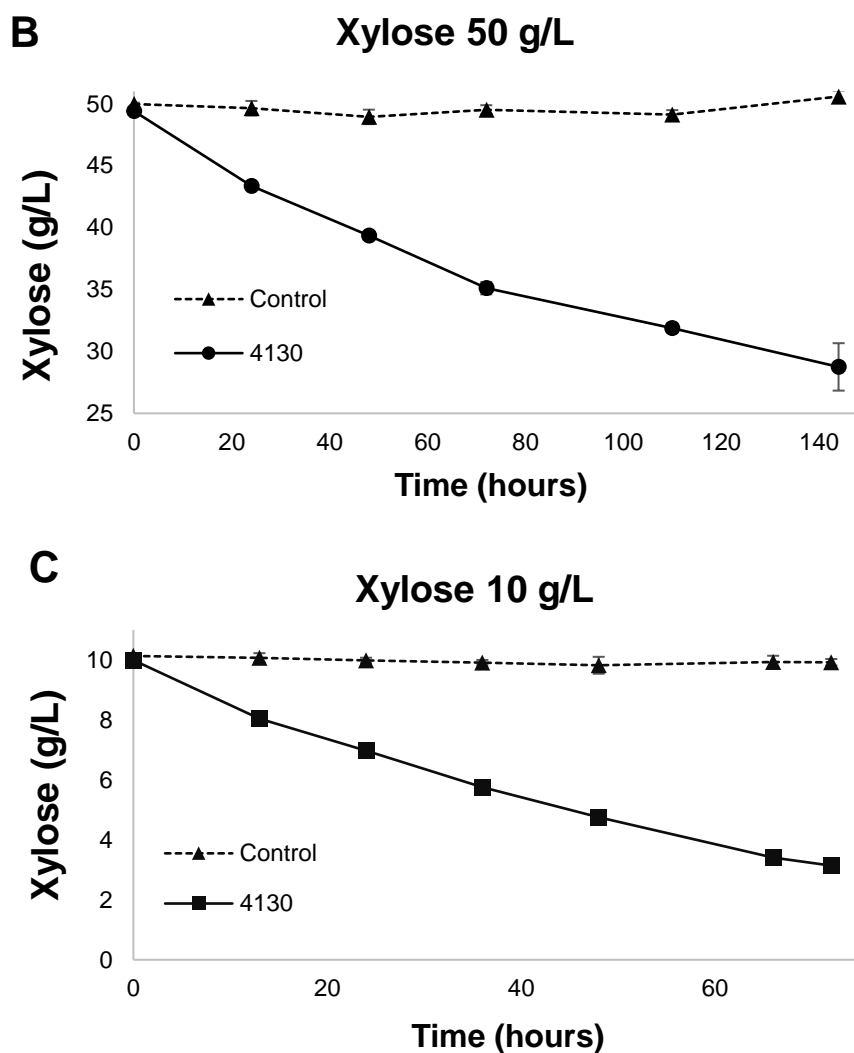


natural inhibition was observed. Previously studies reported that *S. cerevisiae* mutants expressing *GXF1* presented a significant growth improvements only at low xylose concentrations<sup>88,92</sup>. Surprisingly, the *C. sojae* transporters Cs3894 and Cs4130 demonstrated an opposite behavior allowing improvements on growth in intermediary and higher xylose concentrations, 30 g/L and 50 g/L respectively, figure 5A. The EBY\_Xyl1 cell expressing Cs4130 showed a better growth pattern compared to the other transporters in high concentrations of xylose, showing a natural loss of inhibition in contrast to commonly observed in *GXF1*. The spot-assay experiment in different xylose concentrations were repeated three times, and each experiment was conducted in triplicate to confirm the growth patterns.

The spot-assay results were also confirmed in liquid media in a high cell density fermentation assay, using EBY\_Xyl1\_pRS426 as a control strain and xylose as only source of carbon. Since Cs4130 presented the most interesting results, we focus the new experiments only on this facilitator. The yeast cell expressing Cs4130 were able to assimilate xylose and grow in both 10 and 50 g/L, although it has shown higher volumetric productivity in superior concentrations (Figure 5A and B). In both concentrations, in the first hours of collection and analysis, productivity was higher and decreased during fermentation. Besides the constrain in transport, the uptake can be impaired by a slow xylose catabolism present in EBY\_Xyl1, as this strain will have redox balance limitations with the XR-XDH pathway. The degree of xylose uptake in engineered *S. cerevisiae* is dependent on the downstream pathway flux and the xylose concentration in the medium<sup>92</sup>. But it is clearly observed the higher productivity in 50 against 10 g/L of xylose, with productivity differences of up to 50.03% in the first 24 hours of fermentation and 52.12% in 72 hours, reinforcing the non-conventional pattern of growth of Cs4130 observed in spot-assay experiments.







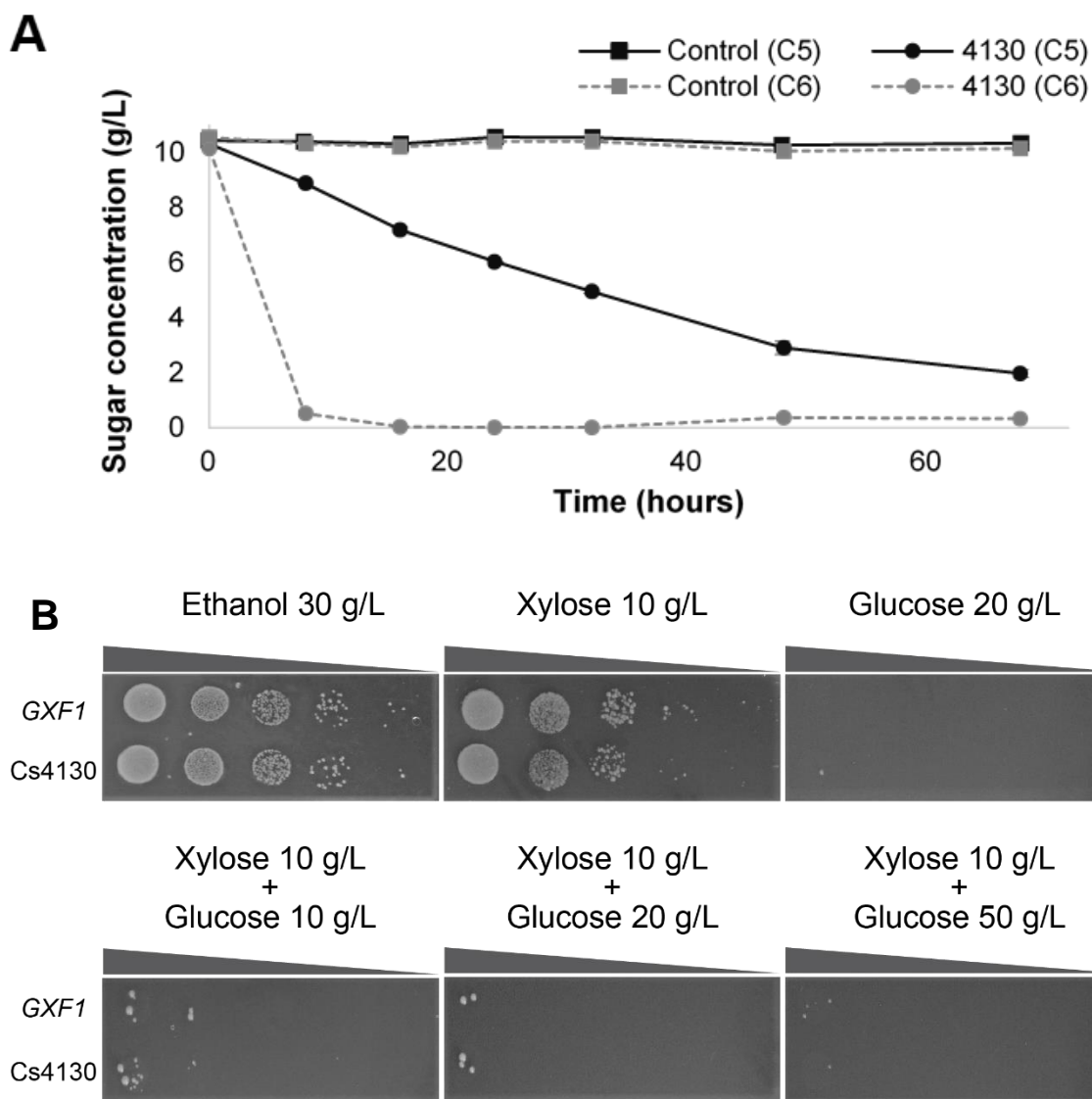
**Figure 5 | Functional profile of concentration-dependence of *C. sojae* transporters.** **A.** Growth on different range of xylose concentrations demonstrating an uptake/concentration dependency. Yeast dilution are represented by the triangle above each spot. All experiments were performed in triplicate. **B.** High cell density fermentation of Cs4130 in xylose at 50 g/L (B) and 10 g/L (C). The fermentations were conducted on YNBX medium at 30°C, 200 rpm. Control represent the EBV\_Xyl1 harboring empty vector pRS426. Yeast dilution are represented by the triangle above each spot. Error bars represent standard deviation from average measures from triplicates.

### Simultaneous consumption of glucose and xylose on Cs4130

The main sugars of 2G hydrolysates from lignocellulosic biomass are glucose and xylose. The efficient co-assimilation of these two sugars are essential to competitively convert biomass into fuels and chemicals<sup>70</sup>. All xylose transporters previously described in literature are severely inhibited by glucose. Membrane transporters present higher affinity for glucose, limiting the efficiency

in capture xylose due to substrate competition and consequently decreasing fermentation efficiency in the presence of both sugars<sup>132</sup>. Therefore, we tested the simultaneous consumption of glucose and xylose on a high-density fermentation using a C6/C5 mixture, with 10 g/L of each sugar and EBY\_Xyl1 strain expressing Cs4130 (Figure 6A). This fermentation showed a faster glucose consumption in the first 10 hours. A very small amount of xylose was consumed in this period, confirming, as expected, a higher affinity for glucose compared to xylose. As an example, the facilitator *GXF1* is known to have a higher glucose affinity ( $K_m \text{ glucose} = 2.0 \pm 0.6$ ) compared to xylose ( $K_m \text{ xylose} = 48.7 \pm 6.5$ ),<sup>84</sup>. The preferential glucose consumption confirmed that Cs4130 does not bypass glucose inhibition, similarly as observed for *GXF1*.

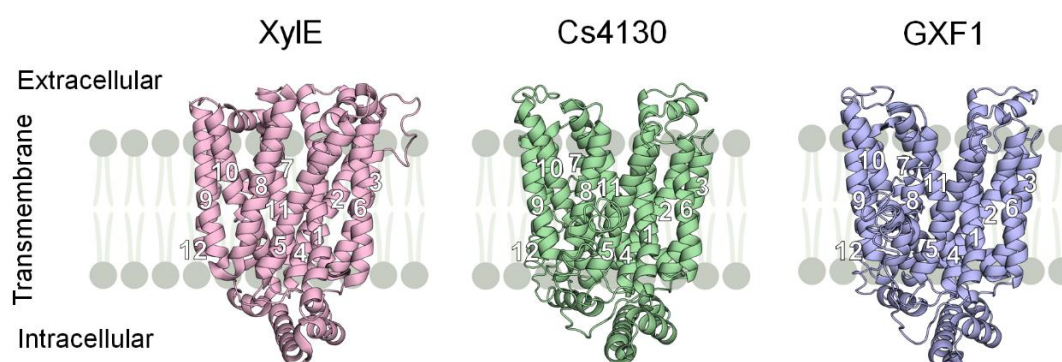
To further evaluate the glucose inhibition of Cs4130, from EBY\_Xyl1, we constructed the strain EBY\_Xyl1\_ *hxxk*<sup>0</sup> by disrupting the glycolytic pathway through gene knockout of three hexokinases: *HXXK1*, *HXXK2* and *GLK1*, and further restored sugar uptake by expression of individual heterologous transporters, Cs4130 and *GXF1*. The derived strain can normally assimilate the glucose molecule from extracellular media but is not capable to metabolize this sugar as a carbon source. Therefore, it is a platform to analyze the glucose inhibition in the mixture of sugars. Since glucose or maltose cannot be used anymore for growth, we used 3% ethanol as a positive control. EBY\_Xyl1\_ *hxxk*<sup>0</sup> strains harboring individual heterologous transporters, showed, as expected, growth on xylose and impairment on glucose uptake. The glucose inhibition test was performed in three different concentrations of C5/C6 mixtures on solid media. The xylose concentration was maintained on 10 g/L and glucose were added to final concentrations of 10 g/L, 20 g/L and 50 g/L, respectively. The inhibition test showed that yeast cells expressing both transporters were unable to restore significant growth on the different mixtures, even on low glucose concentrations (Figure 6B and C). This phenotype indicates that these heterologous transporters show natural higher preference for glucose over xylose, which leads to growth impairments on *hxxk*<sup>0</sup> strains on the presence of both sugars.



**Figure 6 | Glucose inhibition of Cs4130 in glucose and xylose mixtures. A.** High cell density fermentation of Cs4130 in xylose and glucose co-fermentation. The fermentations were conducted on complete medium at 30 °C, 200 rpm. Control represent the EBV\_Xyl1 harboring empty vector pRS426. Error bars represent standard deviation from average measures from triplicates. **B.** Drop-assay analysis on EBV\_Xyl1\_ *hvk*<sup>0</sup> strain expressing the indicated transporters. Cells were grown on complete medium without uracil supplemented with 3% ethanol for 56 hours before a serial dilution. The initial OD<sub>600</sub> was established on 1. Yeast dilution are represented by the triangle above each spot Plates were incubated at 30 °C for 7 days.

## Docking and Normal Mode Analysis (NMA) point out different behavior of xylose bound Cs4130 and GXF1

The modelled structures of Cs4130 and GXF1 were acquired in I-TASSER server<sup>133</sup>, both showing the typical fold of MFS transporters: 12 transmembrane (TM)  $\alpha$ -helix with the N- (TM1-TM6) and C- (TM7-12) termini facing the intracellular side<sup>134</sup> (Figure 7). Both models showed high TM-scores ( $>0.6$ ) with GLUT family of transporters, typical of mammals (PDB code: 5C65, 4YBQ and 4PYP); Xyle and GlcPSE, from bacteria (4GBY, 4LDS); and STP10, from *Arabidopsis thaliana* (6H7D).

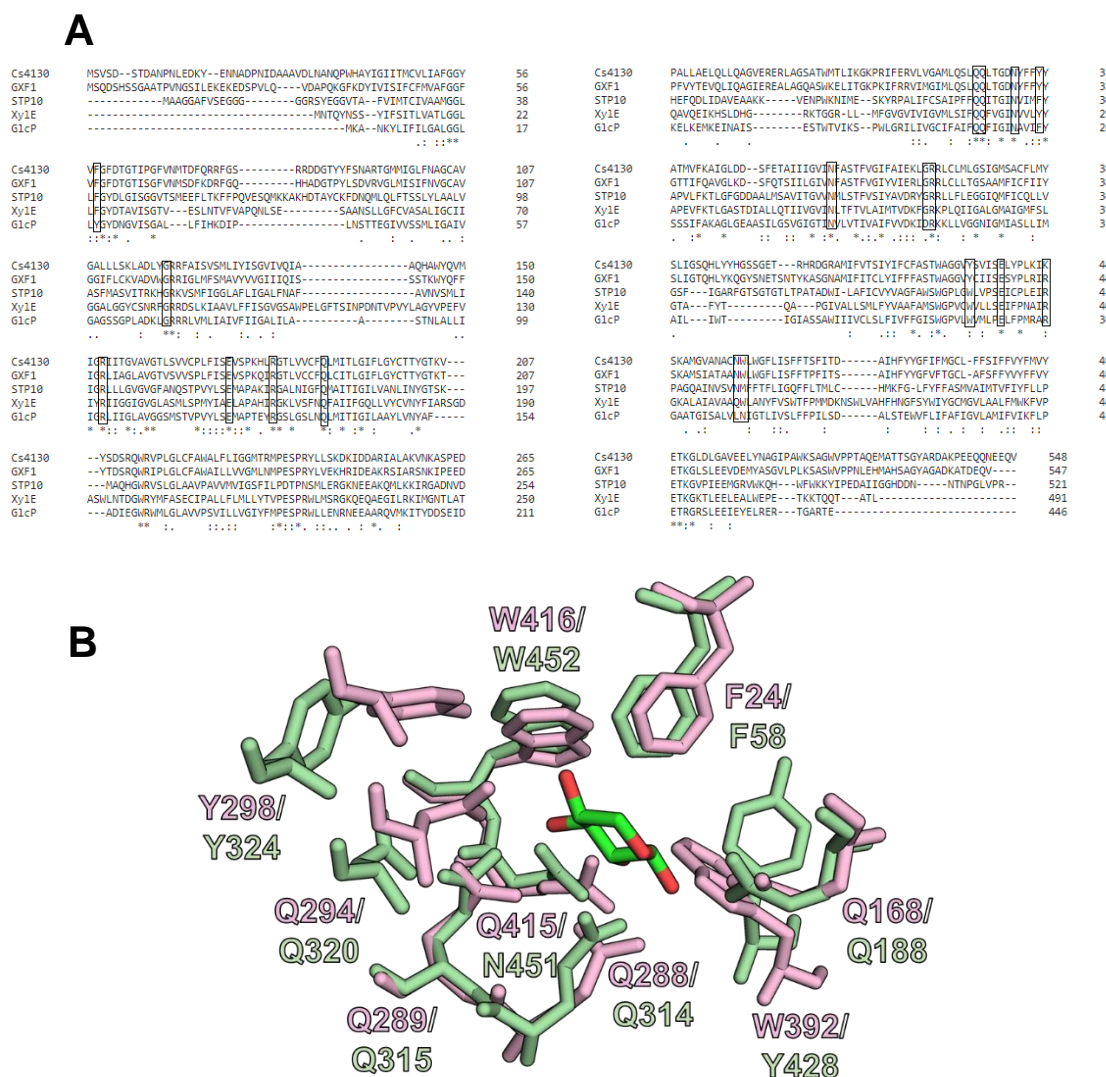


**Figure 7** | Crystallographic and modelled structures of Xyle (4GBY) and Cs4130/GXF1, respectively, embedded in cytoplasmic membrane. Xyle, Cs4130 and GXF1 show the typical fold of MFS transporters. The 12 transmembrane  $\alpha$ -helix in Cs4130 and GXF1 were assigned based in Xyle.

Xyle, a D-xylose:H<sup>+</sup> symporter found in *Escherichia coli* showing 27.02 and 25.56% identity (amino acids) with Cs4130 and GXF1, respectively, was employed as a model in this work. A seminal work revealed the outward-facing/partly occluded crystallographic structure of the xylose (4GBY) and glucose-bound (4GBZ) Xyle. In addition, some residues mutated in Xyle resulted in a decrease or abrogation of xylose transport<sup>135</sup>. Their counterparts in Cs4130, GXF1, STP10 and GlcPSE are conserved suggesting their role in sugar transport (Figure 8A, Table S5).

Through docking analysis, Cs4130 and GXF1 bind to glucose, fructose, galactose, mannose and xylose (Table 3), which agrees with Figure 4 (spot assays with sugars). Even that Cs4130 and GXF1 share the same sugar-binding regions (Figure 8A), the free energy binding (kcal/mol) are distinct taking the same monosaccharide into consideration, except for fructose. It possibly reflects

the differences in the vicinity of the residues bound to the sugar. Xylose showed a slightly low free energy binding, and thus, high binding affinity for Cs4130 compared to *GXF1* (Table 3).



**Figure 8** | Alignment of amino acid sequences of Cs4130, *GXF1*, STP10, Xyle and GlcPSE (A) and superposition of Xyle (pale pink) and Cs4130 (pale green) intermembrane residues related to xylose (green) transport (B). The conserved residues are framed and were assigned based in Sun et al. (2012).

An extensive analysis of Xyle, Cs4130 and *GXF1* was carried out to infer the molecular mechanisms related to the improved assimilation of xylose by Cs4130 over *GXF1*. Functional studies with Xyle<sup>135</sup> provided valuable tips about the behavior of Cs4130 and *GXF1* bound to xylose. The distance of F24, G83, R133, E153, R160, Q168, Q288, Q289, N294, Y298, N325, G340, R341, W392, E397, R404, Q415 and W416 to xylose (Å) was measured using 4GBY as reference (Table 4). Given the superposition of these residues in Xyle, Cs4130

and *GXF1* (Figure 2B, highlighting the TM residues), the distance to the ligand was considered as the same for all targets.

**Table 3** | Free energies of binding (kcal/mol) of Cs4130, *GXF1* and XylE with monosaccharides through docking analysis

GRID SCORE (kcal/mol)					
Transporter	Monosaccharides				
	Glucose	Fructose	Galactose	Mannose	Xylose
<b>Cs4130</b>	-32.9	-29.7	-30.5	-31.6	-28.8
<b><i>GXF1</i></b>	-29.3	-29.9	-29.3	-29.8	-26.0
<b>XylE</b>	-31.0	-30.4	-33.3	-33.9	-27.4

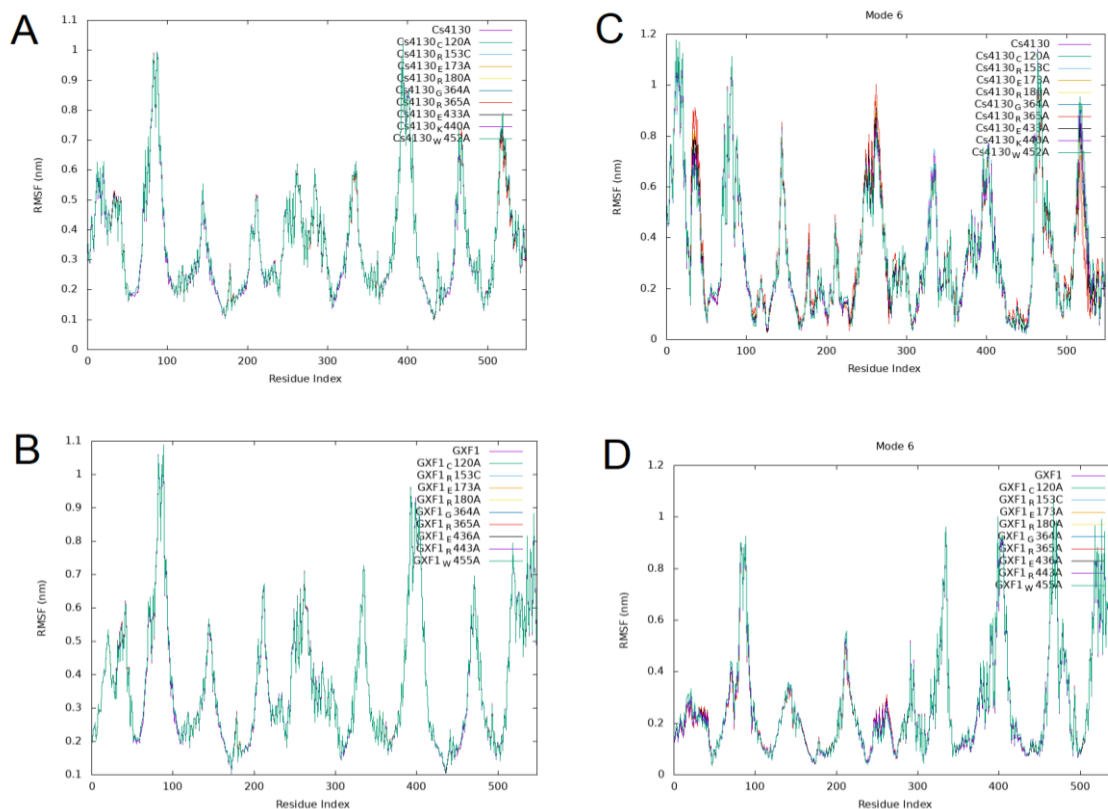
Some residues which interact with xylose by hydrogen bonds [Q168 (2.3 Å from xylose), Q288 (2.9 Å), N294 (2.9 Å) and Y298 (4.3 Å)] were mutated to alanine and evaluated through docking analysis (Table 4). N325 was chosen as a control given that N325A did not affect the transport of xylose, probably because it is not involved in sugar binding (distance of 6.7 Å from the ligand)<sup>135</sup>. It is noteworthy that small variations in the free energy binding (~ +0.7 kcal/mol) (Table 4) change the behavior of XylE given that Q168A, Q288A, N294A and Y298A versions lost their ability to transport xylose<sup>135</sup>. In addition, the free energy binding of the N325A was equal to the wild-type (WT) (-27.4 kcal/mol) validating our docking data by agreeing with the functional studies carried out by Sun et al. (2012)<sup>135</sup>. Taking the results of Table 4 into consideration, mutants 1/5 and 2 of Cs4130 and *GXF1*, respectively, would probably result in altered xylose transport.

Cs4130 seems to be more susceptible to phenotypic changes upon mutations than *GXF1* given the variation in the free energy binding of mutants compared to the respective WT. This behavior was also observed for Cs4130 and *GXF1* mutants I191A, I195A, F322A, F323A, I416A (I419A in *GXF1*), F419A (F422A in *GXF1*) and Y428A (Y431A in *GXF1*) (Table S6). These residues were chosen in Cs4130 based in their closeness to the ligand. Then, changes in xylose transport are expected in mutants Q188A, I191A, I195A, F322A, F324A, N349A, I416A and F419A for Cs4130, but only in I191A, I195A and Q314A regarding *GXF1*.

Docking analysis was applied to investigate the effect of xylose-closeness residues in transport. The possible roles of distant residues in the protein dynamics were calculated using the nonlinear rigid block normal-mode analysis method<sup>136</sup> looking for the low-frequencies normal modes which reveal the behavior of sugar-bound Cs4130 and *GXF1*.

The average of the dynamical regimes (Root Mean Square Fluctuation, RMSF) of WT Cs4130 and *GXF1* did not point out any differences in the transporters' behavior (data not shown). Then, the residues G120, G153, E173, R180, G364, R365, E433 (E436 in *GXF1*), K440 (K443 in *GXF1*) and W452 (W455 in *GXF1*), with distances from xylose underlined in Table 4, were mutated to alanine in Cs4130 and *GXF1* and the dynamical regimes were compared with the respective WTs.

Figure 9A and C show the RMSD average of Cs4130 and *GXF1*, respectively. Mutant R365A Cs4130 changes the behavior of WT Cs4130 in the low-frequency modes 6-8 (Figure 9B, only mode 6), also affecting the residues 50-60 (TM1), 250 (cytoplasmatic region) and 290 (cytoplasmatic region), probably altering the transport of xylose in Cs4130. Regarding *GXF1*, the dynamical regimes were quite stable considering all low-frequency modes (Figure 9D).



**Figure 9** | Average of the dynamical regime of Cs4130 (A) and *GXF1* (B). The low-frequency modes 6 is shown for Cs4130 (C) and *GXF1* (D). The flexibility of the residues were evaluated using the root mean square fluctuation (RMSF).

**Table 4:** Distance of each residue mutated in XylE from xylose (Å) and free energies of binding (kcal/mol) of Cs4130, *GXF1* and XylE with xylose through docking analysis

GRID SCORE (kcal/mol)				
	XylE	Cs4130	GXF1	Distance to xylose (Å)
<b>Wild-type</b>	-27.4	-28.8	-25.8	-
	F24	F58	F58F	<b>3.9</b>
	G83	G120	G120	<u>22.0</u>
	R133	R153	R153	<u>11.7</u>
	E153	E173	E173	<u>17.8</u>
	R160	R180	R180	<u>14.3</u>
<b>Mut 1</b>	Q168 (-28.3/0.9)	Q188 (-26.4/2.4)	Q188 (-25.5/0.3)	<b>2.3</b>
<b>Mut 2</b>	Q288 (-26.4/1.0)	Q314 (-28.1/0.7)	Q314 (-24.9/0.9)	<b>2.9</b>
	Q289	Q315	Q315	<b>3.0</b>
<b>Mut 3</b>	N294 (-28.1/0.7)	N320 (-28.2/0.6)	N320 (-25.2/0.6)	<b>2.9</b>
<b>Mut 4</b>	Y298 (-26.4/1.0)	Y324 (-28.1/0.7)	Y324 (-25.8/0)	<b>4.3</b>
<b>Mut 5</b>	N325 (-27.4/0)	N349 (-26.9/1.9)	N349 (-25.7/0.1)	<u>6.7</u>
	G340	G364	G364	<u>15.8</u>
	R341	R365	R365	<u>14.3</u>
	W392	Y428	Y431	<b>3.1</b>
	E397	E433	E436	<u>14.2</u>
	R404	K440	R443	<u>16.1</u>
	Q415	N451	N454	<b>2.7</b>
	W416	W452	W455	4.8

Residues next and far from ligand (xylose) are highlighted in bold and underlined, respectively.

In all mutants, the native residue was replaced by alanine.

The free energy of binding (kcal/mol) of mutants and the difference from the wild-type are between parentheses.

Mut, mutant.

## Discussion

Metabolic engineering of microbial platforms for 2G biorefineries relies on the development of an efficient transport system to capture all the sugars from biomass<sup>63,64</sup>. The xylose uptake across plasma membrane is a critical limiting step in recombinant yeast strains, reducing the efficiency of downstream pathways and conversion in high value added bioproducts. In the course of evolution, *S. cerevisiae* was highly equipped with a diverse and multipurpose



transport system in order to deal with different conditions and challenging environments. The whole-genome duplication event expanded the repertoire of glucose transporters that *S. cerevisiae* has at its disposal and allows efficient growth in a wide range of concentrations<sup>126</sup>. *S. cerevisiae* present 18 genes encoding hexose transporters - Hxt1p-Hxt17p and Gal2p, that belongs to a Major Facilitator Superfamily (MFS). The endogenous transporters can be defined by the different degree of glucose affinities, from low ( $K_m$  of 50–100 mmol l<sup>-1</sup>) to high-affinity ( $K_m$  of 1–2 mmol l<sup>-1</sup>), and are dependent from glucose concentration in extracellular environment<sup>74,103</sup>. The transport system is also fine regulated at the transcriptional and post-translational levels to adapt to different glucose concentrations<sup>131</sup>. During xylose assimilation, engineered strains of *S. cerevisiae* does not have all these transport system options. And neither has a regulated signaling system that allows the pentose uptake in a range of different concentrations. Using an in vivo biosensor panel associated with three major signalling pathways, a recent work demonstrates that an engineered C5 *S. cerevisiae* strain senses xylose differently from glucose. The signal triggered at high concentrations of xylose was similar to the carbon limitation signal observed at low glucose levels<sup>137,138</sup>. In addition to the differences observed in cell signaling, in order to overcome transport limitations that results in low assimilation rates, a global modification at the xylose transport system is necessary with the integration of facilitators capable to transport xylose in a wide range of sugar concentrations.

The nature already evolved many pentose-assimilating yeast species that transport xylose with high and low affinities and the functional expression of only a few heterologous transporters were demonstrated to improve xylose uptake in recombinant *S. cerevisiae* strains<sup>68,81,83,85,91</sup>. For example, the xylose transporter *SUT1* isolated from *S. stipitis* were able to improve xylose consumption by 25% and increase ethanol yield by 17% on recombinant *S. cerevisiae* cells<sup>68</sup>. The transporters, An254, AXT16, XYLH7, and MGT051968 from C5 yeasts *Neurospora crassa*<sup>85</sup>, *Pichia guilliermondii*<sup>39</sup>, *Debaryomyces hansenii*<sup>86</sup> and *Meyerozyma guilliermondii*<sup>118</sup>, respectively, were also able to restore significant xylose growth on expressing strains. Even *Arabidopsis thaliana* xylose transporters, At5g17010 and At5g59250, were responsible for 25% and 40% of improvements in xylose consumption rate, respectively, on *S. cerevisiae*

strains<sup>117</sup>. As described above, the facilitator *GXF1* isolated from *C. intermedia*<sup>84</sup> is described as the best heterologous xylose transporter, although, it was only capable to capture xylose efficiently at low sugar concentrations<sup>81,89,92</sup>. *GXF1* transporter shown an intermediary affinity to xylose and a glucose preference 20-fold higher<sup>84</sup>.

By assessing microbiomes such as the digestive tract of plague insects and several decayed biomasses, we isolated several pentose yeast species capable of naturally using xylose (Figure 3A). The directed cultivation in xylose with successive batches limited the natural diverse population by substrate competition and allowed the isolation and molecular identification of only few species able to consume xylose more efficiently. Since these strains competed to grow in xylose, an efficient transport system to capture the C5 sugar is expected. Comparative fermentations select the yeast *C. sojae* as a potential source of high-affinity transporters. *C. sojae* was isolated from the intestine of a *Coleoptera* larvae collected internally from an energy cane stalk (Table 2). From genome analysis<sup>122</sup>, well-known eukaryotic xylose transporters, such as *HXT7*<sup>71,72,96</sup>, *GAL2*<sup>71,96</sup>, *XUT1*<sup>81</sup>, *XUT3*<sup>81</sup>, *SUT1*<sup>68,83</sup>, *GXS1*<sup>84,90,140</sup>, *GXF1*<sup>73,84,88,89</sup>, *Xylhp*<sup>81</sup> and others members from Major Facilitator Superfamily (MFS) were used as baits to select the transporter-encoding genes Cs186, Cs2608, Cs3894 and Cs4130. Members of MFS are classified on three transport groups: uniport, antiport and symporter<sup>75,79</sup>. From phylogenetic comparative analysis, figure 2, two of the elect *C. sojae* transporters, Cs4130 and Cs3894 shown proximity sequence to members from uniport transport, as *GXF1*, *HXT7* and *GAL2*, whereas Cs186 is related to *XUT3* from *S. stipitis*. However, Cs2608 was more related to symporter transporters, as *GXS1*. Therefore, these sequence similarities suggest that Cs186, Cs3894 and Cs4130 transporters are novel members of uniport transporters and Cs2608 represent a new xylose/glucose - H<sup>+</sup> symporter.

All selected *C. sojae* transporters were tested on strain EBY\_Xyl1, a modified version from EBY.VW4000<sup>125</sup>, expressing the XR/XDH/XKS xylose consumption pathway (table 1). This analysis shown that *C. sojae* transporters restore growth in all sugars tested, showing a broader substrate specificity (Figure 4). Regarding Cs186, its low xylose transport capacity may be explained not only by the distant relation to known xylose carriers, but also by the presence

of many identical copies on *C. sojae* genome (Cs186, Cs173, Cs183, Cs136) as shown in the polytomy in figure 2D. As transporters are dose-dependent, this low capacity could be compensated by these extra copies, which together supply xylose uptake for *C. sojae* cell. Cs2608 is a symporter and only few examples were functionally expressed on *S. cerevisiae* previously in literature. Despite being functional and restoring EBY\_Xyl1 growth, Cs2608 performed poorly on all tested sugars, especially on glucose and mannose (Figure 4).

The growth profile observed in xylose were different comparing *C. sojae* transporters and concentration-dependent (Figure 5A). While the traditional xylose transporter *GXF1* is inhibited at concentrations above 10 g/L, the strains expressing Cs3894 and Cs4130 show superior xylose uptake, not affected at concentrations up to 50 g/L (Figure 5A). Considering that xylose concentrations in 2G hydrolysates can reaches high values, Cs3894 and Cs4130 are profitable candidates for xylose uptake due to natural loss of inhibition. Interestingly, Cs4130 restores a better growth of recombinant strains at high xylose concentrations than at low concentrations presenting an opposite behavior to *GXF1*. Since Cs4130 enabled better growth compared to Cs3894, especially in higher concentrations, we decided to focus the analysis with only the first facilitator. The fermentation in liquid media confirmed the non-conventional pattern of growth of Cs4130 observed in spot-assay experiments, presenting a higher productivity in 50 against 10 g/L of xylose.

Simultaneous uptake of glucose and xylose are highly desirable in engineered 2G strains. The cells expressing a heterologous xylose assimilation pathway preferentially consume glucose and a diauxic growth is observed in hydrolysates. The glucose repression effect is caused mainly by competition for uptake through hexose transporters<sup>141</sup>. To investigate if the transport mechanism of Cs4130 alleviates the glucose inhibitory effect, we tested different growth conditions with mixtures of glucose and xylose. High-cell density fermentation with an equal concentration of both sugars, figure 6A, shown that glucose was consumed preferentially in the first 10 hours, indicating a higher preference for glucose over xylose. We disrupted the glycolysis in EBY\_Xyl1\_ *hvk*<sup>0</sup>, uncoupling the glucose uptake from its metabolization, similarly as the strategy described by Farwick, et al<sup>98</sup>. The mutant cell is unable to metabolize glucose through knockout of hexokinase genes (*HXK1*, *HXK2* and *GLK1*)<sup>98,132</sup>. EBY\_Xyl1\_ *hvk*<sup>0</sup> expressing

*GXF1* and Cs4130 showed strong glucose inhibition on a mixture of C5 and C6 (Figure 6B). Higher glucose concentration on mixtures of C5 and C6 decreases overall the growth of the strain. The strong glucose inhibition on *HXK* null strain were seen previously for *GAL2* native transporter which showed a high affinity for glucose over xylose.

Sugar transport in cells is a dynamic event where transmembrane proteins oscillate between the outward/occluded and inward/occluded states. Some residues make the path for ligand mainly by hydrogen bonds which are constantly redesigned in order to pull the sugar to the cytoplasmatic face<sup>142</sup>. By measuring the residue-ligand distance, some valuable information about the residues directly related to sugar transport as those responsible for protein dynamic properties such as atom fluctuation and conformational changes in side-chains can be obtained<sup>143</sup>.

Several residues related with xylose transport were highlighted in the work carried out by Sun et al. (2012)<sup>135</sup>. Even sharing the same sugar binding regions, the response of Cs4130 and *GXF1* to sugars is different, probably given the vicinity of residues in contact with sugars. The potential contribution of residues distant to xylose to transport in Cs4130 and *GXF1* was evaluated by Normal Mode Analysis (NMA). NMA assumes that modes exhibiting the lower frequencies describe the largest movements in a protein considering that alpha carbon motions describe the entire backbone motions<sup>144,145</sup>. The mutation R365A in Cs4130 changed the dynamics regime of the residues 50-60, 250 and 290 (Figure 9B) while the RMSF of *GXF1* remained quite stable in all 10 low-frequency modes evaluated for all protein variants (Figure 9D). The TM1 region harbors the residue F24, essential for xylose transport in *Xyle*<sup>135</sup>, and L50-F60 in Cs4130. Then, the variant R365A possible will result in a different phenotype in Cs4130. Interestingly, Young et al.<sup>114</sup> showed that the conserved motif G-G/F-XXX-G in TM1 is related to the preferential consumption of xylose over glucose in MFS family transporters. Changes in the variable region, represented by X, improve xylose uptake and attenuate glucose transport. Cs4130 presents the same conserve motif, however, the variable region shows a unique sequence: Y56, V57, and F58 (Figure 8). Amino acid changes in the variable region of the symporter *GXS1* isolated from *C. intermedia* resulted in improved xylose affinity<sup>90,114</sup>. Qureshi et al. (2020)<sup>146</sup> showed that polar interactions of residues

about 15Å from sugar are determinant for the transport of glucose and fructose in PfHT1 from *Plasmodium falciparum*. The authors concluded that the substrate promiscuity is not devoid to sugar-binding site, but substrate-gating dynamics.

The transport of sugars through membrane is not an obvious event in nature. Several bonds are made and broken in a restrictive and insoluble micro-environment. From the functional and in silico data, it is hypothesized that Cs4130 is a more promiscuous transporter than *GXF1* given it supports higher concentrations of xylose. In addition, mutations in Cs4130 will probably result in a diversity of phenotypes not observed in *GXF1*, for example, placing Cs4130 as a model to obtain transporters with improved properties regarding the assimilation of substrates other than glucose.

### **Authors' contributions**

JGRB performed the molecular biology, fermentation assays and wrote the manuscript, GB isolated the yeasts cells from diverse microbiomes, TLRC and LCO performed and analyzed the structure model data, MG and JJ performed the phylogenetic analysis, GAGP supervised the work and revised the manuscript, LVS designed/supervised the work, wrote and revised the manuscript. All authors read and approved the final manuscript.

### **Acknowledgements**

We are thankful to Centro Nacional de Pesquisas em Energia e Materiais (CNPEM), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Grant Number: 2017/08519-6, 2017/05078-9, 2018/00888-5), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq Universal: 430291/2018-3) and Serrapilheira Institute (grant number Serra-1708-16205). Image from yeast were made by Biorender.

### **Competing interests**

The authors declare that they have no competing interests.

**Availability of data and materials**

All data generated or analyzed during this study are included in this published article and its additional files.

**Funding**

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Grant Number: 2017/08519-6, 2017/05078-9, 2018/00888-5), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq Universal: 430291/2018-3) and Serrapilheira Institute (grant number Serra-1708-16205).

## Supplementary material

**Table S1** | Fermentation performance of isolated C5 yeast species in YPX 20 g/L. Yield (Yp/s) in grams of product (g<sub>p</sub>) per grams of xylose (g<sub>xyi</sub>) [g<sub>p</sub>/g<sub>xyi</sub><sup>-1</sup>]. Xylose consumption was measured at 54 hours of fermentation.

Yeast species	Xylose consumed (%)	Yp/s <sup>Xylitol</sup>	Yp/s <sup>EtOH</sup>	Yp/s <sup>Glycerol</sup>
<i>Blastobotrys (Arxula) adenivorans</i>	66.5	0.11	0.12	0.15
<i>Candida boidinii</i>	89.6	0.25	0.02	0.02
<i>Candida pseudointermedia</i>	99.75	0.6	0.01	0.01
<i>Candida sojae</i>	100	0.67	0	0
<i>Candida tropicalis</i>	100	0.6	0	0.02
<i>Cryptococcus laurentii</i>	36	0.28	0.02	0.02
<i>Kodameae ohmeri</i>	38.85	0.06	0	0
<i>Meyerozyma caribbica</i>	68	0.3	0	0.01
<i>Pichia sp.</i>	27.1	0.13	0.07	0.07
<i>Pseudozyma sp.</i>	36.1	0.09	0.02	0
<i>Trichosporon asahii</i>	36.75	0.28	0.02	0.01
<i>Trichosporon coremeiformi</i>	28.5	0.11	0.03	0
<i>Wickerhamomyces anomalus</i>	49.35	0.54	0.02	0

**Table S2** | Xylose transporter proteins used in phylogenetic analysis.

ID	Gene name	Accession
Q5A1D7	<i>XUT1</i>	<a href="https://www.uniprot.org/uniprot/Q5A1D7">https://www.uniprot.org/uniprot/Q5A1D7</a>
G4TIRP7	<i>PIIN_05137</i>	<a href="https://www.uniprot.org/uniprot/G4TIR7">https://www.uniprot.org/uniprot/G4TIR7</a>
Q2MDH1	<i>GXF1</i>	<a href="https://www.uniprot.org/uniprot/Q2MDH1">https://www.uniprot.org/uniprot/Q2MDH1</a>
P32467	<i>HXT4</i>	<a href="https://www.uniprot.org/uniprot/P32467">https://www.uniprot.org/uniprot/P32467</a>
P39003	<i>HXT6</i>	<a href="https://www.uniprot.org/uniprot/P39003">https://www.uniprot.org/uniprot/P39003</a>
P39004	<i>HXT7</i>	<a href="https://www.uniprot.org/uniprot/P39004">https://www.uniprot.org/uniprot/P39004</a>
P13181	<i>GAL2</i>	<a href="https://www.uniprot.org/uniprot/P13181">https://www.uniprot.org/uniprot/P13181</a>
O07563	<i>glcP</i>	<a href="https://www.uniprot.org/uniprot/O07563">https://www.uniprot.org/uniprot/O07563</a>
Q8J0V1	<i>mstA</i>	<a href="https://www.uniprot.org/uniprot/Q8J0V1">https://www.uniprot.org/uniprot/Q8J0V1</a>
Q2MEV7	<i>GXS1</i>	<a href="https://www.uniprot.org/uniprot/Q2MEV7">https://www.uniprot.org/uniprot/Q2MEV7</a>
A3GHU5	<i>XUT3</i>	<a href="https://www.uniprot.org/uniprot/A3GHU5">https://www.uniprot.org/uniprot/A3GHU5</a>
Q64L87	<i>Xylhp</i>	<a href="https://www.uniprot.org/uniprot/Q64L87">https://www.uniprot.org/uniprot/Q64L87</a>
Q5A1D7	<i>XUT1</i>	<a href="https://www.uniprot.org/uniprot/Q5A1D7">https://www.uniprot.org/uniprot/Q5A1D7</a>

**Table S3** | Primers used in this study. Homology to promoter (Pr) and terminator (Ter) *TDH1* are shown in bold.

Primer name	Sequence 5' to 3'	Description
Cs186_Forward	ATGCACGGTGGAGATTTAG	ORF amplification
Cs186_Reverse	TTATTCACTGATAACATCAGAATTTGTA	
Cs186_Pr_Forward	<b>CACACACAAAAACAGTACTTCACTAA</b> <b>ATTTACACACAAAAACAAA</b> ATGCACGGT GGAGATTTAG	pCS186 assemble
Cs186_Ter_Reverse	<b>AAATCATTATCCTCATCAAGATTGCTTT</b> <b>ATTTATTCACTGATAACATCAGAATTTGT</b> A	
Cs2608_Forward	ATGGGTTTTAAAGATAGTGAATTGG	ORF amplification
Cs2608_Reverse	TTAGACTGAATCTTCTTCAACATAA	
Cs2608_Pr_Forward	<b>CACACACAAAAACAGTACTTCACTAA</b> <b>ATTTACACACAAAAACAAA</b> ATGGGTTTTA AAGATAGTGAATTGG	pCS2608 assemble
Cs2608_Ter_Reverse	<b>AAATCATTATCCTCATCAAGATTGCTTT</b> <b>ATTTAGACTGAATCTTCTTCAACATAA</b>	
Cs3894_Forward	ATGAGTGCTACAGAACAAAATTCA	ORF amplification
Cs3894_Reverse	TTATGCAGATTCAGTTGCAGTT	
Cs3894_Pr_Forward	<b>CACACACAAAAACAGTACTTCACTAA</b> <b>ATTTACACACAAAAACAAA</b> ATGAGTGCT ACAGAACAAAATTCA	pCS3894 assemble
Cs3894_Ter_Reverse	<b>AAATCATTATCCTCATCAAGATTGCTTT</b> <b>ATTTATGCAGATTCAGTTGCAGTT</b>	
Cs4130_Forward	ATGTCTGTCTCCGACTC	ORF amplification
Cs4130_Reverse	CTAGACTTGTTCTTCATTTTGTTGT	
Cs4130_Pr_Forward	<b>CACACACAAAAACAGTACTTCACTAA</b> <b>ATTTACACACAAAAACAAA</b> ATGTCTGTCT CCGACTC	pCS4130 assemble
Cs4130_Ter_Reverse	<b>AAATCATTATCCTCATCAAGATTGCTTT</b> <b>ATCTAGACTTGTTCTTCATTTTGTTG</b>	
ITS1_Forward	TCCGTAGGTGAACCTGCGG	Molecular identification
ITS4_Reverse	TCCTCCGCTTATTGATATGC	



**Table S4** | Plasmids used in the study

Plasmids	Parental	Relevant features	Source
pRS426		Multi copy plasmid, <i>URA3</i>	160
pSH65		Cre recombinase, zeocin resistance	151
pRS304		<i>TRP1</i>	160
pSsXRXDH	pRS304	pPGK1-XYL1-tPGK1, loxP-URA3-loxP, pTDH1-XYL2-tTDH1	Santos, unpublished
pScXKS1		pADH1-XKS1-tADH1, loxP-URA3-loxP	Santos, unpublished
pCS186		<i>pTDH1-Cs186-tTDH1</i>	This study
pCS2608		<i>pTDH1-Cs2608-tTDH1</i>	This study
pCS3894	pRS426	<i>pTDH1-Cs3894-tTDH1</i>	This study
pCS4130		<i>pTDH1-Cs4130-tTDH1</i>	This study
pCS4133		<i>pTDH1-Cs4133-tTDH1</i>	This study
pCIGXF1		<i>pTDH1-CIGXF1-tTDH1</i>	This study

**Table S5** | Variants of Xyle (Sun et al., 2012) and equivalent residues in Cs4130, *GXF1*, STP10 (Paulsen et al., 2019) and GlcP<sub>SE</sub> (Iancu et al., 2013)

Xyle	Cs4130	GXF1	STP10	GlcP <sub>SE</sub>
F24, decreases xylose transport	F58	F58	F39	<b>Y19</b>
G83, abolishes xylose transport	G120	G120	G110	G71
R133, abolishes xylose transport	R153	R153	R142	R102
E153, abolishes xylose transport	E173	E173	E162	E122
R160, abolishes xylose transport	R180	R180	R169	R129
Q168, abolishes xylose transport	Q188	Q188	Q177	Q137
Q288, abolishes xylose transport	Q314	Q314	Q295	Q250
Q289, strongly decreases xylose transport	Q315	Q315	Q296	Q251
N294, abolishes xylose transport	N320	N320	N301	N256
Y298, abolishes xylose transport	Y324	Y324	<b>F305</b>	<b>F260</b>
N325, No effect in xylose transport	N349	N349	N332	N287
G340, abolishes xylose transport	G364	G364	G347	<b>D302</b>
R341, abolishes xylose transport	R365	R365	R348	R303
W392, abolishes xylose transport	<b>Y428</b>	<b>Y431</b>	W410	W357
E397, abolishes xylose transport	E433	E436	E415	E362
R404, strongly decreases xylose transport	<b>K440</b>	R443	R422	R369
Q415, strongly decreases xylose transport	<b>N451</b>	<b>N454</b>	<b>N433</b>	<b>L380</b>
W416, strongly decreases xylose transport	W452	W455	<b>M434</b>	<b>N381</b>

The amino acids were mutated to alanine (Sun et al., 2012). The only exception is 133R, which was mutated by C, H or L, but the phenotype was the same for all variants. Substitutions of amino acids among the transporters are highlighted in bold.

**Table S6** | Free energies of binding (kcal/mol) of Cs4130, *GXF1* mutants with xylose through docking analysis

	GRID SCORE (kcal/mol)	
	Cs4130	<i>GXF1</i>
<b>Wild-type</b>	-28.8	-25.8
<b>Mut 6</b>	I191 (-27.4/1.4)	I191 (-26.5/0.7)
<b>Mut 7</b>	I195 (-26.9/1.9)	I195 (-27.2/1.4)
<b>Mut 8</b>	F322 (-27.1/1.7)	R53 (-25.7/0.1)
<b>Mut 9</b>	F323 (-26.9/1.9)	F323 (-25.6/0.2)
<b>Mut 10</b>	I416 (-27.7/1.1)	I419 (-25.8/0)
<b>Mut 11</b>	F419 (-27.3/1.5)	F422 (-25.5/0.3)
<b>Mut 12</b>	Y428 (-28.5/0.3)	Y431 (-25.3/0.5)

In all mutants, the native residue was replaced by alanine.  
 The free energy of binding (kcal/mol) of mutants and the difference from the wild-type are between parentheses.  
 Mut, mutant.

## Capítulo 2

### **Design e construção de linhagens modificadas 2G de *Saccharomyces cerevisiae* para fermentação simultânea de xilose e glicose**

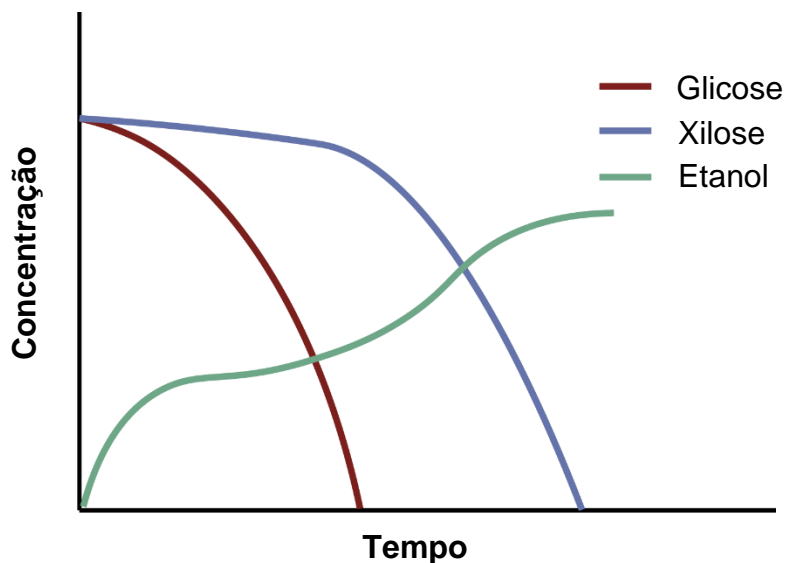
## Introdução

O desenvolvimento de tecnologias para o uso da biomassa lignocelulósica como matéria prima e produção de bioprodutos renováveis como biocombustíveis e bioquímicos tem avançado ao longo dos anos<sup>1-3</sup>. Entretanto para que todo o potencial dessa biomassa seja aproveitado, os açúcares provenientes dessa matéria prima devem ser utilizados de forma mais eficiente. O açúcar de seis carbonos (C6) glicose é o componente mais abundante presente na biomassa, sendo utilizado naturalmente pela plataforma microbiana escolhida na indústria de segunda geração (2G), a levedura *Saccharomyces cerevisiae*. Entretanto, o açúcar de cinco carbonos (C5) xilose, segundo açúcar mais presente na biomassa e componente da fração hemicelulósica, não é metabolizado naturalmente por linhagens selvagens de *S. cerevisiae*, necessitando de procedimentos de engenharia metabólica para o eficiente consumo dessa fonte de carbono<sup>4-7</sup>.

Após a construção de linhagens modificadas de *S. cerevisiae* capazes de fermentar xilose com alta eficiência<sup>8,9</sup>, uma outra limitação ainda existente ocorre nas fermentações com misturas dos açúcares C5 e C6 presentes nos hidrolisados 2G. Nessa mistura, as leveduras naturalmente irão consumir inicialmente a glicose e convertê-la em etanol. Dessa forma, a assimilação da xilose é iniciada somente quando a glicose está presente em baixas concentrações no meio, não ocorrendo o consumo simultâneo dos açúcares C5 e C6 (Figura 1)<sup>10-13</sup>. Esse mecanismo de preferência da glicose em detrimento a outro açúcar é comumente conhecido como repressão catabólica<sup>14-16</sup>. Essa preferência de consumo ocorre primeiramente devido a maior afinidade por glicose ao sistema endógeno de transporte presente na levedura em comparação aos outros açúcares<sup>17,18</sup>.

Moléculas complexas como os carboidratos necessitam de transportadores de membrana que facilitam a entrada desses açúcares para dentro da célula<sup>19</sup>. Esses transportadores de membrana são membros principalmente da “*Major Facilitator Superfamily*” (MFS), encontrados em todos os reinos biológicos<sup>20-23</sup>. Os transportadores da MFS são divididos em três classes: transporte uniporte, simporter e antiporte<sup>21</sup>. Além disso, a superfamília MFS pode ser dividida em 74 subfamílias<sup>24</sup>, onde estão descritos transportadores de xilose que foram

funcionalmente expressos em *S. cerevisiae*, como os endógenos *GAL2*<sup>17,25–28</sup>, *HXT7*<sup>17,25–27</sup> e os heterólogos *GXF1*<sup>25,29,30</sup> e *GXS1*<sup>31–33</sup>.



**Figura 1 | Fermentação em mistura de glicose e xilose de uma linhagem modificada de *S. cerevisiae*.** Uma linhagem de levedura consumidora de xilose, na presença da mistura dos açúcares, inicia o consumo do açúcar C5 somente quando a glicose é depletada do meio, levando a um extensivo tempo de fermentação (adaptado de Kim, *et al* 2012<sup>67</sup>).

Os transportadores endógenos de *S. cerevisiae* *GAL2* e *HXT7* já foram descritos como sendo os melhores transportadores capazes de carrear xilose na levedura<sup>25</sup>. Entretanto, eles apresentam uma baixa afinidade por esse açúcar, com  $K_m = 225.6 \pm 15.8$  mM e  $K_m = 200.3 \pm 13.2$  mM, respectivamente<sup>34</sup>. Por outro lado, esses mesmos transportadores demonstram uma grande afinidade para glicose,  $K_m = 1.9 \pm 0.2$  mM para *GAL2*<sup>34</sup> e  $K_m = 0.5 \pm 0.1$  mM para *HXT7*<sup>34</sup>. Assim, em um cenário industrial 2G, onde temos a presença tanto de glicose quanto de xilose em altas concentrações, os transportadores de *S. cerevisiae* irão apresentar uma maior preferência para glicose em relação a xilose devido a essa grande diferença de afinidade. Por esse motivo, grupos de pesquisa atualmente direcionam seus esforços para a obtenção de um transportador de xilose mais eficiente e com preferência exclusiva para esse açúcar C5<sup>26,34–40</sup>.

A presença de glicose intracelular pode ter um efeito inibidor no metabolismo de xilose<sup>41–43</sup>. Uma das estratégias encontradas para se estudar os efeitos da inibição por glicose é através de um mutante com deleções nos genes que codificam as hexoquinases na via Embden-Meyerhof (*GLK1*, *HXK1* e *HXK2*)<sup>37</sup>, interrompendo o primeiro passo da via glicolítica. A linhagem obtida

perde a capacidade de consumir a glicose, mesmo sendo capaz de transportar esse açúcar, acumulando intracelularmente. Poucos estudos foram apresentados e os mecanismos moleculares envolvidos na assimilação preferencial de xilose e na perda da inibição pela glicose nesses transportadores ainda não foi totalmente elucidado<sup>13,32,35,44</sup>. A competição de açúcares pelos transportadores de membrana acarreta no aumento do tempo de fermentação e consequentemente diminui a produtividade do processo de obtenção do etanol<sup>13</sup>.

Diversos transportadores heterólogos já foram analisados e expressos funcionalmente em linhagens de *S. cerevisiae*, visando otimizar a eficiência de assimilação de xilose pela levedura. Contudo, o melhor transportador heterólogo descrito na literatura *GXF1*<sup>25</sup>, isolado de *Candida intermedia*, apresenta uma afinidade por xilose de  $K_m = 48,6 \pm 6,5 \text{ mM}$ <sup>45</sup>, superior aos transportadores endógenos *GAL2* e *HXT7*. Curiosamente, apesar de apresentar essa maior afinidade para o substrato C5, linhagens de *S. cerevisiae* com o *GXF1* apresentaram um crescimento ligeiramente inferior comparado ao transportador endógeno de *GAL2*<sup>25</sup>. Mesmo sendo um dos melhores transportadores já descritos, ele permite a assimilação da xilose apenas em baixas concentrações, limitando seu uso em condições industriais de produção de etanol 2G<sup>25,29,46</sup>. Além disso, os transportadores *XUT1* e *SUT1* de *Scheffersomyces stipitis*, demonstraram maior transporte de xilose em condição onde os açúcares C5 e C6 estão presentes<sup>25,47</sup>. Apesar da seletividade no transporte de xilose, o crescimento das linhagens que expressam esses transportadores ainda não apresenta alta eficiência em meio contendo xilose como fonte única de carbono<sup>25,47</sup>.

Em um trabalho anterior envolvendo a temática de prospecção de novos transportadores a partir de leveduras naturalmente fermentadoras de xilose, identificamos um novo transportador com interessantes características para indústria 2G e que pode ser utilizado como base para futuras modificações (Capítulo 1)<sup>48</sup>. Com base nesse trabalho, o transportador Cs4130 da levedura *Candida sojae* demonstrou melhores performances em altas concentrações de xilose no meio (50 g/L), enquanto o tradicional transportador *GXF1* apresentou forte inibição. Assim como *GXF1*, o novo transportador heterólogo Cs4130 ainda apresenta uma forte preferência no transporte de glicose em relação a xilose,

levando a uma grande inibição na assimilação de xilose em misturas desses dois açúcares. Assim sendo, etapas de modificações visando o aumento da afinidade em xilose devem ocorrer para que Cs4130 seja considerado um potencial alvo biotecnológico no transporte C5.

Nesse contexto, o desenvolvimento de novas plataformas microbianas que possam ser utilizadas em metodologias de evolução adaptativa em meios com substratos C5 e C6 é um passo primordial para modificação de linhagens de *S. cerevisiae* capazes de consumir xilose e glicose. Dessa forma, a partir dessas linhagens evoluídas é possível identificar mutações relacionadas com o aumento da fermentação de xilose na presença de glicose. Essas mutações podem estar presentes tanto nos transportadores<sup>35,44,49</sup>, quanto em vias de regulação e sinalização associadas ao metabolismo de açúcares<sup>50</sup>. A compreensão do papel de cada uma dessas mutações e dos mecanismos e bases moleculares envolvidos na regulação, inibição ao substrato e co-consumo de açúcares C5 e C6, é essencial para desenvolver novos transportadores e leveduras que possam ser utilizados como alvos de engenharia metabólica em linhagens industriais.

## **Resultados e discussão**

### **Deleção das hexoquinases na linhagem LVY\_X5 e seleção dos mutantes gerados**

A deleção das hexoquinases em uma linhagem fermentadora de xilose é um evento necessário na estratégia de evolução adaptativa visando identificar bases moleculares responsáveis pela co-fermentação de C5 e C6. Para esse fim, utilizamos a metodologia de deleção gênica múltipla via CRISPR/Cas9 em um passo único. O plasmídeo que apresenta a maquinaria da enzima Cas9 denominado pCRCT<sup>182</sup>, também apresenta um “*cluster*” de gRNA’s acoplados com os DNA’s doadores, o que possibilita uma tentativa de deleção múltipla de genes por um único passo. A linhagem fermentadora de xilose LVY\_X5, transformada com os plasmídeos pDELHexo5 e pCRCT (controle negativo) foi submetida a uma etapa de seleção contínua pela metodologia descrita utilizando 2’ Deoxiglicose (2’ DOG) como fator de seleção. Esse composto quando fosforilado de 2’DOG para 2’DOG-6-Fosfato, pela ação das hexoquinases,

interrompe a via glicolítica que não poderá gerar energia para a célula. Assim, as linhagens que apresentam as cópias selvagens do gene das hexoquinases não irão crescer na presença de 2'DOG ao longo do tempo, mas os mutantes com deleções nesses genes terão a capacidade de crescer na presença em meio com a presença de 2'DOG e outras fontes de carbono, uma vez que esse composto inibitório não apresenta mais ação nesses mutantes gerados.

A estratégia desenhada utilizando o sistema CRISPR/Cas9 remove oito pb da sequência e gera um *stop codon* prematuro levando a perda de função da proteína. A tabela 1 abaixo especifica o tamanho inicial da proteína selvagem e após o evento de excisão.

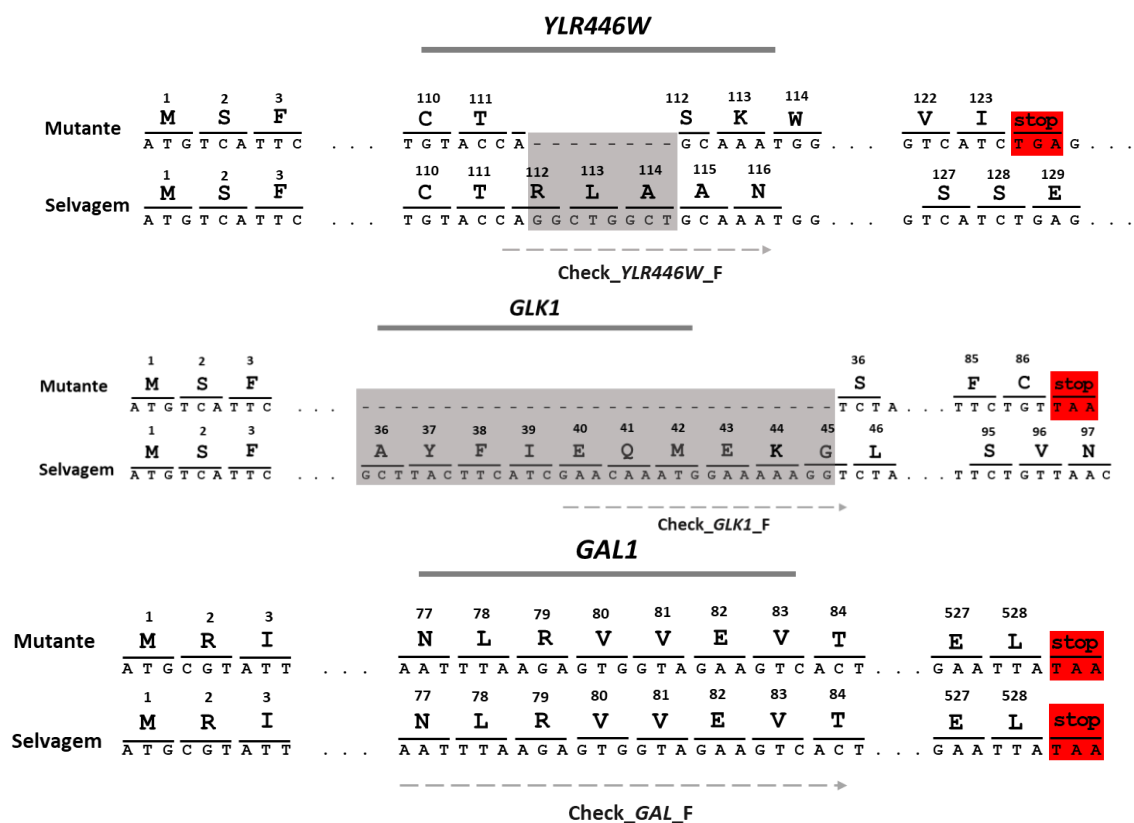
**Tabela 1** | Número de resíduos de aminoácidos remanescentes no peptídeo de cada gene após a indução com o plasmídeo pDELHexo5.

	Nº de aminoácidos proteína selvagem	Nº de aminoácidos proteína mutante
<b><i>HXK2</i></b>	486	219
<b><i>HXK1</i></b>	485	58
<b><i>GLK1</i></b>	500	93
<b><i>YLR446W</i></b>	433	123
<b><i>GAL1</i></b>	528	303

Após cultivo em 2' DOG, colônias isoladas foram selecionadas para confirmação das deleções. A confirmação foi realizada por PCR, onde um dos primers (*check\_F*) para amplificação de cada gene foi desenhado para se anelar exatamente na região onde pretende-se excisar oito pb do gene. Assim sendo, em linhagens onde ocorreu o evento de deleção, não ocorrerá o anelamento dos primers na região específica, impedindo a amplificação. Para confirmação do evento de CRISPR, foi realizado um sequenciamento na plataforma Sanger cobrindo a região alvo de cada gene (Figura 2).

O resultado do sequenciamento confirmou a deleção dos genes *GLK1* e *YLR446W*. Curiosamente, foi observado que 29 nucleotídeos foram deletados em *GLK1* ao invés dos oito pb esperados. Entretanto, isso não afetou a formação do "*stop codon*" prematuro levando a perda de função da proteína. Nesse caso, o tamanho do peptídeo remanescente em *GLK1* foi de 86 aminoácidos.



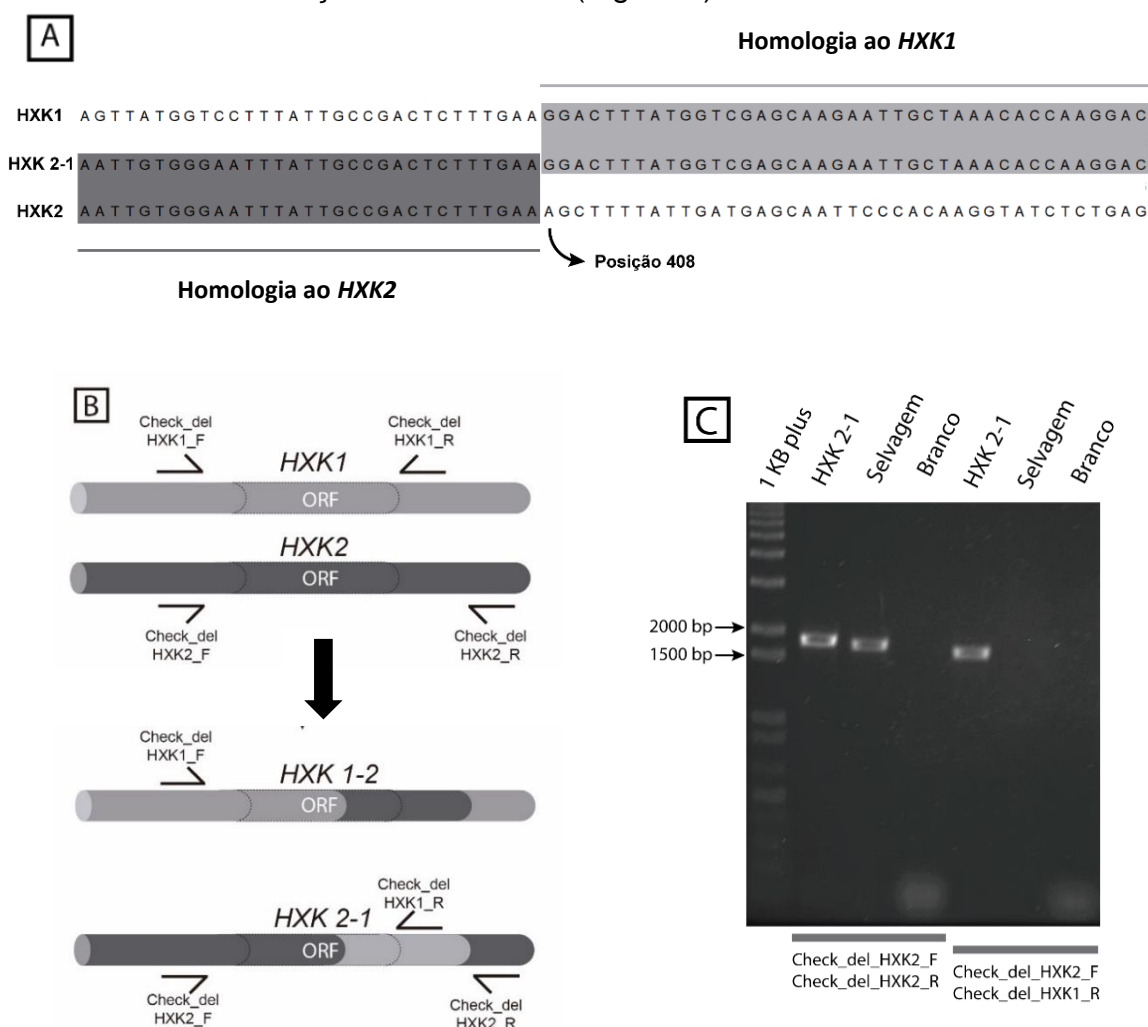


**Figura 2 | Sequenciamento YLR446W, GLK1 e GAL1.** Análise dos reads do sequenciamento Sanger confirmando as deleções presentes com relação ao gene selvagem, quadro cinza. As regiões de anelamento dos primers de confirmação estão destacadas na figura (seta pontilhada).

Por outro lado, o gene *GAL1* não foi deletado utilizando a metodologia descrita, uma vez que o seu gRNA/DNA doador estavam localizados na última posição do cassete de deleção, diminuindo a eficiência da deleção. Tal evento foi previamente descrito por Bao et al., 2015<sup>182</sup>. Quanto maior o número de posições dentro do cassete de deleção em pCRCT, menor será a probabilidade do evento ocorrer nos últimos alvos. Dessa forma, a falha na deleção em *GAL1* pode ser atribuída à sua posição dentro do cassete CRISPR do plasmídeo pDELHexo5.

Curiosamente, na análise do sequenciamento do gene *HXK2* foi observado a formação de um evento de recombinação entre os genes *HXK1* e *HXK2*, originando um gene quimérico funcional (Figura 3). Esse *contig* apresenta a primeira metade da sequência em correspondência ao *HXK2* fusionado a partir da posição 408 na ORF que codifica *HXK1*. Além do sequenciamento mostrar esse fenômeno de recombinação entre *HXK1* e *HXK2*, o evento foi confirmado por PCR (Figura 3-C). Os genes *HXK1* e *HXK2* são parálogos, ou seja, genes que originaram e evoluíram a partir de um evento de duplicação gênica,

compartilhando, dessa forma, alta similaridade de sequência<sup>183</sup>. Esse processo de recombinação entre duas regiões de dupla quebra (*DSB - double strand break*) no DNA já foi relatado por gerar recombinações entre cromossomos<sup>184-186</sup>. Assim, essas recombinações ocorreram entre as regiões das quebras geradas causando translocações cromossômicas indesejadas. Por combinação específica de primers, foi evidenciado que ocorreu apenas um evento de recombinação parcial a partir da posição 408 da região codificante de *HXK2*, sem o evento de translocação cromossomal (Figura 3).



**Figura 3 | Evento de formação do gene quimérico *HXK2-1*.** **A** – A análise do sequenciamento Sanger obtido da região onde ocorreu a troca entre *HXK2* e *HXK1*. Em destaque estão as regiões onde o gene quimérico *HXK2-1* possui similaridade com *HXK1* e *HXK2*. **B** – Modelo de recombinação parcial entre os parálogos *HXK1* e *HXK2*, originando os genes quiméricos e representação da região de anelamento dos primers de confirmação. **C** – Gel de eletroforese da confirmação do evento de recombinação com combinações específicas dos primers utilizados destacados na parte inferior. Marcador 1kb plus Thermo Scientific.



Esse problema pode ter ocorrido devido ao processo de seleção dos mutantes na metodologia proposta. Já foi descrito que um processo de evolução laboratorial em uma linhagem com deleções nos genes da hexoquinases (*hvk<sup>0</sup>*), mutações no gene *GAL1* revertiam o fenótipo levando a linhagem a consumir glicose<sup>95</sup>. Como esse gene se manteve inalterado devido a limitação da técnica descrita anteriormente, pode ser que ao longo do cultivo essa linhagem tenha adquirido alguma mutação levando a essa reversão do fenótipo. Além disso, culturas longas na presença de 2'DOG podem selecionar para mutações que levam as células a possuírem a capacidade de crescer na presença desse composto<sup>187,188</sup>. Dessa forma, uma nova metodologia e abordagem por deleção via CRISPR/Cas9 deve ser utilizada evitando o uso de 2'DOG.

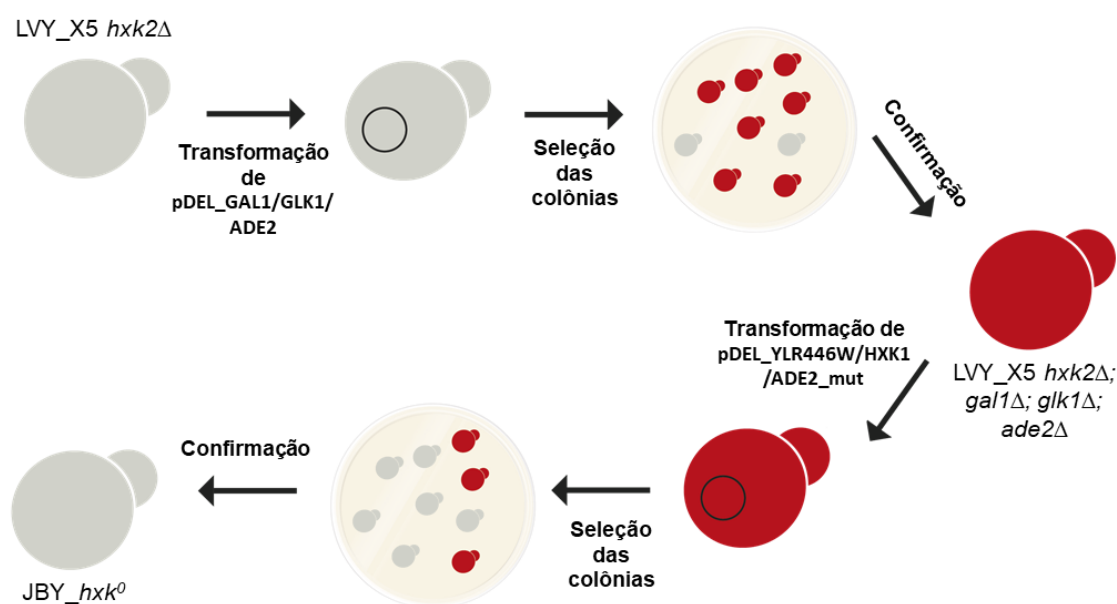
### **Gene repórter para seleção dos transformantes deletados para *HXK1*, *HXK2* e *GLK1* via CRISPR-CAS9**

O uso da metodologia CRISPR/Cas9 com o plasmídeo pDELHexo5 e seleção por meio do composto 2-deoxiglicose se mostrou desafiadora e gerou alguns resultados inesperados que prejudicaram o prosseguimento da construção da linhagem. Um dos fatores que é desfavorável no uso de CRISPR/Cas9 é o método de seleção dos transformantes. Comumente é utilizado um *screening* das colônias obtidas. Sendo alta a eficiência da técnica de CRISPR, os eventos de recombinação positivos são encontrados e reportados<sup>182,189,190</sup>. Entretanto, um *screening* de deleções múltiplas de genes que resultam em uma queda de fitness na linhagem recombinante é uma tarefa mais desafiadora<sup>191</sup>. Dessa forma, uma metodologia que tenha auxílio de uma seleção e contra-seleção possibilitará um *screening* mais rápido e eficiente das colônias positivas (Figura 5).

A partir de um novo conjunto de cassetes de deleção via CRISPR, poderemos deletar os genes das hexoquinases com uma metodologia de seleção visual. O plasmídeo pDel\_GAL1/GLK1/ADE2 foi desenhado contendo o gene repórter *ADE2*, componente da via de biossíntese da adenina, na última posição do cassete. Em meio com uma concentração menor de adenina do que o usual<sup>192</sup>, a deleção desse gene leva ao acúmulo de um composto intermediário deixando a levedura com uma coloração vermelha<sup>193</sup>. A deleção do último gene

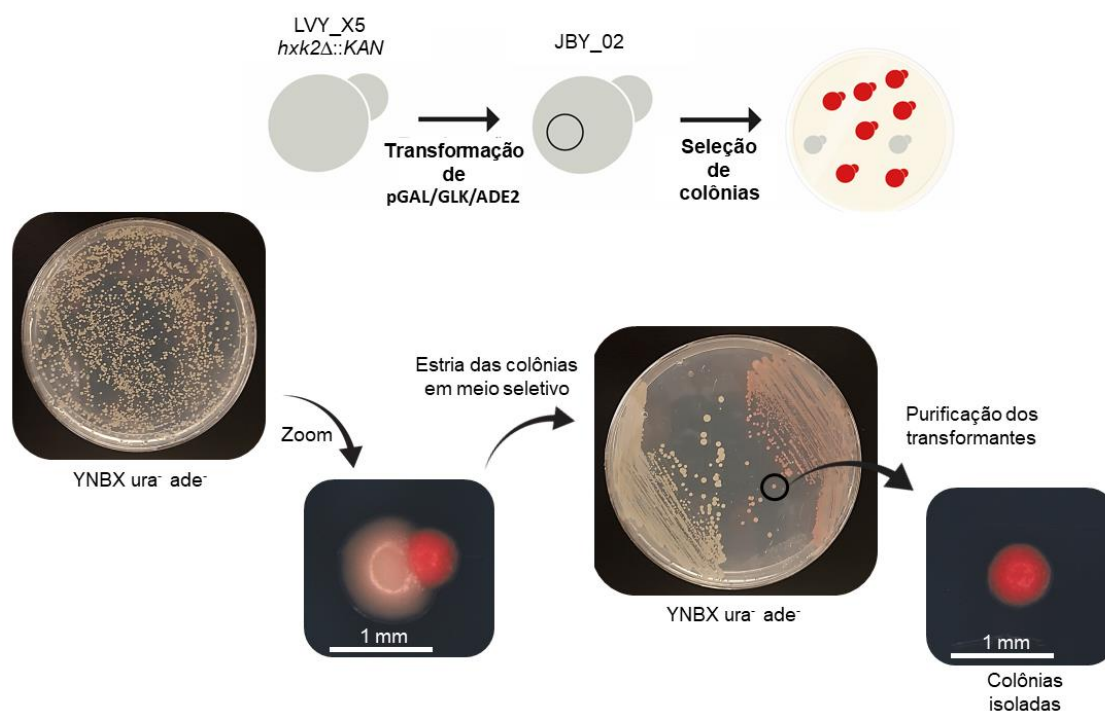
do cassete, o repórter *ADE2*, indica que provavelmente os genes das primeiras posições também foram deletados, *GAL1* e *GLK1*. A deleção de *ADE2* acontecerá pela inserção de uma sequência denominada de “*shuffer*”, que apresenta a mudança de 20 pb dentro da região codificante do gene, sendo que dentro dessa sequência “*shuffer*” contêm um stop códon prematuro. Com isso reduzimos nosso “*screening*” dos mutantes do primeiro plasmídeo identificando as leveduras com coloração vermelha.

O segundo plasmídeo, pDel\_YLR446W/HXK1/ADE2\_mut utiliza o mesmo princípio de seleção visual, restituindo o *ADE2* mutado no passo anterior. Nesse segundo passo, a condição selvagem do *ADE2* restitui a coloração usual da levedura, creme. O “*screening*” dos mutantes do segundo plasmídeo possibilitará a seleção das linhagens que tiveram a deleção dos genes que codificam as hexoquinases *YLR446W* e *HXK1* (Figura 5). O gene *HXK2* foi deletado previamente na linhagem inicial para evitar possíveis rearranjos cromossômicos indesejados, como observados anteriormente em experimentos com CRISPR/Cas9.



**Figura 5 | Nova estratégia de deleção das hexoquinases via CRISPR/Cas9 em linhagens fermentadoras de xilose.** Passo a passo da metodologia de geração dos mutantes utilizando a seleção e contra seleção por coloração das linhagens. O plasmídeo pDel\_GAL1/GLK1/ADE2 leva a deleção dos genes *GAL1*, *GLK1* e *ADE2*. O plasmídeo pDel\_YLR446W/HXK1/ADE2\_mut resulta na deleção dos genes *HXK1* e *YLR446W*, e a restituição do gene repórter *ADE2*. Após cada etapa de seleção será feito um sequenciamento dos genes deletados para confirmação.

Dessa forma, após a transformação da linhagem base LVY\_X5 *hxx2Δ* com o primeiro plasmídeo da nova metodologia, um “pool” de células foi induzido em meio líquido que apresenta a auxotrofia para seleção do plasmídeo (YNBX ura<sup>-</sup>), em um procedimento semelhante ao que foi feita com a primeira metodologia de CRISPR-Cas9. Dessa forma, após três passagem de crescimento as células induzidas foram plaqueadas em meio seletivo que apresenta metade da concentração recomendada de adenina (10 mg/L)<sup>192</sup> até que fosse possível verificar colônias de coloração vermelha. Após os crescimentos das colônias nas placas de “*screening*”, foi possível verificar a presença de 4 colônias vermelhas (Figura 6). Essas colônias foram isoladas, estriadas em meio seletivo e confirmadas molecularmente.



**Figura 6 | Primeira etapa da metodologia via ADE2.** Após a transformação do plasmídeo pdeL\_GAL1/GLK1/ADE2 na linhagem parental LVY\_X5 *hxx2Δ* e indução em meio seletivo, as colônias que se apresentaram vermelhas foram selecionadas para se obter uma colônia isolada.

Foram realizados PCR de confirmação com as 4 colônias candidatas utilizando regiões específicas de deleção acarretada pelo primeiro plasmídeo da estratégia, *GAL1*, *GLK1* e *ADE2*. Os DNA's doadores foram desenhados para que fosse retirada toda ORF de cada gene alvo que codificas hexoquinases, com exceção de *ADE2* onde apenas 20 pb são modificados. As figuras abaixo representam a validação dos mutantes.

A partir da análise dos resultados, é possível verificar que as ampliações de cada uma das regiões-alvo do plasmídeo pDEL\_GAL1/GLK1/ADE2 foi semelhante ao controle negativo, selvagem, mostrando que as deleções não ocorreram da forma como foram planejadas, dados não mostrados. Contudo, o fenótipo de coloração vermelha em condições de baixa concentração de adenina demonstra que ocorreu a mutação de perda de função do gene *ADE2* com sucesso. Essa deleção pode ter ocorrido através de um processo de “Non Homologous End Joining” (NHEJ), onde a maquinaria de reparo da célula concerta os DSB através de inserções aleatórias na região da quebra do DNA, podendo a levar a uma mudança no frame de leitura. Contudo, o DNA doador não é utilizado para esse reparo<sup>194,195</sup>.

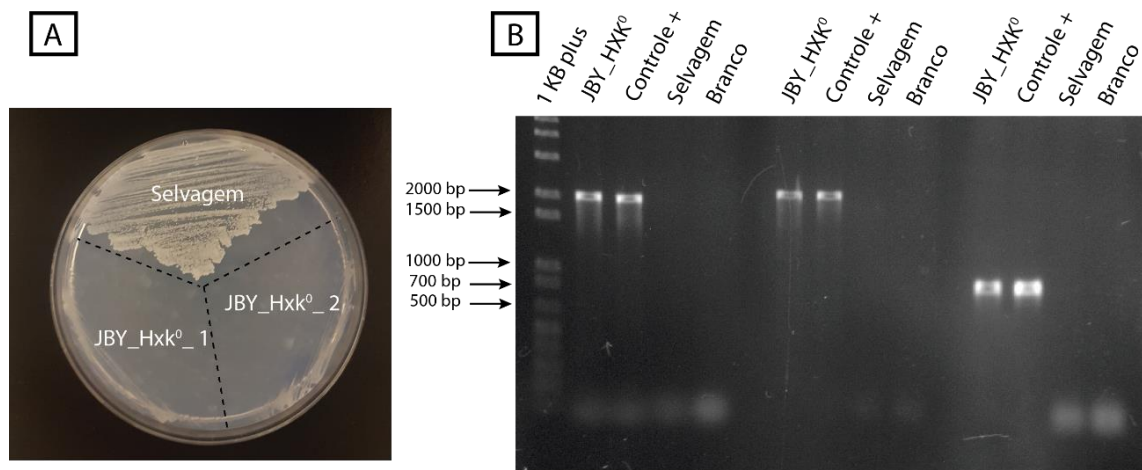
Esse processo de reparo explicaria a falha na confirmação do gene repórter *ADE2*, mesmo na presença do fenótipo. Além disso, se esse processo de reparo por NHEJ ocorreu nos outros genes alvo, *GAL1* e *GLK1*, a confirmação por primers específicos para os eventos de recombinação com os DNAs doadores não será possível, sendo necessário o sequenciamento dessas regiões.

### **Construção da linhagem JBY\_ *hxx*<sup>0</sup> por recombinação homóloga**

Devido aos problemas enfrentados na estratégia de construção da linhagem JBY\_ *hxx*<sup>0</sup> via CRISPR-Cas9, foi desenvolvida em paralelo a construção pela metodologia tradicional por recombinação homóloga. Genes que codificam marcadores de resistência a antibióticos foram utilizados para deletar as três hexoquinases *HXX1*, *GLK1* e *HXX2* na linhagem fermentadora de xilose LVY\_X5. A construção da linhagem JBY\_ *hxx*<sup>0</sup> através da metodologia tradicional por inserção de cassetes de deleção é mais desafiadora pelo fato de que após a deleção de cada gene, a taxa de crescimento da linhagem fica cada vez menor, dificultando a recuperação e seleção das colônias<sup>191</sup>.

Para a deleção dos genes *HXX1*, *GLK1* e *HXX2*, utilizou-se os marcadores que codificam resistência aos antibióticos higromicina, zeocina e geneticina, respectivamente. Cada etapa de transformação foi confirmada por crescimento em placa na presença do antibiótico e posteriormente via PCR. Ao final da construção da linhagem JBY\_ *hxx*<sup>0</sup> o fenótipo foi confirmado em placa

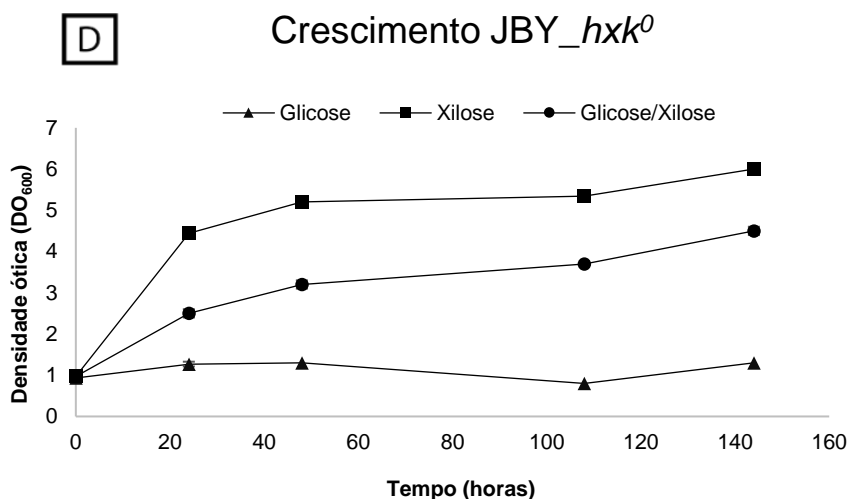
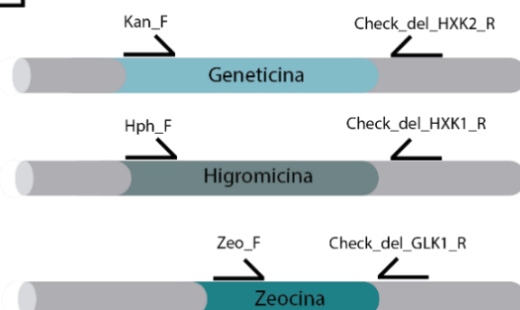
onde somente glicose se apresentava como fonte de carbono (Figura 7-A). Posteriormente foi realizado um teste em meio líquido em três condições: **(i)** xilose 40 g/L, **(ii)** glicose 40 g/L e **(iii)** xilose e glicose (20g/L de cada), para que possamos ver a inibição na glicose na presença de outra fonte de carbono (Figura 8-D). O teste em líquido podemos ver que a linhagem apresenta uma performance melhor em xilose em relação a mistura em glicose e xilose, mostrando que a presença da glicose inibe o crescimento da linhagem *hxx<sup>0</sup>*,



YNB enquanto em glicose observado.

**C**

nenhum crescimento foi observado.





**Figura 7 | Construção da linhagem JBY\_ *hxx*<sup>0</sup>.** **A** – crescimento de duas colônias candidatas em placa com meio mínimo com glicose como única fonte de carbono (YNB completo). A placa foi incubada por 4 dias em 30 °C. **B** – Confirmação molecular da presença dos cassetes de deleção no locus dos genes *HXK2*, *HXK1* e *GLK1*. Controle positivo refere-se a linhagem EBY\_Xyl1\_ *hxx*<sup>0</sup> que possui as mesmas deleções nos três genes das hexoquinases. Marcador 1kb plus Thermo Scientific. **C** – Representação esquemática da inserção dos cassetes de deleção e região de anelamento dos primers de verificação. **D** – Crescimento da linhagem JBY\_ *hxx*<sup>0</sup> em xilose (40 g/L), glicose (40 g/L) e mistura de xilose e glicose (20 g/L de cada).

Dessa forma, utilizando da metodologia tradicional de deleção gênica via recombinação homóloga, construímos uma linhagem recombinante capaz de fermentar xilose eficientemente e incapaz de metabolizar glicose. Sem a presença do primeiro passo da via glicolítica, essa nova plataforma microbiana reconhece as moléculas de glicose como um inibidor do metabolismo e ao transporte de outras fontes de carbono<sup>93,98</sup>. Portanto, é uma linhagem apropriada para experimentos de evolução adaptativa que evidenciem mutações relacionadas ao alto metabolismo e transporte de xilose sem que ocorra a inibição pela glicose.

### **Deleção das hexoquinases na linhagem *hxt*<sup>0</sup>, EBY\_Xyl1**

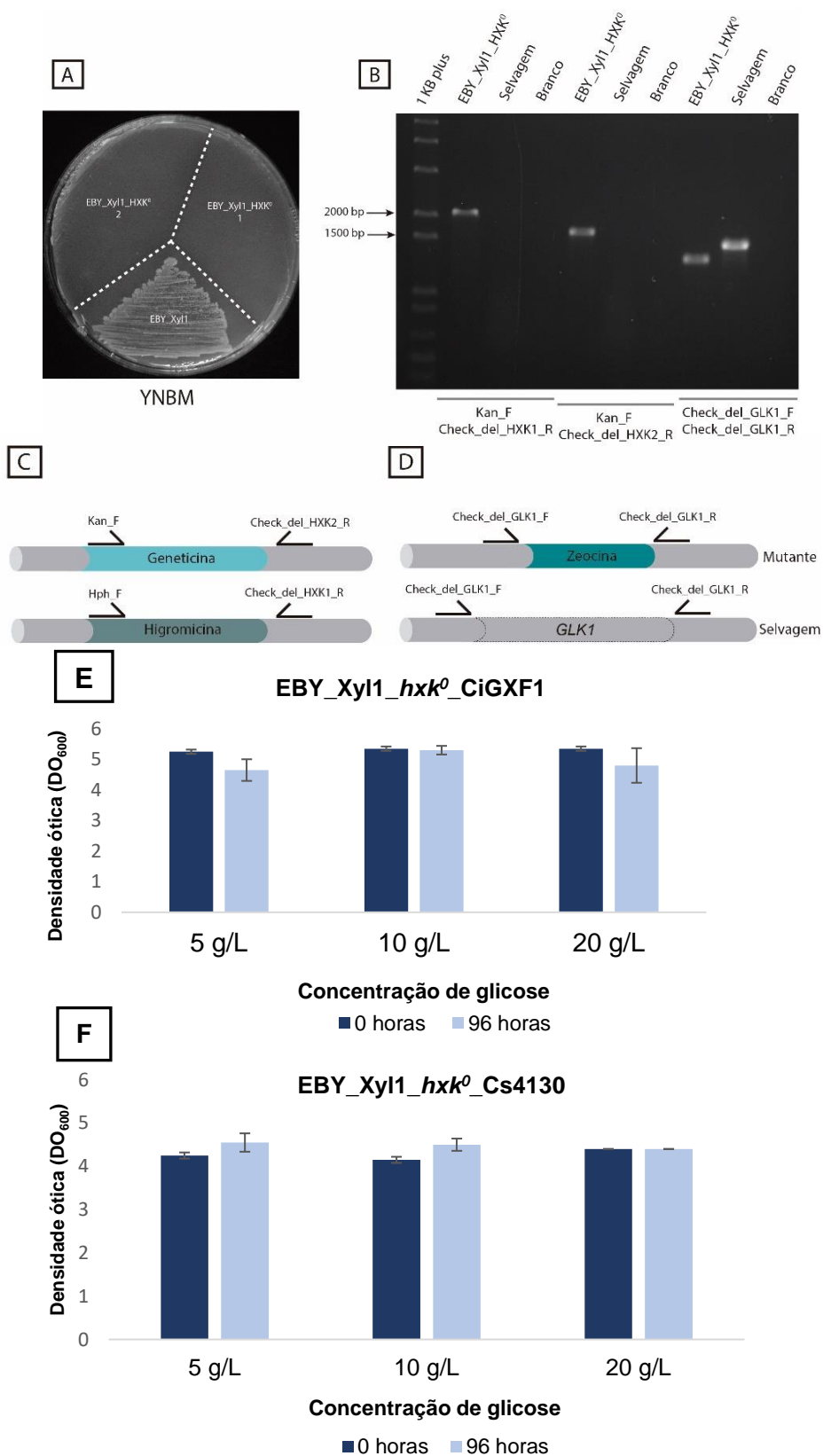
A grande maioria dos transportadores capazes de carrear xilose apresentam uma maior afinidade por glicose em relação aos outros açúcares, resultando em uma competição por substrato e consumo preferencial da glicose<sup>67,69–71</sup>. Sendo assim, o consumo da xilose em hidrolisados de segunda-geração inicia somente quando a glicose é depletada do meio<sup>67,132</sup>. Visando reduzir o tempo total da fermentação 2G pela assimilação simultânea de C5 e C6, torna-se imprescindível o desenvolvimento de transportadores capazes de carrear xilose com alta afinidade na presença de glicose, além da elucidação de mecanismos moleculares envolvidos na repressão catabólica em xilose. Para esse fim, é necessário conduzir o processo de evolução em uma cepa com a via glicolítica não funcional (*hxx*<sup>0</sup>) e sem transportadores endógenos de glicose e xilose (*hxt*<sup>0</sup>).

Farwick et al., 2014 descreveu uma estratégia similar de evolução, obtendo um transportador específico para xilose e sem inibição pela glicose<sup>82</sup>. A mudança de um único aminoácido na sequência do transportador resultou em uma diferença conformacional na região de ligação entre os açúcares com o transportador. Como existe diferença no tamanho da molécula de glicose em

relação a de xilose, um menor poro translocacional levou a glicose não conseguir mais se ligar ao transportador mutante, mantendo a ligação da xilose inalterada, permitindo a co-fermentação<sup>82</sup>.

A construção da nova plataforma microbiana *hxt<sup>0</sup>* e *hvk<sup>0</sup>*, ocorreu a partir da linhagem EBY\_Xyl1<sup>180</sup>, modificada da levedura EBY.VW4000 que possui a deleção nos transportadores da família *HXT* e *GAL2*<sup>125</sup> e com a expressão da via de conversão da xilose de *S. stipitis*. Como essa linhagem EBY.VW4000 não apresenta nenhum transportador capaz de carrear glicose, não é possível utilizar a metodologia de CRISPR-Cas9 desenvolvida para a linhagem LVY\_X5, pois o processo de seleção utiliza o análogo de glicose - 2' Deoxiglicose (2'DOG) - como componente essencial da técnica de seleção dos mutantes. Dessa maneira, a deleção dos genes que codificam hexoquinases nessa cepa foi conduzida por uma estratégia de recombinação homóloga tradicional utilizando marcadores de resistência ao antibiótico.

A deleção dos genes *HXK1*, *GLK1* e *HXK2* utilizou os mesmos marcadores de seleção utilizados na construção da linhagem JBY\_ *hvk<sup>0</sup>*, higromicina, zeocina e geneticina, respectivamente. Cada etapa de transformação foi confirmada por crescimento em placa na presença do antibiótico e posteriormente via PCR. Assim, ao final das etapas de inserção dos cassetes de deleções, confirmamos o fenótipo da linhagem final. Como a linhagem EBY\_Xyl1 originalmente já é incapaz de carrear glicose, a confirmação da ausência da via glicolítica não pode ser feita com *screening* funcional. As moléculas de maltose são carreadas por outro sistema de transporte em *S. cerevisiae* que está intacto na linhagem EBY\_Xyl1. Uma vez dentro da célula, o dissacarídeo maltose é hidrolisado em duas moléculas de glicose<sup>196</sup>. Dessa forma, duas colônias candidatas da linhagem, EBY\_Xyl1\_ *hvk<sup>0</sup>* foram testadas em meio sólido YNBM contendo maltose como fonte única de carbono. As duas colônias não foram capazes de crescer na presença desse açúcar, mostrando que a via glicolítica estava ausente (Figura 8-A). A deleção das hexoquinases foi confirmada molecularmente em uma dessas colônias, dando origem a linhagem final EBY\_Xyl1\_ *hvk<sup>0</sup>* (Figura 8-B).



**Figura 8 | Construção da linhagem EBY\_Xyl1\_hxk<sup>0</sup>.** **A** – Estria em meio completo YNBM com maltose como fonte única de carbono da linhagem parental EBY\_Xyl1 e dos mutantes com deleções em três hexoquinases. **B** – Confirmação da inserção do cassete de deleção das hexoquinases na linhagem EBY\_Xyl1\_hxk<sup>0</sup>. Marcador 1kb plus Thermo Scientific. **C** – Esquema de amplificação específica do cassete *hxk2Δ::KanMX4* (resistência a canamicina) e

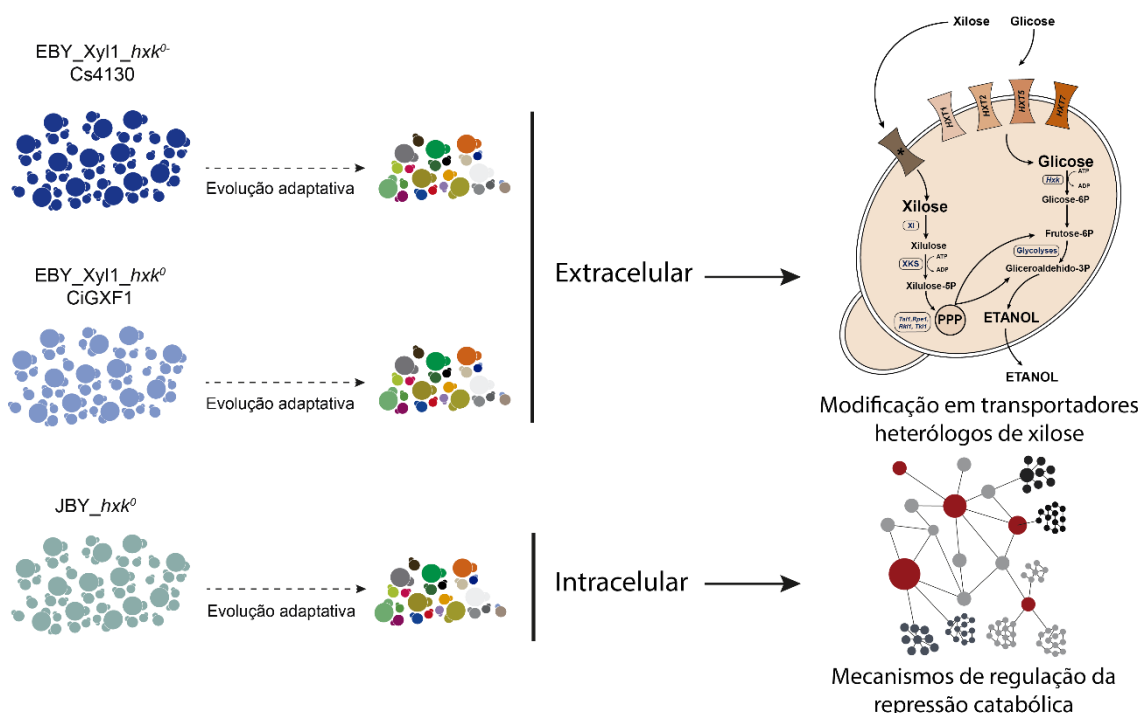
*hvk1Δ::hphMX4* (resistência a higromicina). **D** – Confirmação da deleção do gene *GLK1* por diferença de tamanho através da inserção do cassete *glk1Δ::bleR* (resistência a zeocina). Performance em meio líquido em glicose como única fonte de carbono com variada concentração do açúcar em 96 horas **E** (EBY\_Xyl1\_*hvk0*\_CiGXF1) e **F** (EBY\_Xyl1\_*hvk0*\_Cs4130).

Por fim, essa linhagem foi transformada com plasmídeos pCiGXF1 e pCs4130 que codificam os dois transportadores heterólogos de xilose que serão submetidos ao processo de evolução adaptativa. Um teste em meio líquido com essas linhagens transformantes onde glicose é a única fonte de carbono para os dois transportadores demonstrou que após 96 horas nenhum crescimento foi visto (Figura 9-E e 9-F). Dessa forma, podemos ver que mesmo na presença de dois transportadores de alta afinidade por glicose, o metabolismo desse açúcar está inativado na linhagem EBY\_Xyl1\_*hvk0*.

### **Evolução adaptativas das linhagens transformantes**

Finalizada as construções das novas plataformas EBY\_Xyl1\_*hvk0* e JBY\_*hvk0*, é possível iniciar o processo de evolução adaptativas em ambas as linhagens em meio com mistura de xilose e glicose, iniciando com concentrações baixas de glicose. Ao longo do processo de evolução espera-se que mutações aleatórias surjam. Entretanto, aquelas que são positivas na condição de seleção tendem a ser fixadas na população. Gradativamente, as concentrações de glicose no meio serão aumentadas conforme se observe o crescimento das linhagens<sup>93,98</sup>. Dessa maneira, a pressão seletiva irá se manter e selecionará as linhagens mais adaptadas a essa condição. O sequenciamento genômico dos mutantes de JBY\_*hvk0* e dos transportadores isolados de EBY\_Xyl1\_*hvk0*\_CiGXF1 e EBY\_Xyl1\_*hvk0*\_Cs4130, permitem a identificação de alvos que elucidem os mecanismos de repressão catabólica tanto em um nível intracelular, JBY\_*hvk0*, quanto extracelular, EBY\_Xyl1\_*hvk0*, (Figura 9).

A validação das mutações relacionadas a melhora no metabolismo de xilose na presença de glicose poderá ser utilizada para construções de novas plataformas microbianas mais adaptadas para os hidrolisados de 2G.



**Figura 9 | Evolução adaptativa em linhagens JBY\_hxk<sup>0</sup> e EBY\_Xyl1\_hxk<sup>0</sup>.** Representação esquemática dos processos de evolução nas linhagens EBY\_Xyl1\_hxk<sup>0</sup>\_Cs4130, EBY\_Xyl1\_hxk<sup>0</sup>\_CiGXF1 e JBY\_hxk<sup>0</sup> para elucidação de alvos moleculares relacionados a repressão catabólica por glicose. O processo de evolução em linhagens EBY\_Xyl1\_hxk<sup>0</sup> (extracelular) que expressam transportadores heterólogos (Cs4130 e GXF1) modificará estruturalmente os transportadores a fim de carrear xilose com maior eficiência na presença de glicose (repressão catabólica extracelular). O processo de evolução adaptativa na linhagem JBY\_hxk<sup>0</sup> (intracelular) visa identificar alterações em mecanismos de regulação gênica e sinalização envolvidos no processo de repressão catabólica em nível intracelular.

## Perspectivas Futuras

### Modificação racional dos transportadores de xilose

A metodologia proposta nesse projeto utiliza como base o procedimento de evolução adaptativa em linhagens de *S. cerevisiae* onde a pressão seletiva ocorre nos transportadores heterólogos expressos. Assim, as mutações que serão geradas ocorrerão de forma randômica e se fixarão na população conforme sejam selecionadas ao longo do tempo. As linhagens recombinantes de *S. cerevisiae* que expressam transportadores heterólogos apresentam diferenças de crescimento quando a concentração de açúcar no ambiente extracelular é alterada<sup>88,92,180</sup>. Assim, essa diferença apresentada no fenótipo das linhagens recombinantes pode ser devido ao processo de regulação dessas proteínas de membrana. Os transportadores de açúcares são regulados

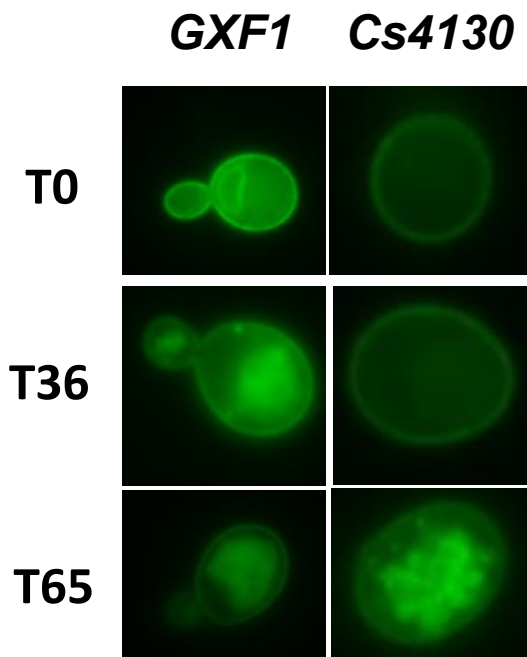
conforme as concentrações de açúcares presente no meio extracelular, através de dois níveis de regulação, transcricional e pós-traducional<sup>128,197</sup>.

O primeiro nível regula os níveis de expressão de cada transportador. Dependendo do ambiente extracelular, a sinalização celular ativa e inibe grupos específicos de transportadores de membrana.<sup>127</sup> Já o segundo nível regula a estabilidade dos transportadores na membrana plasmática, sendo que transportadores de alta afinidade, como *HXT7*, são degradados em um processo de ubiquitinação em condições de concentrações superiores de glicose<sup>129</sup>. Em contrapartida, para transportadores de baixa afinidade, como *HXT3*, a degradação ocorre em ambientes de baixa concentração de glicose. O transportador *GXF1* é considerado um carreador de alta afinidade para glicose,  $K_m = 2,0 \pm 0,6 \text{ mM}$ <sup>84</sup> e sua regulação é similar ao *HXT7*, sendo induzida em condições de baixa concentração de glicose ou por fontes não fermentescíveis<sup>130,198</sup> e degradado em condições de alta concentrações<sup>129</sup>.

Os transportadores heterólogos *GXF1* e Cs4130 apresentam diferentes padrões de transporte nas linhagens de *S. cerevisiae* recombinantes conforme a concentração de açúcar no ambiente extracelular (Capítulo 1)<sup>180</sup>. A expressão de *GXF1* demonstrou melhoras no crescimento em linhagens recombinantes em condições de baixa concentração de xilose, enquanto que em elevadas concentrações, o transportador é inibido<sup>88,180</sup>. O transportador Cs4130 apresenta melhora na performance nas linhagens recombinantes em condições de altas concentrações de xilose, enquanto em baixas concentrações desse açúcar o transporte é inibido. Ainda não é elucidado como esse mecanismo de regulação dos transportadores de membrana, tanto endógenos quanto heterólogos, ocorrem na presença de xilose e como esse açúcar atua em toda cascata de sinalização intracelular.

Um teste com os transportadores Cs4130 e *GXF1* acoplados a proteína fluorescente mUkG1<sup>161</sup>, mostrou que Cs4130 apresenta uma maior intensidade na membrana plasmática ao longo do crescimento em 50 g/L de xilose (Figura 10). O facilitador *GXF1* perde essa intensidade na periferia da célula e se observa um aumento da fluorescência na região citoplasmática, o que pode ser indício de um processo de degradação desse transportador. A confirmação dessa diferença na permanência dos transportadores na membrana plasmática

explicaria a diferença de performance das linhagens em altas concentrações de xilose.



**Figura 10 | Fluorescência dos transportadores GXF1 e Cs4130.** Análise da fluorescência ao longo do crescimento em 50 g/L de xilose. GXF1 apresenta aumento da fluorescência no ambiente citoplasmático antes do que o observado para Cs4130.

Os transportadores GXF1 e Cs4130 apresentam possíveis sítios de ubiquitinação que podem sinalizar o processo de degradação por meio de endocitose de proteínas de membrana (Figura 11). Futuras modificações nesses sítios de ubiquitinação podem limitar esse processo de degradação e consequentemente aumentar a permanência desses transportadores na membrana ao longo de toda a fermentação. Por fim, por meio de mutações pontuais e racionais nos transportadores de xilose poderemos melhorar as performances das linhagens recombinantes em altas concentrações de xilose.

### **GxF1 – *Candida intermedia***

```
MSQDSHSSGAATPVNGSILEKEKEDSPVLQVDAPQKGFKDYIVISIFCFMVAFGGFVFGFDTGTISGF
VNMSDFKDRFGQHHADGTPYLSDVVRVGLMISIFNVGCAVGGIFLCKVADVWGRRIGLMFMAVYVVGII
IIQISSSTKWYQFFIGRLIAGLAVGTVSVVSPLEFISEVSPKQIRGTLVCCFQLCITLGLIFLGYCTTYG
TKTYTDSRQWRIPGLGLCFAWAILLVVGMLNMPESPRYLVEKHRIDEAKRSIARSNKIPEEDPFVYTEV
QLIQAGIEREALAGQASWKEKITGKPKIFRRVIMGIMLQSLQQLTGDNFFYYGTTIFQAVGLKDSFQ
TSIILGIVNFASTFVGIYVIERLGRRLCLLTGSAAMFICFIIYSLIGTQHLYKQGYSNETSNTYKASG
NAMIFITCLYIFFFASTWAGGVYCIISESYPLRIRSKAMSIAAANLWLGFLISFFTPFITSAIHFFYY
GFVFTGCLAFSFFYVYFFVYETKGLSLEEVDEMYASGVLPLKSASWVPPNLEHMAHSAGYAGADKATD
EQV
```

Vermelho → Alta probabilidade  
Azul → Média probabilidade

## Cs4130 – *Candida sojae*

MSVSDSTDANPNLEDK<sup>Y</sup>ENNADPNIDAAAVDLNANQPWHAYIGIITMCVLIAFGGYVFGFDTGTIPGF  
 VNMTDFQRRFGSRDDGTYYSNARTGMMIGLNFNAGCAVGALLLSKLADLYGRRFAISVSMLIYISGV  
 IVQIAAQHAWYQVMIGRIITGVAVGTLSVVCPLFISEVSPKHLRGTLVVCFQLMITLGIIFLGCTTYG  
 TKVYSDSRQWRVPLGLCFAWALFLIGGMTRMPESPRYLLSKDKIDDARIALAKVNKASPEDPALLAEL  
 QLLQAGVERERLAGSATWMTLIKGPRIFERVLVGAMLQSLQQLTGDNYFFYYATMVFKAIGLDDSF  
 TAIIGVINFASTFVGIFAIEKLGRRCLMLGSI GMSACFLMYSLIGSQHLYYHGSSGETRHRDGRAM  
 IFVTSIYIFCFASTWAGGVYSVISELYPLKIKSKAMGVANACNWLWGFLISFFTSFITDAIHFYGF  
 FMGCLFFSIFVYFMVYETKGLDLGAVEELYNAGIPAWKSAGWVPPTAQEMATTSGYARDA<sup>K</sup>PEEQQN  
 EEQV

Vermelho → Alta probabilidade  
 Azul → Média probabilidade

**Figura 11 | Análise dos prováveis sítios de ubiquitinação nos transportador Cs4130 e GXF1.** Assinalado em vermelho e azul são representados os aminoácidos de alta e média probabilidade, respectivamente, para sítios de ubiquitinação. (<http://www.ubpred.org>).

## Conclusões

Para melhor utilização do potencial biotecnológico da biomassa lignocelulósica devemos utilizar a totalidade dos principais açúcares presente nessa biomassa, glicose e xilose. As linhagens industriais de *S. cerevisiae* desenvolvidas para a indústria 2G devem ter a capacidade de metabolizar esses dois açúcares de forma simultânea, levando ao aumento da produtividade dos produtos desejados, como o etanol. Dessa forma, entender e elucidar os mecanismos de repressão catabólica e os efeitos da glicose na presença de outras fontes de carbono nos níveis moleculares é essencial para construção de linhagens industriais 2G.

O primeiro passo para entender a repressão catabólica é obter linhagens plataforma capazes de identificar genes alvos que desvendem os mecanismos envolvidos. A linhagem construída JBY\_*h<sub>xk</sub><sup>0</sup>* mostrou uma performance pior em mistura com glicose e xilose do que somente em xilose. Isso evidencia que a glicose apresenta um forte papel inibitório no metabolismo interno da levedura mesmo na presença de outras fontes de carbono. Evoluir a linhagem JBY\_*h<sub>xk</sub><sup>0</sup>* em misturas de glicose e xilose e estudar as mutações presentes nas cepas evoluídas irão auxiliar a compreender o papel inibitório global que a glicose exerce na presença de outras fontes de carbono.

Adicionalmente, considerando uma linhagem *h<sub>xt</sub><sup>0</sup>* (sem os transportadores endógenos presentes no genoma) e *h<sub>xk</sub><sup>0</sup>*, temos uma segunda ferramenta capaz de ajudar na compreensão do processo de competição na



assimilação entre glicose e xilose pelos transportadores escolhidos. A construção da linhagem *EBY\_Xyl1\_hxk<sup>0</sup>* é essencial para a obtenção e “screening” de transportadores mutantes gerados em linhagens evoluídas capazes de consumir simultaneamente substratos C5 e C6.

Com a construção das duas linhagens, *JBY\_hxk<sup>0</sup>* e *EBY\_Xyl1\_hxk<sup>0</sup>*, poderemos identificar novos mecanismos de repressão catabólica e inibição por glicose em dois níveis moleculares, intracelular e extracelular. Os alvos a serem identificados possuem grande potencial para serem utilizados na construção de linhagens otimizadas ao ambiente industrial 2G.

## Metodologia

### Meios e linhagens utilizadas

Os meios utilizados foram preparados conforme descrito previamente em Ausubel et al., 2003<sup>147</sup>. Foi utilizado o meio LB (1% Triptona, 0,5% extrato de levedura e 1% NaCl) para o cultivo de *E. coli* DH5 $\alpha$ . A levedura *S. cerevisiae* foi cultivada em meio YPD (1% extrato de levedura, 2% Peptona e 2% Glicose) ou YPX (1% extrato de levedura, 2% Peptona e 2% Xilose). As linhagens *S. cerevisiae* *LVY\_X5* e *EBY\_Xyl1*<sup>180</sup> foram utilizadas como base para o estudo, tabela S1. *LVY\_X5* apresenta superexpressão do gene *XKS1* e dos quatro genes da PPP, *TAL1*, *TKL1*, *RPE1* e *RKI1*, bem como a deleção dos genes *GRE3* e *URA3*. *EBY\_Xyl1* apresenta deleção de 21 transportadores e é derivada da linhagem *EBY.VW4000*<sup>125</sup>.

- **LVY\_X5** (*MAT $\alpha$* ; *CEN5::pTDH1-xyIA-tTDH1*; *gre3 $\Delta$* ; *CEN2::pADH1-XKS1-tADH1*; *CEN8::pADH1-XKS1-tADH1*; *CEN12::pTDH1-TAL1-tTDH1-pPGK1-RKI1-tPGK1*; *CEN13::pTDH1-TKL1-tTDH1-pPGK1-RPE1-tPGK1*; *isu1 $\Delta$*  e *XylIA*).
- **EBY\_Xyl1** (*EBY.VW4000* - *hx15 $\Delta$* , *hxt16 $\Delta$* , *hxt13 $\Delta$* , *hxt14 $\Delta$* , *hxt12 $\Delta$* , *hxt9 $\Delta$* , *hxt11 $\Delta$* , *hxt10 $\Delta$* , *hxt8 $\Delta$* , *hxt4-1-5 $\Delta$* , *hxt2 $\Delta$* , *hxt3-6-7 $\Delta$* , *gal2 $\Delta$* , *stl1 $\Delta$* , *agt1 $\Delta$* , *mph2 $\Delta$*  e *mph3 $\Delta$* ; *CEN5::pPGK1-XYL1-tPGK1-pTDH1-XYL2-tTDH1*; *CEN2::pADH1-XKS1-tADH1*)

Para os experimentos de transformação, as linhagens foram cultivadas em YNB (0,67% Difco yeast nitrogen base, 1% drop out completo para leveduras, 0,5% sulfato de amônio, 2% glicose) e YNBX (0,67% Difco yeast nitrogen base, 1% drop out completo para leveduras, 0,5% sulfato de amônio, 2% xilose). Os

meios foram autoclavados a 121 °C durante 20 minutos, com exceção do YNB e YNBX, que foram filtrados em membrana 0,2 µm. O meio YNB pode ter a uracila removida da composição do “*drop-out*” para a seleção dos transformantes.

### **Extração de DNA plasmidial de bactéria**

O protocolo de extração de DNA plasmidial seguiu a metodologia de lise alcalina<sup>147</sup>. A linhagem de *E. coli* transformada foi inoculada em meio LB overnight e incubada a 37°C sob a agitação de 200 rpm. Após a extração, o DNA foi quantificado em nanodrop para determinar a eficiência da extração. Esse procedimento foi utilizado para extrair DNA plasmidial das bactérias contendo os plasmídeos pDelHexo5, pCl\_GXF1, pCs186, pCs2608, pCs3894 e pCs4130 que foram utilizados nas transformações de leveduras. Plasmídeos utilizados ao longo desse capítulo se encontram na tabela S2

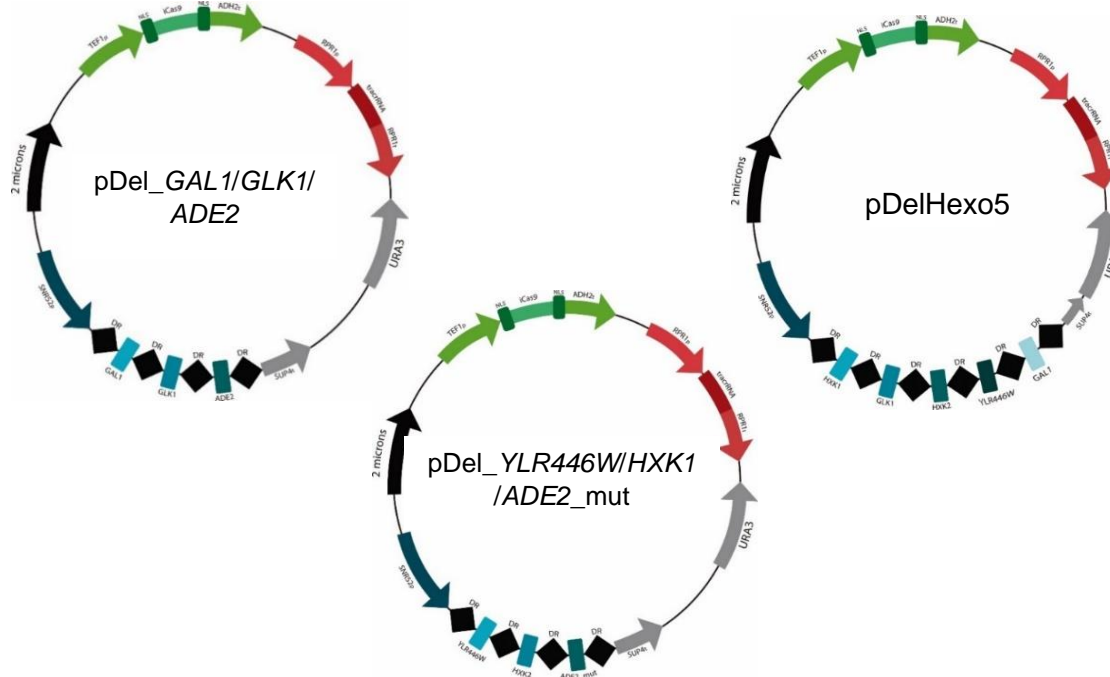
### **Transformação de *S. cerevisiae*.**

As transformações de *S. cerevisiae* foram realizadas utilizando o método de Acetato de Lítio (LiAc)<sup>199</sup>. Esse protocolo é composto por três etapas: i. preparação de células de levedura competentes, ii. transformação com o plasmídeo/DNA de interesse e iii. seleção dos transformantes. A indução da competência nas leveduras utiliza uma solução de 1x TE|LiAc. A transformação da cepa utilizada ocorreu com a adição do DNA a ser transformado nas células de levedura e soluções de 40% PEG/1x LiAc (80% PEG 50%; 10% TE buffer, 10% 10x acetato de lítio 1M), DMSO e DNA de espermatozoides de salmão como carreador do DNA exógeno. A seleção das colônias transformantes foi feita em YPX suplementado com o antibiótico necessário ou em meio YNB sintético sem uracila, de acordo com a seleção e fonte de carbono necessária

### **Plasmídeos de deleção das hexoquinases via CRISPR-Cas**

Foram sintetizados plasmídeos para a deleção das hexoquinases com base no plasmídeo já disponível pCRCT, descrito por Bao et al., 2015<sup>182</sup>. O plasmídeo denominado pDELHexo5 (Figura 12) tem uma construção que irá gerar deleções de oito pb nos genes que codificam as hexoquinases *GLK1*, *HXK1*, *HXK2*, *YLW446W* (putativa hexoquinase)<sup>200</sup> e *GAL1* (galactose quinase).

Esses dois últimos já foram descritos como genes que ao longo do processo de evolução podem obter mutações que os levem a adquirirem funções de hexoquinases<sup>98,132,201</sup>. O plasmídeo pDelGAL1/GLK1/ADE2 visa gerar a deleção pela retirada da região codificante dos genes *GAL1* e *GLK1*, enquanto a deleção o gene repórter *ADE2* ocorre pela a inserção de um stop códon prematuro. Nessa estratégia, os genes *YLR446W* e *HXK1* são deletados da mesma maneira pelo plasmídeo pDel\_YLR446W/HXK2/ADE\_mut. Entretanto a deleção do gene *ADE2* gerado pela ação do plasmídeo pDel\_GAL1/GLK1/ADE2 é restituída pela remoção desse stop códon prematuro. Os cassetes de deleção via Crispr-Cas9, DNA's doadores e gRNA's utilizados se encontram nas tabelas tabela S4, S5 e S6, respectivamente.



**Figura 12 | Plasmídeo de deleção das hexoquinases.** A construção dos plasmídeos pDelHexo5, pDel\_GAL1/GLK1/ADE2 e pDel\_YLR446W/HXK1/ADE2\_mut utilizou como base o plasmídeo pCRCT. O plasmídeo apresenta uma região codificadora da proteína *iCas9*, com marcas de sinalização para transporte intranuclear (NLS) e uma região de tracrRNA, formando o complexo funcional base da metodologia de deleção. Os RNA's guias estão juntamente com os DNA's disruptivos de cada gene alvo, representados pelo retângulo com o nome do gene, sob a ação do promotor SNR52 e intercalados por uma sequência "direct repeat" (DR). (Adaptado de Bao *et al.* 2015)<sup>182</sup>.

### Deleção das hexoquinases em LVY\_X5 e EBY\_Xyl1

Para realização da metodologia CRISPR-Cas em LVY\_X5, usamos uma estratégia de construção contendo apenas um plasmídeo que tem as regiões de reconhecimento aos genes de interesse *GLK1*, *HXK1*, *HXK2*, *YLR446W* e *GAL1*

e as regiões que compõem o complexo funcional da maquinaria do CRISPR<sup>182</sup>. Essa metodologia utiliza uma combinação em conjunto de gRNA's (*guide RNA*) e DNA doador de cada gene alvo em um único plasmídeo. Sendo possível através dessa construção particular uma deleção multiplex de genes em *S. cerevisiae*. Os genes serão deletados por meio do uso do CRISPR que está presente no plasmídeo, sendo dessa forma uma metodologia de passo único.

A partir de uma colônia isolada, foi feito um pré-inóculo em YNBX ura<sup>-</sup> (50 g/L) e posteriormente as linhagens foram transferidas para um meio contendo YNBX ura<sup>-</sup> (10 g/L) + 2' Deoxiglucose (2 g/L). A presença de 2' Deoxiglucose (2'DOG) foi descrito previamente como um composto que não possibilita o crescimento de linhagens que possuam a capacidade de fosforilar glicose<sup>132</sup>. Dessa forma, as linhagens mutantes para as hexoquinases selecionadas terão uma vantagem seletiva em meio com a presença de 2'DOG. Ao longo da seleção contínua, alíquotas de 200 uL foram transferidas para um novo meio de cultivo, visando selecionar linhagens com deleções das hexoquinases e capacidade de crescer em xilose na presença de 2'DOG.

As deleções das hexoquinases em *EBY\_Xyl1* aconteceu através da metodologia tradicional por inserções de cassete de resistência a antibióticos, por recombinação homóloga, dentro da região codificante do gene a ser deletado<sup>202</sup>. Dessa forma, a deleção de *HXK2*, *HXK1* e *GLK1* prosseguiu com a amplificação dos genes de resistência presente plasmídeos pFA6KanMX4 (geneticina)<sup>149</sup>, pAG32 (higromicina)<sup>150</sup> e pSH65 (zeocina)<sup>151</sup>, respectivamente.

### **Microscopia e linhagens *S. cerevisiae* fluorescentes**

A sequência otimizada da proteína fluorescente *mUKG1*<sup>161</sup> foi sintetizada pela Genscript e fusionada ao plasmídeo pCs4130 e pCiGXF1<sup>180</sup>. As linhagens *EBY\_Xyl1* com a expressão das construções Cs4130-mUkG1 e GXF1-mUkG1 foram crescidas em meio YNB sem uracila, com a respectiva fonte de carbono. As linhagens foram analisadas no microscópio de fluorescência Carl Zeiss (Jena, Germany) utilizando objetiva de 100x por óleo de imersão (EC Plan-Neofluar, NA1.3). Imagens de contraste e fluorescência foram obtidas pela câmera AxioCam (Carl Zeiss) e processadas no software AxioVision software versão 3.1.

## Material suplementar

**Tabela suplementar S1.** Linhagens usadas nesse estudo

<b>Linhagem</b>	<b>Genótipo</b>	<b>Parental</b>	<b>Referência</b>
EBY.VW4000	$\Delta$ hxt1-17 gal2 $\Delta$ ::loxP st1 $\Delta$ ::loxP agt1 $\Delta$ ::loxP mph2 $\Delta$ ::loxP mph3 $\Delta$ ::loxP	CEN.PK2-1C	125
EBY_XYL1	CEN5::pPGK1-XYL1-tPGK1; pTDH1-XYL2-tTDH, CEN2::pADH1-XKS1-tADH1	EBY.VW4000	180
EBY_Xyl1_hxk <sup>0</sup>	Kan $\Delta$ HXK2, Zeo $\Delta$ GLK1, Hph $\Delta$ HXK1	EBY_XYL1	Este estudo
LVY_X5	pOXylATy1 + seleção em evolução adaptativa	PE-2	Dos santos, comunicação pessoal
JBY_hxk <sup>0</sup>	Kan $\Delta$ HXK2, Zeo $\Delta$ GLK1, Hph $\Delta$ HXK1	LVY_X5	Este estudo
CEN.PK2-1C	MAT $\alpha$ ura3-52 trp1-1 his3 $\Delta$		
LVYA1	MAT $\alpha$	PE-2	63
PE-2	Linhagem industrial de produção de etanol PE-2, MAT $\alpha$ / $\alpha$		CBMAI 0959

Tabela suplementar S2. Plasmídeos utilizados nesse estudo

Plasmídeo	Plasmídeo parental	Característica relevante	Referência
pCRCT		Plasmídeo " <i>ultra high-copy</i> " com iCas9	182
pAG32		HphMX4 - Resistencia a Higromicina B	150
pFA6KanMX4		KanMX4 - Resistência a geneticina (G418)	149
pSH65		Cre recombinase, Resistência a zeocina	151
pDELHexo5	pCRCT	Cassete de deleção <i>HXK1</i> , <i>HXK2</i> , <i>GLK1</i> , <i>YLR446W</i> e <i>GAL1</i>	Este estudo
pGAL1/GLK1/ADE2	pCRCT	Cassete de deleção <i>GAL1</i> , <i>GLK1</i> e <i>ADE2</i>	Este estudo
pHXK1/YLR/ADE2_mut	pCRCT	Cassete de deleção <i>HXK1</i> , <i>YLR446W</i> e <i>ADE2_mut</i>	Este estudo
pCiGXF1	pRS426	Transportador GXF1 clonado no vetor multicópia de levedura	180
pCs4130	pRS426	Transportador Cs4130 clonado no vetor multicópia de levedura	180
pCiGXF1-GFP	pCiGXF1	Transportador GXF1 acoplado a proteína fluorescente eGFP	Este estudo
pCs4130-GFP	pCs4130	Transportador Cs4130 acoplado a proteína fluorescente eGFP	180

Tabela suplementar S3. Primers utilizados no trabalho

Primer	Sequência 5' --> 3'	Descrição
Hph_HXK1_F	ATAAGATGGTTCATTTAGGTCCAAAGAAACCACAGGCTAGAAAGCCAGCTGAAGCTTCGTA	Deleção de HXK1
Hph_HXK1_R	TTAAGCGCCAATGATACCAAGAGACTTACCTTCGGCAATTTAGGGAGACCGGCAGATC	Deleção de HXK1
Kan_HXK2_F	TAATAAGTACGCTAATTAATAAAAATGGTTCATTTAGGTCCAGCCAGCTGAAGCTTCGTA	Deleção de HXK2
Kan_HXK2_R	TAAGTTTAAGCACCGATGATACCAACGGACTTACCTTCAGCATAGGGAGACCGGCAGATC	Deleção de HXK2
Zeo_GLK1_F	ACAACCTATATACACACAAGATGTCATTTCGACGACTTACACACTCGAAGGCTTTAATTTG	Deleção de GLK1
Zeo_GLK1_R	CGGCACCCACTCCGGAACCATCCTTGGCAATCTTCAAGTGC GCGTAATACGACTCACTATA	Deleção de GLK1
Check_del_HXK1_F	CATAAGGGCATCACTCATAAG	Confirmação de deleção HXK1
Check_del_HXK1_R	CATAAGGGCATCACTCATAAG	Confirmação de deleção HXK1
Check_del_HXK2_F	CACGAAATTACCTCCTGCT	Confirmação de deleção HXK2
Check_del_HXK2_R	TTATCGTCACGAATAAATCC	Confirmação de deleção HXK2
Check_del_GLK1_F	CTCAGCTTCCGTAAACAC	Confirmação de deleção GLK1
Check_del_GLK1_R	CGGTGGGATACGTACACA	Confirmação de deleção GLK1
Check_del_YLR446W_R	TCCTCTACACCAACATCACA	Confirmação de deleção YLR446W
Check_HXK2_F	CAAAGTACCGGTAGTGTC	Confirmação de deleção HXK2 via CRISPR
Check_HXK1_F	GACAAAGAAGGGAGGTAAC	Confirmação de deleção GLK1 via CRISPR
Check_GLK1_F	TCGAACAAATGGAAAAAGGT	Confirmação de deleção GLK1 via CRISPR
Check_YLRW_F	CAGGCTGGCTGCAAAT	Confirmação de deleção YLR446W via CRISPR
Check_pCRCT_F	GAGAAAGTTGTTAAGCACTT	Amplificação do cassete de deleção em pDELhexo5
Check_pCRCT_R	TGCTTGGGTGTTTTGAAG	Amplificação do cassete de deleção em pDELhexo5
HXK1_Seq_del_F	ATATAAGCAATCGGTTTTAC	Sequenciamento de HXK1
HXK1_Seq_del_2R	CTTCGTTAATCTGTTTTGG	Sequenciamento de HXK1
HXK2_Seq_del_F	AAGCCAGAAAGGGTTCC	Sequenciamento de HXK2
HXK2_Seq_del_R	TAGTTTTCCCTGTAGCTTTTC	Sequenciamento de HXK2
GLK1_Seq_del_F	ATTATAAGTGGTGTGCCG	Sequenciamento de GLK1
GLK1_Seq_del_R	GTCTACAGGGTATGAGAAAG	Sequenciamento de GLK1
GAL1_Seq_del_F	TGTTGTTAAAGCGAATATGG	Sequenciamento de GAL1
GAL1_Seq_del_R	AGATTCAGAATACACATGCT	Sequenciamento de GAL1
Check_del_YLR446W_F	TGCCAGGTATTGTTAGCG	Sequenciamento de YLR446W
Check_del_YLR446W_R	TCCTCTACACCAACATCACA	Sequenciamento de YLR446W
KAN_F	GCCAGCTGAAGCTTCGTA	Amplificação do gene KanMX4
Hph_F	GCCAGCTGAAGCTTCGTA	Amplificação do gene hphMX4
Zeo_F	CCATCACGAGATTTTCGATT	Amplificação do gene bleR (resistência a Zeocina)

Tabela suplementar S4. Cassetes de deleção via CRISPR-Cas9 usados

Nome do crRNA	Sequência 5' --> 3'	Descrição
Cassete DelHexo5	CTTTGGTCTCACCAAAACGAGAAAGGTTGTTAAGCACTTTATCGACGAATTGAATAAAGGTTTGACAAGGTAACATTCCAA TGATTCCCAGTTGGGTCATGGAATCCCAACAGGTAAGAATAAAGGTTTGACAAAGAGTTTTAGAGCTATGCTGTTTTGAA TGGTCCCAAAACCGAGGTTACCCCGGAGAAGCTGGACGAATTAAGTCTTACTTCATCGAACAAAGGTCTAGCTCCACCA AAGGAAGGCCACACATTGGCCTCGGACAAAAGGTGCTTACTTCATCGAACAAAGTTTTAGAGCTATGCTGTTTTGAATGGT CCCAAAACGATAACACCCATCTTAGTTTTCTGGGTCAGTGTAGTAAGAAGCAACCAAGGTGTCGTTTTATCAAAGCAACAAC TTCAATTGGGATATTCCTCTTAGTGATGTAAGAAGCAACCAAGTACGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAA ACCATCGTTGATTCGAATTTCTTTAATCAAATCATCTATAACAATATGTACCAGCAAATGGTTATATAAAAAAAGAACGAG TCATCTGAGGCGTCTAAATTCATACAATATGTACCAGGCGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAACCTTG TTGATCTAACAAGTTTTGAAACCGCCCAACCAACTATAATTTAAGAAGTCACTACAGCTGCAAATGTTTTAGCTGCCACG TACGGTGTGTTTTAACCAACTATAATTTAAGAGGTTTTAGAGAGAGACCTTTC	Cassete para deleção de <i>HXK1</i> , <i>HXK2</i> , <i>GLK1</i> , <i>YLR446W</i> e <i>GAL1</i>
Cassete Del <i>GAL1/GLK1/ADE2</i>	CTTTGGTCTCACCAAAACTAATCTTACTTATATGTTGTGGAATGTAAAGAGCCCCATTATCTTAGCCACACAGGGACAAA ATTCTTGATATGCTTTCAACCGCTGCGTTTTGGATACAACCAACTATAATTTAAGAGGTTTTAGAGCTATGCTGTTTTGAAT GGTCCCAAAACCCACTAATACAACCTCTATCATACACAAGATGTCATTTCGACGACTTACACACCAAGGATGGTCCGGAGT GGGTGCCGCTTGTGTGCGCTTGTAGCATGATGCTTACTTCATCGAACAAAGTTTTAGAGCTATGCTGTTTTGAATGGTC CCAAAACCCAAACAAATAAGCAACTCCAATGACCACGTTAATGGCTCCTCGAGAACCTCTCCCGCATCTGAATAAACTAG CTGAAAAATGTGATGTGCTAACGATTGGATATCAAGAGGATTGAAAGTTTTAGAGAGAGACCTTTC	Cassete para deleção de <i>GAL1</i> , <i>GLK1</i> e <i>ADE2</i>
Cassete Del <i>HXK1/YLR/ADE2</i> _mut	CTTTGGTCTCACCAAAACCGAGGATACTGAACACCGGGAAGTTAACTGAGTATATCGGTTATCATTACAACCAAGTTCAA TAATATATTTTCATCATTTCTTACCATCTTCTCGATCTCTATAACAATATGTACCAGGCGTTTTAGAGCTATGCTGTTTTGAATG GTCCCAAAACCTTAGGACCGTTGAGAGGAATAGTAACAAGTGAACGCAACAAAGATTGTTGAAAAAAGAATTGCCGAAG GTAAGTCTCTTGGTATCATTGGCGCTTAATGAATAAAGGTTTGACAAAGAGTTTTAGAGCTATGCTGTTTTGAATGGTCCC AAAACCCAAACAAATAAGCAACTCCAATGACCACGTTAATGGCTCCTTTTCCAATCCTCTTGATATCGAAAAACTAGCTGA AAAATGTGATGTGCTAACGATTGAAGATGCGGGAGAGGTTCTCGGTTTTAGAGAGAGACCTTTC	Cassete para deleção de <i>HXK1</i> , <i>YLR446W</i> e <i>ADE2_mut</i>



Tabela suplementar S5. DNAs doadores usados para experimentos com CRISPR-Cas9

Nome do DNA doador	Sequência 5' --> 3'	Descrição
<i>HXK1</i> doador Hexo 5	GAGAAAGGTTGTTAAGCACTTTATCGACGAATTGAATAAAGGTTTGACAAGGTAACATTCCA ATGATTCCCGGTTGGGTCATGGAATTCCTCAACAGGTAA	DNA doador para cassete DelHexo5
<i>HXK2</i> doador Hexo 5	GATAACACCCATCTTAGTTTCTGGGTCAGTGTAGTAAGAAGCAACCAAAGGTGTCGTTTATC AAAGCAACAACCTCAATTGGGATATTCCTCTTAGTGAT	DNA doador para cassete DelHexo5
<i>GLK1</i> doador Hexo 5	CGAGGTTACCCCGAGAAGCTGGACGAATTAAGTCTTACTTTCATCGAACAAAGGTCTAGC TCCACCAAAGGAAGGCCACACATTGGCCTCGGACAAAGG	DNA doador para cassete DelHexo5
<i>GAL1</i> doador Hexo 5	CCTTGTTGTATCTAACAAGTTTGAACCGCCCAACCAACTATAATTTAAGAAGTCACTACA GCTGCAAATGTTTTAGCTGCCACGTACGGTGTGTTTT	DNA doador para cassete DelHexo5
<i>YLR446W</i> doador Hexo 5	CATCGTTGATTGCAATTTCTTTAATCAAATCATCTATAACAATATGTACCAGCAAATGGTTATA TAAAAAAAAAGAACGAGTCATCTGAGGCGTCTAAATT	DNA doador para cassete DelHexo5
<i>GLK1</i> doador	CCACTAATACAACCTCTATCATAACAAGATGTCATTGACGACTTACACACCAAGGATGGTT CCGGAGTGGGTGCCGCCTTGTGTGCGCTTGTAGCATGA	DNA doador para cassete Del <i>GAL1/GLK1/ADE2</i>
<i>GAL1</i> doador	TAATCTTACTTATATGTTGTGGAAATGTAAAGAGCCCCATTATCTTAGCCACACAGGGACAA AATTCTTGATATGCTTTCAACCGCTGCGTTTTGGATAC	DNA doador para cassete Del <i>GAL1/GLK1/ADE2</i>
<i>ADE2</i> doador	CCAAACAAATAAGCAACTCCAATGACCACGTTAATGGCTCCTCGAGAACCTCTCCCGCATC TGAATAAACTAGCTGAAAAATGTGATGTGCTAACGATTG	DNA doador para cassete Del <i>GAL1/GLK1/ADE2</i> e Del <i>HXT1-7/GAL2/ADE2</i>
<i>YLR446W</i> doador	CAGGGATACTGAACACCGGGGAAGTAACTGAGTATATCGGTTATCATTACAACCAAGTTC AATAATATATTTTCATCATTCTTACCATCTTCTCGATCT	DNA doador para cassete Del <i>HXK1/YLR/ADE_mut</i>
<i>HXK1</i> doador	CTTAGGACCGTTGAGAGGAATAGTAACAAGTGAACGCAACAAAGATTGTTGAAAAAAGAA TTGCCGAAGGTAAGTCTCTTGGTATCATTGGCGCTTAAT	DNA doador para cassete Del <i>HXK1/YLR/ADE_mut</i>
<i>ADE_mut</i> doador	CCAAACAAATAAGCA'ACTCCAATGACCACGTTAATGGCTCCTTTTCCAATCCTCTTGATATC GAAAAACTAGCTGAAAAATGTGATGTGCTAACGATTGA	DNA doador para cassete Del <i>HXK1/YLR/ADE_mut</i>

Tabela suplementar S6. gRNAs utilizados no trabalho

Nome do gRNA	Sequência 5' --> 3'	Descrição
<i>HXK1</i> Hexo5	GAATAAAGGTTTGACAAAGA	gRNA usado no cassete DelHexo5
<i>HXK2</i> Hexo 5	GTAAGAAGCAACCAAAGTAC	gRNA usado no cassete DelHexo5
<i>GLK1</i> Hexo 5	TGCTTACTTCATCGAACAAA	gRNA usado no cassete DelHexo5
<i>YLR446W</i> Hexo 5	CTATACAATATGTACCAGGC	gRNA usado no cassete DelHexo5
<i>GAL1</i> Hexo 5	AACCAACTATAATTTAAGAG	gRNA usado no cassete DelHexo5
<i>GLK1</i>	TGCTTACTTCATCGAACAAA	gRNA usado no cassete Del <i>GAL1/GLK1/ADE2</i>
<i>GAL1</i>	AACCAACTATAATTTAAGAG	gRNA usado no cassete Del <i>GAL1/GLK1/ADE2</i>
<i>ADE2</i>	GATATCAAGAGGATTGGAAA	gRNA usado no cassete Del <i>GAL1/GLK1/ADE2</i>
<i>YLR446W</i>	CTATACAATATGTACCAGGC	gRNA usado no cassete Del <i>HXK1/GLK1/ADE2_mut</i>
<i>HXK1</i>	GAATAAAGGTTTGACAAAGA	gRNA usado no cassete Del <i>HXK1/GLK1/ADE2_mut</i>
<i>ADE_mut</i>	AGATGCGGGAGAGGTTCTCG	gRNA usado no cassete Del <i>HXK1/GLK1/ADE2_mut</i>

## 2. Referências

1. Keeling, C. D. Industrial production of carbon dioxide from fossil fuels and limestone. *Tellus* **25**, 174–198 (1973).
2. Andres, R. J. *et al.* Carbon dioxide emissions from fossil-fuel use, 1751-1950. *Tellus, Ser. B Chem. Phys. Meteorol.* **51**, 759–765 (1999).
3. Bin, S. & Dowlatabadi, H. Consumer lifestyle approach to US energy use and the related CO2 emissions. *Energy Policy* **33**, 197–208 (2005).
4. Vitousek, P. M. Human Domination of Earth's Ecosystems. *Science (80-. )*. **277**, 494–499 (1997).
5. Dalal, R. C. & Allen, D. E. Greenhouse gas fluxes from natural ecosystems. *Aust. J. Bot.* **56**, 369–407 (2008).
6. Owen, N. a., Inderwildi, O. R. & King, D. a. The status of conventional world oil reserves—Hype or cause for concern? *Energy Policy* **38**, 4743–4749 (2010).
7. WMO. World Climate Conference A Conference of Experts on Climate and Mankind. in *World climate Conference: A conference of experts on climate and mankind* 12–23 (1979). doi:10.1016/0004-6981(78)90305-0
8. Gupta, J. A history of international climate change policy. *Wiley Interdiscip. Rev. Clim. Chang.* **1**, 636–653 (2010).
9. Bolin, B. A history of the science and politics of climate change: The role of the intergovernmental panel on climate change. *A Hist. Sci. Polit. Clim. Chang. Role Intergov. Panel Clim. Chang.* **9780521880**, 1–277 (2007).
10. FEDERATIVE REPUBLIC OF BRAZIL & INTENDED. *Intended Nationally Determined Contribution: Towards achieving the objective of the United Nations Framework Convention on Climate Change. Intended Nationally Determined Contribution TOWARDS ACHIEVING THE OBJECTIVE OF THE UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE* **9**, (2015).
11. Naik, S. N., Goud, V. V., Rout, P. K. & Dalai, A. K. Production of first and second generation biofuels: A comprehensive review. *Renew. Sustain. Energy Rev.* **14**, 578–597 (2010).
12. Santos, F. A., Queiróz, J. H. De, Colodette, J. L., Fernandes, S. A. & Guimarães, V. M. POTENCIAL DA PALHA DE CANA-DE-AÇÚCAR PARA PRODUÇÃO DE ETANOL. *Quim. Nova* **35**, 1004–1010 (2012).
13. Andrietta, G. S., Steckelberg, C. & Andrietta, R. *Bioetanol – Brasil, 30 anos na vanguarda.*

- (2006).
14. Matsuoka, S., Ferro, J. & Arruda, P. The Brazilian experience of sugarcane ethanol industry. *Vitr. Cell. Dev. Biol. - Plant* **45**, 372–381 (2009).
  15. Groom, M. J., Gray, E. M. & Townsend, P. a. Biofuels and biodiversity: Principles for creating better policies for biofuel production. *Conserv. Biol.* **22**, 602–609 (2008).
  16. Macrelli, S. Ethanol from sugarcane lignocellulosic residues. (2014).
  17. Matsushika, A., Inoue, H., Kodaki, T. & Sawayama, S. Ethanol production from xylose in engineered *Saccharomyces cerevisiae* strains: current state and perspectives. *Appl. Microbiol. Biotechnol.* **84**, 37–53 (2009).
  18. Stephen, J. D., Mabee, W. E. & Saddler, J. N. Will second-generation ethanol be able to compete with first-generation ethanol? Opportunities for cost reduction. *Biofuels Bioprod. Bioref* 159–176 (2012). doi:10.1002/bbb
  19. dos Santos, L. V. *et al.* Second-Generation Ethanol: The Need is Becoming a Reality. *Ind. Biotechnol.* **12**, 40–57 (2016).
  20. Vanholme, R., Demedts, B., Morreel, K., Ralph, J. & Boerjan, W. Lignin biosynthesis and structure. *Plant Physiol.* **153**, 895–905 (2010).
  21. Hernández-Beltrán, J. U. *et al.* Insight into pretreatment methods of lignocellulosic biomass to increase biogas yield: Current state, challenges, and opportunities. *Appl. Sci.* **9**, (2019).
  22. Zhao, Q. *et al.* Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. (2010). doi:10.1073/pnas.1009170107/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1009170107
  23. Balat, M. Production of bioethanol from lignocellulosic materials via the biochemical pathway: A review. *Energy Convers. Manag.* **52**, 858–875 (2011).
  24. Dias, M. O. S. *et al.* Second generation ethanol in Brazil: can it compete with electricity production? *Bioresour. Technol.* **102**, 8964–71 (2011).
  25. Söderström, J., Galbe, M. & Zacchi, G. Effect of washing on yield in one- and two-step steam pretreatment of softwood for production of ethanol. *Biotechnol. Prog.* **20**, 744–749 (2004).
  26. Söderström, J., Pilcher, L., Galbe, M. & Zacchi, G. Two-step steam pretreatment of softwood by dilute H<sub>2</sub>SO<sub>4</sub> impregnation for ethanol production. *Biomass and Bioenergy* **24**, 475–486 (2003).

27. Schell, D. J. & Harwood, C. Milling of lignocellulosic biomass - Results of pilot-scale testing. *Appl. Biochem. Biotechnol.* **45–46**, 159–168 (1994).
28. Zheng, Y., Zhao, J., Xu, F. & Li, Y. Pretreatment of lignocellulosic biomass for enhanced biogas production. *Prog. Energy Combust. Sci.* **42**, 35–53 (2014).
29. Canilha, L. *et al.* Bioconversion of sugarcane biomass into ethanol: An overview about composition, pretreatment methods, detoxification of hydrolysates, enzymatic saccharification, and ethanol fermentation. *J. Biomed. Biotechnol.* **2012**, (2012).
30. Das, P., Ganesh, A. & Wangikar, P. Influence of pretreatment for deashing of sugarcane bagasse on pyrolysis products. *Biomass and Bioenergy* **27**, 445–457 (2004).
31. Sarkar, N., Ghosh, S. K., Bannerjee, S. & Aikat, K. Bioethanol production from agricultural wastes: An overview. *Renew. Energy* **37**, 19–27 (2012).
32. Chaturvedi, V. & Verma, P. An overview of key pretreatment processes employed for bioconversion of lignocellulosic biomass into biofuels and value added products. *3 Biotech* **3**, 415–431 (2013).
33. Bensah, E. C. & Mensah, M. Chemical pretreatment methods for the production of cellulosic ethanol: Technologies and innovations. *Int. J. Chem. Eng.* **2013**, (2013).
34. Gírio, F. M. *et al.* Hemicelluloses for fuel ethanol: A review. *Bioresour. Technol.* **101**, 4775–800 (2010).
35. Kumar, P., Barrett, D. M., Delwiche, M. J. & Stroeve, P. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Ind. Eng. Chem. Res.* **48**, 3713–3729 (2009).
36. Chandel, A. K., Chandrasekhar, G., Silva, M. B. & Silvério Da Silva, S. The realm of cellulases in biorefinery development. *Crit. Rev. Biotechnol.* **32**, 187–202 (2012).
37. Cragg, S. M. *et al.* Lignocellulose degradation mechanisms across the Tree of Life. *Curr. Opin. Chem. Biol.* **29**, 108–119 (2015).
38. Hellden, G. 19-Year-Old Students' Views of the Development of Their Own Understanding of Biological Processes. *Proc. Third Int. Conf. Sci. Educ. Res. Knowl. Based Soc. Vol. 1* **65**, 435–437 (2001).
39. Gusakov, A. V. Alternatives to *Trichoderma reesei* in biofuel production. *Trends Biotechnol.* **29**, 419–425 (2011).
40. Van Den Brink, J. & De Vries, R. P. Fungal enzyme sets for plant polysaccharide

- degradation. *Appl. Microbiol. Biotechnol.* **91**, 1477–1492 (2011).
41. Vlasenko, E., Schüle, M., Cherry, J. & Xu, F. Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases. *Bioresour. Technol.* **101**, 2405–2411 (2010).
  42. Andlar, M. *et al.* Lignocellulose degradation: An overview of fungi and fungal enzymes involved in lignocellulose degradation. *Eng. Life Sci.* **18**, 768–778 (2018).
  43. Eliasson, A., Christensson, C., Fredrik, C. & Wahlbom, C. F. Anaerobic Xylose Fermentation by Recombinant *Saccharomyces* and XKS1 in Mineral Medium Chemostat Cultures Anaerobic Xylose Fermentation by Recombinant *Saccharomyces*. *Appl. Environmental Microbiol.* **66**, 3381–3386 (2000).
  44. Nielsen, J., Larsson, C., van Maris, A. & Pronk, J. Metabolic engineering of yeast for production of fuels and chemicals. *Curr. Opin. Biotechnol.* **24**, 398–404 (2013).
  45. Jeffries, T. W. Engineering yeasts for xylose metabolism. *Curr. Opin. Biotechnol.* **17**, 320–326 (2006).
  46. Toivari, M. H., Salusjarvi, L., Ruohonen, L. & Penttila, M. Endogenous Xylose Pathway in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **70**, 3681–3686 (2004).
  47. Watanabe, S. *et al.* Ethanol production from xylose by recombinant *Saccharomyces cerevisiae* expressing protein-engineered NADH-preferring xylose reductase from *Pichia stipitis*. *Microbiology* **153**, 3044–3054 (2007).
  48. Liang, L., Zhang, J. & Lin, Z. Altering coenzyme specificity of *Pichia stipitis* xylose reductase by the semi-rational approach CASTing. *Microb. Cell Fact.* **6**, 1–11 (2007).
  49. Runquist, D., Hahn-Hägerdal, B. & Bettiga, M. Increased ethanol productivity in xylose-utilizing *Saccharomyces cerevisiae* via a randomly mutagenized xylose reductase. *Appl. Environ. Microbiol.* **76**, 7796–7802 (2010).
  50. Watanabe, S. *et al.* Ethanol production from xylose by recombinant *Saccharomyces cerevisiae* expressing protein engineered NADP<sup>+</sup>-dependent xylitol dehydrogenase. *J. Biotechnol.* **130**, 316–319 (2007).
  51. Matsushika, A. *et al.* Expression of protein engineered NADP<sup>+</sup>-dependent xylitol dehydrogenase increases ethanol production from xylose in recombinant *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **81**, 243–255 (2008).
  52. Bera, A. K., Ho, N. W. Y., Khan, A. & Sedlak, M. A genetic overhaul of *Saccharomyces cerevisiae* 424A(LNH-ST) to improve xylose fermentation. *J. Ind. Microbiol. Biotechnol.* **38**, 617–626 (2011).

53. Hahn-Hägerdal, B. *et al.* Metabolic engineering of *Saccharomyces cerevisiae* for xylose utilization. *Adv. Biochem. Eng. Biotechnol.* **73**, 53–84 (2001).
54. Sarthy, A. V. *et al.* Expression of the *Escherichia coli* xylose isomerase gene in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **53**, 1996–2000 (1987).
55. Temer, B. *et al.* Conversion of an inactive xylose isomerase into a functional enzyme by co-expression of GroEL-GroES chaperonins in *Saccharomyces cerevisiae*. *BMC Biotechnol.* **17**, 1–13 (2017).
56. Hou, J., Jiao, C., Peng, B., Shen, Y. & Bao, X. Mutation of a regulator Ask10p improves xylose isomerase activity through up-regulation of molecular chaperones in *Saccharomyces cerevisiae*. *Metab. Eng.* **38**, 241–250 (2016).
57. Brat, D., Boles, E. & Wiedemann, B. Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **75**, 2304–2311 (2009).
58. Kuyper, M., Winkler, A. a., Van Dijken, J. P. & Pronk, J. T. Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: A proof of principle. *FEMS Yeast Res.* **4**, 655–664 (2004).
59. Kuyper, M. *et al.* High-level functional expression of a fungal xylose isomerase: The key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res.* **4**, 69–78 (2003).
60. 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).
61. Sonderegger, M. & Sauer, U. Evolutionary engineering of *Saccharomyces cerevisiae* for anaerobic growth on xylose. *Appl. Environ. Microbiol.* **69**, 1990–1998 (2003).
62. Scalcinati, G. *et al.* Evolutionary engineering of *Saccharomyces cerevisiae* for efficient aerobic xylose consumption. *FEMS Yeast Res.* **12**, 582–597 (2012).
63. dos Santos, L. V. *et al.* Unraveling the genetic basis of xylose consumption in engineered *Saccharomyces cerevisiae* strains. *Sci. Rep.* **6**, 38676 (2016).
64. Verhoeven, M. D., Lee, M., Kamoen, L. & Broek, M. Van Den. Mutations in PMR1 stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Nat. Publ. Gr.* 1–11 (2017).  
doi:10.1038/srep46155
65. Sato, T. K. *et al.* Directed Evolution Reveals Unexpected Epistatic Interactions That Alter

- Metabolic Regulation and Enable Anaerobic Xylose Use by *Saccharomyces cerevisiae*. 1–31 (2016). doi:10.1371/journal.pgen.1006372
66. 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**, 688–99 (2010).
  67. Kim, S. R., Ha, S. J., Wei, N., Oh, E. J. & Jin, Y. S. Simultaneous co-fermentation of mixed sugars: A promising strategy for producing cellulosic ethanol. *Trends Biotechnol.* **30**, 274–282 (2012).
  68. Katahira, S. *et al.* Improvement of ethanol productivity during xylose and glucose co-fermentation by xylose-assimilating *S. cerevisiae* via expression of glucose transporter Sut1. *Enzyme Microb. Technol.* **43**, 115–119 (2008).
  69. Moysés, D., Reis, V., Almeida, J., Moraes, L. & Torres, F. Xylose Fermentation by *Saccharomyces cerevisiae*: Challenges and Prospects. *Int. J. Mol. Sci.* **17**, 207 (2016).
  70. Hou, J., Qiu, C., Shen, Y., Li, H. & Bao, X. Engineering of *Saccharomyces cerevisiae* for the efficient co-utilization of glucose and xylose. *FEMS Yeast Res.* **17**, 1–11 (2017).
  71. Sedlak, M. & Ho, N. W. Y. Characterization of the effectiveness of hexose transporters for transporting xylose during glucose and xylose co-fermentation by a recombinant *Saccharomyces* yeast. *Yeast* **21**, 671–684 (2004).
  72. Saloheimo, A. *et al.* Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases. *Appl. Microbiol. Biotechnol.* **74**, 1041–1052 (2007).
  73. Leandro, M. J., Spencer-Martins, I. & Gonçalves, P. The expression in *Saccharomyces cerevisiae* of a glucose/xylose symporter from *Candida intermedia* is affected by the presence of a glucose/xylose facilitator. *Microbiology* **154**, 1646–55 (2008).
  74. Boles, E. & Hollenberg, C. P. The molecular genetics of hexose transport in yeasts. *FEMS Microbiology Reviews* **21**, 85–111 (1997).
  75. Quistgaard, E. M., Löw, C., Guettou, F. & Nordlund, P. Understanding transport by the major facilitator superfamily (MFS): Structures pave the way. *Nature Reviews Molecular Cell Biology* **17**, 123–132 (2016).
  76. Henderson, P. J. F. & Baldwin, S. A. This is about the in and the out. *Nat. Struct. Mol. Biol.* **20**, 654–655 (2013).
  77. Pao, S. S., Paulsen, I. T. & Saier, M. H. Major facilitator superfamily. *Microbiol. Mol. Biol.*



- Rev.* **62**, 1–34 (1998).
78. Forrest, L. R., Krämer, R. & Ziegler, C. The structural basis of secondary active transport mechanisms. *Biochim. Biophys. Acta - Bioenerg.* **1807**, 167–188 (2011).
  79. Marger, M. D. & Saier, M. H. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. *Trends Biochem. Sci.* **18**, 13–20 (1993).
  80. Reddy, V. S., Shlykov, M. A., Castillo, R., Sun, E. I. & Saier, M. H. The major facilitator superfamily (MFS) revisited. *FEBS J.* **279**, 2022–35 (2012).
  81. Young, E., Poucher, A., Comer, A., Bailey, A. & Alper, H. Functional survey for heterologous sugar transport proteins, using *Saccharomyces cerevisiae* as a host. *Appl. Environ. Microbiol.* **77**, 3311–3319 (2011).
  82. Farwick, A., Bruder, S., Schadeweg, V., Oreb, M. & Boles, E. Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc. Natl. Acad. Sci.* **111**, 5159–5164 (2014).
  83. Weierstall, T., Hollenberg, C. P. & Boles, E. Cloning and characterization of three genes (SUT1-3) encoding glucose transporters of the yeast *Pichia stipitis*. *Mol. Microbiol.* **31**, 871–883 (1999).
  84. Leandro, M. J., Gonçalves, P. & Spencer-Martins, I. Two glucose/xylose transporter genes from the yeast *Candida intermedia*: first molecular characterization of a yeast xylose–H<sup>+</sup> symporter. *Biochem. J.* **395**, 543–549 (2006).
  85. Du, J., Li, S. & Zhao, H. Discovery and characterization of novel d-xylose-specific transporters from *Neurospora crassa* and *Pichia stipitis*. *Mol. Biosyst.* **6**, 2150 (2010).
  86. Ferreira, D. *et al.* XYLH encodes a xylose/H<sup>+</sup> symporter from the highly related yeast species *Debaryomyces fabryi* and *Debaryomyces hansenii*. *FEMS Yeast Res.* **13**, 585–596 (2013).
  87. dos Reis, T. F. *et al.* Identification and characterization of putative xylose and cellobiose transporters in *Aspergillus nidulans*. *Biotechnol. Biofuels* **9**, 204 (2016).
  88. Runquist, D., Fonseca, C., Rådström, P., Spencer-Martins, I. & Hahn-Hägerdal, B. Expression of the Gxf1 transporter from *Candida intermedia* improves fermentation performance in recombinant xylose-utilizing *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **82**, 123–130 (2009).
  89. Fonseca, C. *et al.* The glucose/xylose facilitator Gxf1 from *Candida intermedia* expressed in a xylose-fermenting industrial strain of *Saccharomyces cerevisiae* increases xylose uptake

- in SSCF of wheat straw. *Enzyme Microb. Technol.* **48**, 518–525 (2011).
90. Young, E. M., Comer, A. D., Huang, H. & Alper, H. S. A molecular transporter engineering approach to improving xylose catabolism in *Saccharomyces cerevisiae*. *Metab. Eng.* **14**, 401–411 (2012).
  91. Moon, J., Lewis Liu, Z., Ma, M. & Slininger, P. J. New genotypes of industrial yeast *Saccharomyces cerevisiae* engineered with YXI and heterologous xylose transporters improve xylose utilization and ethanol production. *Biocatal. Agric. Biotechnol.* **2**, 247–254 (2013).
  92. Runquist, D., Hahn-Hagerdal, B. & Radstrom, P. Comparison of heterologous xylose transporters in recombinant *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **3**, 5 (2010).
  93. Nijland, J. G. *et al.* Engineering of an endogenous hexose transporter into a specific D-xylose transporter facilitates glucose-xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **7**, 168 (2014).
  94. Reider Apel, A., Ouellet, M., Szmidt-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).
  95. Shin, H. Y. *et al.* An engineered cryptic Hxt11 sugar transporter facilitates glucose–xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **8**, 176 (2015).
  96. Hamacher, T., Becker, J., Gárdonyi, M., Hahn-Hägerdal, B. & Boles, E. Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* **148**, 2783–2788 (2002).
  97. Li, H., Schmitz, O. & Alper, H. S. Enabling glucose/xylose co-transport in yeast through the directed evolution of a sugar transporter. *Appl. Microbiol. Biotechnol.* **100**, 10215–10223 (2016).
  98. Farwick, A., Bruder, S., Schadeweg, V., Oreb, M. & Boles, E. Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5159–64 (2014).
  99. Stambuk, B. U., Franden, M. A., Singh, A. & Zhang, M. D-Xylose Transport by *Candida succiphila* and *Kluyveromyces marxianus*. *Appl. Biochem. Biotechnol.* **105**, 255–263 (2003).
  100. Bhosale, S. H., Rao, M. B. & Deshpande, V. V. Molecular and industrial aspects of glucose isomerase. *Microbiol. Rev.* **60**, 280–300 (1996).
  101. Karhumaa, K., Wu, B. & Kielland-Brandt, M. C. Conditions with high intracellular glucose

- inhibit sensing through glucose sensor SnF3 in *Saccharomyces cerevisiae*. *J. Cell. Biochem.* **110**, 920–925 (2010).
102. Jansen, M. L. a, De Winde, J. H. & Pronk, J. T. Hxt-carrier-mediated glucose efflux upon exposure of *Saccharomyces cerevisiae* to excess maltose. *Appl. Environ. Microbiol.* **68**, 4259–4265 (2002).
  103. Reznicek, O. *et al.* Improved xylose uptake in *Saccharomyces cerevisiae* due to directed evolution of galactose permease Gal2 for sugar co-consumption. *J. Appl. Microbiol.* **119**, 99–111 (2015).
  104. Nanda, S., Azargohar, R., Dalai, A. K. & Kozinski, J. A. An assessment on the sustainability of lignocellulosic biomass for biorefining. *Renew. Sustain. Energy Rev.* **50**, 925–941 (2015).
  105. Lee, R. A. & Lavoie, J.-M. From first- to third-generation biofuels: Challenges of producing a commodity from a biomass of increasing complexity. *Anim. Front.* **3**, 6–11 (2013).
  106. Nigam, P. S. & Singh, A. Production of liquid biofuels from renewable resources. *Prog. Energy Combust. Sci.* **37**, 52–68 (2011).
  107. Jordan, D. B. *et al.* Plant cell walls to ethanol. *Biochem. J.* **442**, 241–252 (2012).
  108. Nielsen, J. & Keasling, J. D. Engineering Cellular Metabolism. *Cell* **164**, 1185–1197 (2016).
  109. Lu, H. *et al.* A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nat. Commun.* **10**, (2019).
  110. Choi, K. R. *et al.* Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. *Trends Biotechnol.* **37**, 817–837 (2019).
  111. Long, C. P. & Antoniewicz, M. R. How adaptive evolution reshapes metabolism to improve fitness: recent advances and future outlook. *Curr. Opin. Chem. Eng.* **22**, 209–215 (2018).
  112. Maier, A., Völker, B., Boles, E. & Fuhrmann, G. F. Characterisation of glucose transport in *Saccharomyces cerevisiae* with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. *FEMS Yeast Res.* **2**, 539–550 (2002).
  113. Aeling, K. A. *et al.* Co-fermentation of xylose and cellobiose by an engineered *saccharomyces cerevisiae*. *J. Ind. Microbiol. Biotechnol.* **39**, 1597–1604 (2012).
  114. Young, E. M., Tong, A., Bui, H., Spofford, C. & Alper, H. S. Rewiring yeast sugar transporter preference through modifying a conserved protein motif. *Proc. Natl. Acad. Sci.* **111**, 131–136 (2014).

115. Bracher, J. M. *et al.* The *Penicillium chrysogenum* transporter PcAraT enables high-affinity, glucose-insensitive l-arabinose transport in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **11**, 63 (2018).
116. Parachin, N. S., Bergdahl, B., van Niel, E. W. J. & Gorwa-Grauslund, M. F. Kinetic modelling reveals current limitations in the production of ethanol from xylose by recombinant *Saccharomyces cerevisiae*. *Metab. Eng.* **13**, 508–517 (2011).
117. Hector, R. E., Qureshi, N., Hughes, S. R. & Cotta, M. a. Expression of a heterologous xylose transporter in a *Saccharomyces cerevisiae* strain engineered to utilize xylose improves aerobic xylose consumption. *Appl. Microbiol. Biotechnol.* **80**, 675–684 (2008).
118. Wang, C. *et al.* Cloning and characterization of heterologous transporters in *Saccharomyces cerevisiae* and identification of important amino acids for xylose utilization. *Metab. Eng.* **30**, 79–88 (2015).
119. Silva, S. S. & Afschar, A. S. Microbial production of xylitol from D-xylose using *Candida tropicalis*. *Bioprocess Eng.* **11**, 129–134 (1994).
120. Parajó, J. C., Domínguez, H. & Domínguez, J. M. Biotechnological production of xylitol. Part 2: Operation in culture media made with commercial sugars. *Bioresour. Technol.* **65**, 203–212 (1998).
121. Guo, C. *et al.* Screening and characterization of yeasts for xylitol production. *J. Appl. Microbiol.* **101**, 1096–1104 (2006).
122. Borelli, G., José, J., Teixeira, P. J. P. L., dos Santos, L. V. & Pereira, G. A. G. De Novo Assembly of *Candida sojae* and *Candida boidinii* Genomes, Unexplored Xylose-Consuming Yeasts with Potential for Renewable Biochemical Production. *Genome Announc.* **4**, e01551-15 (2016).
123. Borelli, G. *et al.* Positive Selection Evidence in Xylose-Related Genes Suggests Methylglyoxal Reductase as a Target for the Improvement of Yeasts' Fermentation in Industry. *Genome Biol. Evol.* **11**, 1923–1938 (2019).
124. Conant, G. C. Comparative genomics as a time machine: How relative gene dosage and metabolic requirements shaped the time-dependent resolution of yeast polyploidy. *Mol. Biol. Evol.* **31**, 3184–3193 (2014).
125. Wiczorke, R. *et al.* Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett.* **464**, 123–128 (1999).
126. Tiukova, I. A. *et al.* Identification and characterisation of two high-affinity glucose

- transporters from the spoilage yeast *Brettanomyces bruxellensis*. *FEMS Microbiol. Lett.* **366**, 1–9 (2019).
127. Horák, J. Regulations of sugar transporters: insights from yeast. *Current genetics* **59**, 1–31 (2013).
  128. Ozcan, S. & Johnston, M. Function and regulation of yeast hexose transporters. *Microbiol. Mol. Biol. Rev.* **63**, 554–69 (1999).
  129. Krampe, S. & Boles, E. Starvation-induced degradation of yeast hexose transporter Hxt7p is dependent on endocytosis, autophagy and the terminal sequences of the permease. *FEBS Lett.* **513**, 193–196 (2002).
  130. Ye, L., Berden, J. A., van Dam, K. & Kruckeberg, A. L. Expression and activity of the Hxt7 high-affinity hexose transporter of *Saccharomyces cerevisiae*. *Yeast* **18**, 1257–1267 (2001).
  131. Snowdon, C. & van der Merwe, G. Regulation of Hxt3 and Hxt7 Turnover Converges on the Vid30 Complex and Requires Inactivation of the Ras/cAMP/PKA Pathway in *Saccharomyces cerevisiae*. *PLoS One* **7**, 1–10 (2012).
  132. Subtil, T. & Boles, E. Competition between pentoses and glucose during uptake and catabolism in recombinant *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **5**, 14 (2012).
  133. Yang, J. & Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **43**, W174–W181 (2015).
  134. Quistgaard, E. M., Löw, C., Moberg, P., Trésaugues, L. & Nordlund, P. Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. *Nat. Struct. Mol. Biol.* **20**, 766–768 (2013).
  135. Sun, L. *et al.* Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* **490**, 361–366 (2012).
  136. Grudinin, S. NOLB: Nonlinear Rigid Block Normal-Mode Analysis Method. (2017). doi:10.1021/acs.jctc.7b00197
  137. Brink, D. P., Borgström, C., Tueros, F. G. & Grauslund, M. F. G. Real - time monitoring of the sugar sensing in *Saccharomyces cerevisiae* indicates endogenous mechanisms for xylose signaling. *Microb. Cell Fact.* 1–17 (2016). doi:10.1186/s12934-016-0580-x
  138. Osiro, K. O. *et al.* Assessing the effect of d-xylose on the sugar signaling pathways of *Saccharomyces cerevisiae* in strains engineered for xylose transport and assimilation. 1–15 (2018). doi:10.1093/femsyr/fox096

139. Knoshaug, E. P. *et al.* Novel transporters from *Kluyveromyces marxianus* and *Pichia guilliermondii* expressed in *Saccharomyces cerevisiae* enable growth on l -arabinose and d -xylose. *Yeast* **32**, 615–628 (2015).
140. Young, E. M., Tong, A., Bui, H., Spofford, C. & Alper, H. S. Rewiring yeast sugar transporter preference through modifying a conserved protein motif. *Proc. Natl. Acad. Sci.* **111**, 131–136 (2014).
141. Lane, S. *et al.* Glucose repression can be alleviated by reducing glucose phosphorylation rate in *Saccharomyces cerevisiae*. *Sci. Rep.* **8**, 2613 (2018).
142. Wisedchaisri, G., Park, M., Iadanza, M. G., Zheng, H. & Gonen, T. Proton-coupled sugar transport in the prototypical major facilitator superfamily protein XylE. *Nat. Commun.* **5**, 4521 (2014).
143. Petrović, D., Risso, V. A., Kamerlin, S. C. L. & Sanchez-Ruiz, J. M. Conformational dynamics and enzyme evolution. *J. R. Soc. Interface* **15**, 20180330 (2018).
144. Bauer, J. A., Pavlović, J. & Bauerová-Hlinková, V. Normal mode analysis as a routine part of a structural investigation. *Molecules* **24**, (2019).
145. Skjaerven, L., Hollup, S. M. & Reuter, N. Normal mode analysis for proteins. *J. Mol. Struct. THEOCHEM* **898**, 42–48 (2009).
146. Qureshi, A. A. *et al.* The molecular basis for sugar import in malaria parasites. *Nature* (2020). doi:10.1038/s41586-020-1963-z
147. Ausubel, F. M. *et al.* *Current Protocols in Molecular Biology*. *Molecular Biology* **1**, (2003).
148. Cadete, R. M. *et al.* *Spathaspora brasiliensis* sp. nov., *Spathaspora suhii* sp. nov., *Spathaspora roraimanensis* sp. nov. and *Spathaspora xylofermentans* sp. nov., four novel d-xylose-fermenting yeast species from Brazilian Amazonian forest. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **103**, 421–431 (2013).
149. Wach, A., Brachat, A., Pöhlmann, R. & Philippsen, P. New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* **10**, 1793–808 (1994).
150. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
151. 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).

152. Fell, J. W., Boekhout, T., Fonseca, A., Scorzetti, G. & Statzell-Tallman, A. Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. *Int. J. Syst. Evol. Microbiol.* **50**, 1351–1371 (2000).
153. Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
154. Kurtzman, C. P. Use of gene sequence analyses and genome comparisons for yeast systematics. *Int. J. Syst. Evol. Microbiol.* **64**, 325–332 (2014).
155. Li, L. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Res.* **13**, 2178–2189 (2003).
156. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772–772 (2012).
157. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164–1165 (2011).
158. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
159. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
160. Christianson, T. W., Sikorski, R. S., Dante, M., Shero, J. H. & Hieter, P. Multifunctional yeast high-copy-number shuttle vectors. *Gene* **110**, 119–22 (1992).
161. Kaishima, M., Ishii, J., Matsuno, T., Fukuda, N. & Kondo, A. Expression of varied GFPs in *Saccharomyces cerevisiae*: Codon optimization yields stronger than expected expression and fluorescence intensity. *Sci. Rep.* **6**, 1–15 (2016).
162. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
163. Lang, P. T. *et al.* DOCK 6: Combining techniques to model RNA-small molecule complexes. *Rna* **15**, 1219–1230 (2009).
164. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
165. Sagermann, M. & Matthews, B. W. Crystal structures of a T4-lysozyme duplication-extension mutant demonstrate that the highly conserved  $\beta$ -sheet region has low intrinsic folding propensity. *J. Mol. Biol.* **316**, 931–940 (2002).

166. Pronk, S. *et al.* Structural bioinformatics molecular simulation toolkit *J. Mol. Biol.* **29**, 845–854 (2013).
167. Buck, M., Bouguet-bonnet, S., Pastor, R. W. & Mackerell, A. D. Importance of the CMAP Correction to the CHARMM22 Protein Force Field : Dynamics of Hen Lysozyme. 36–38 (2006). doi:10.1529/biophysj.105.078154
168. Delano, W. L. & Ph, D. PyMOL : An Open-Source Molecular Graphics Tool.
169. Kumar, R., Singh, S. & Singh, O. V. Bioconversion of lignocellulosic biomass: Biochemical and molecular perspectives. *J. Ind. Microbiol. Biotechnol.* **35**, 377–391 (2008).
170. Wettstein, S. G., Martin Alonso, D., Gürbüz, E. I. & Dumesic, J. A. A roadmap for conversion of lignocellulosic biomass to chemicals and fuels. *Curr. Opin. Chem. Eng.* **1**, 218–224 (2012).
171. Zhou, C. H., Xia, X., Lin, C. X., Tong, D. S. & Beltramini, J. Catalytic conversion of lignocellulosic biomass to fine chemicals and fuels. *Chem. Soc. Rev.* **40**, 5588–5617 (2011).
172. 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).
173. Kuyper, M. *et al.* Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. *FEMS Yeast Res.* **5**, 925–934 (2005).
174. Jojima, T., Omumasaba, C. A., Inui, M. & Yukawa, H. Sugar transporters in efficient utilization of mixed sugar substrates: Current knowledge and outlook. *Appl. Microbiol. Biotechnol.* **85**, 471–480 (2010).
175. Matsushika, A. & Sawayama, S. Characterization of a Recombinant Flocculent *Saccharomyces cerevisiae* Strain that Co-ferments Glucose and Xylose: I. Influence of the Ratio of Glucose/Xylose on Ethanol Production. *Appl. Biochem. Biotechnol.* **169**, 712–721 (2013).
176. Stambuk, B. U., Eleutherio, E. C. A., Florez-Pardo, L. M., Souto-Maior, A. M. & Bon, E. P. S. Brazilian potential for biomass ethanol: Challenge of using hexose and pentose cofermenting yeast strains. *J. Sci. Ind. Res. (India)*. **67**, 918–926 (2008).
177. Gancedo, J. M. Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev.* **62**, 334–61 (1998).
178. Kayikci, Ö. & Nielsen, J. Glucose repression in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **15**, fov068 (2015).



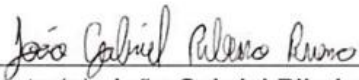
179. Boles, E. & Hollenberg, C. P. The molecular genetics of hexose transport in yeasts. *FEMS Microbiol. Rev.* **21**, 85–111 (1997).
180. Bueno, J. G. *et al.* Direct from nature - novel xylose transporters and molecular mechanisms insights for improved uptake on high sugar concentrations. *Pers. Commun.* (2020).
181. Nijland, J. G. *et al.* Improved Xylose Metabolism by a CYC8 Mutant of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **83**, 1–12 (2017).
182. Bao, Z. *et al.* Homology-Integrated CRISPR-Cas (HI-CRISPR) System for One-Step Multigene Disruption in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **4**, 585–594 (2015).
183. Byrne, K. P. & Wolfe, K. H. The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* **15**, 1456–1461 (2005).
184. Fleiss, A., O'Donnell, S., Agier, N., Delmas, S. & Fischer, G. Reshuffling yeast chromosomes with CRISPR/Cas9. *bioRxiv* 415349 (2018). doi:10.1101/415349
185. Lee, H. J., Kweon, J., Kim, E., Kim, S. & Kim, J. S. Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. *Genome Res.* **22**, 539–548 (2012).
186. Brunet, E. *et al.* Chromosomal translocations induced at specified loci in human stem cells. *Proc. Natl. Acad. Sci.* **106**, 10620–10625 (2009).
187. McCartney, R. R., Chandrashekarappa, D. G., Zhang, B. B. & Schmidt, M. C. Genetic analysis of resistance and sensitivity to 2-deoxyglucose in *Saccharomyces cerevisiae*. *Genetics* **198**, 635–646 (2014).
188. Loba, Z. & Maitra, P. K. Resistance to 2-deoxyglucose in yeast: A direct selection of mutants lacking glucose-phosphorylating enzymes. *MGG Mol. Gen. Genet.* **157**, 297–300 (1977).
189. Wijsman, M. *et al.* A toolkit for rapid CRISPR-SpCas9 assisted construction of hexose-transport-deficient *Saccharomyces cerevisiae* strains. *FEMS Yeast Res.* **19**, 1–12 (2019).
190. Biot-Pelletier, D. & Martin, V. J. J. Seamless site-directed mutagenesis of the *Saccharomyces cerevisiae* genome using CRISPR-Cas9. *J. Biol. Eng.* **10**, 1–5 (2016).
191. Ma, H., Bloom, L. M., Walsh, C. T. & Botstein, D. Isolation and Characterization of Mutations in the HXK2 Gene of *Saccharomyces cerevisiae*. **9**, 5630–5642 (1989).
192. DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M. & Church, G. M. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat. Biotechnol.* **33**, 1250–1255 (2015).


193. Dorfman, B. Z. The isolation of adenylosuccinate synthetase mutants in yeast by selection for constitutive behavior in pigmented strains. *Genetics* **61**, 377–89 (1969).
194. Jakočiūnas, T., Jensen, M. K. & Keasling, J. D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* **34**, 44–59 (2015).
195. Lee, H. & Kim, J. S. Unexpected CRISPR on-target effects. *Nat. Biotechnol.* **36**, 9–10 (2018).
196. Hu, Z., Nehlin, J. O., Ronne, H. & Michels, C. A. MIG1-dependent and MIG1-independent glucose regulation of MAL gene expression in *Saccharomyces cerevisiae*. *Curr. Genet.* **28**, 258–266 (1995).
197. Ozcan, S. & Johnston, M. Three different regulatory mechanisms enable yeast hexose transporter (HXT) genes to be induced by different levels of glucose. *Mol. Cell. Biol.* **15**, 1564–1572 (1995).
198. Roy, A., Kim, Y.-B., Cho, K. H. & Kim, J.-H. Glucose starvation-induced turnover of the yeast glucose transporter Hxt1. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 2878–2885 (2014).
199. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc / SS carrier DNA / PEG method. *Nat. Protoc.* **2**, 31–35 (2008).
200. Balakrishnan, R. *et al.* Fungal BLAST and Model Organism BLASTP Best Hits: New comparison resources at the *Saccharomyces* Genome Database (SGD). *Nucleic Acids Res.* **33**, 374–377 (2005).
201. Shin, H. Y., Nijland, J. G., de Waal, P. P. & Driessen, A. J. M. The amino-terminal tail of Hxt11 confers membrane stability to the Hxt2 sugar transporter and improves xylose fermentation in the presence of acetic acid. *Biotechnol. Bioeng.* **114**, 1937–1945 (2017).
202. Baudin, A., Ozier-kalogeropoulos, O., Denouel, A., Lacroute, F. & Cullin, C. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **21**, 3329–3330 (1993).

**Anexo I - Termo de autorização****Declaração**

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Engenharia genética de *Saccharomyces cerevisiae* para co-fermentação de xilose e glicose e produção de etanol de segunda geração**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 20 de fevereiro de 2020

Assinatura :   
Nome do(a) autor(a): **João Gabriel Ribeiro Bueno**  
RG n.º 41.729.460-8

Assinatura :   
Nome do(a) orientador(a): **Leandro Vieira dos Santos**  
RG n.º 30.805.457-x

## Anexo II - Declaração de Bioética/Biossegurança



Campinas, 14 de maio 2019

Para o Programa de pós-graduação em Genética e Biologia Molecular (PPG-GBM)

Assunto - Aprovação de projeto na CIBio - CNPEM

Prezado coordenador,

Declaramos para os devidos fins que o aluno João Gabriel Ribeiro Bueno desenvolveu projeto de pesquisa associado à sua tese, intitulada "Expressão heteróloga de putativos transportadores de xilose de *Candida sojae* em *Saccharomyces cerevisiae* visando a produção de etanol 2G", sob orientação do Dr. Leandro Vieira dos Santos. O projeto utilizou organismos geneticamente modificados classe de risco I, cumprindo os requisitos de biossegurança necessários. O projeto de mestrado integra o protocolo CIBio 2017-15, Engenharia metabólica e evolutiva de *Saccharomyces cerevisiae* para processos 1G e 2G, tendo sido APROVADO para execução no CNPEM.

Atenciosamente,

---

Leandro Vieira dos Santos, PhD  
Orientador, pesquisador responsável

---

Marcio Chaim Bajgelman, PhD  
Presidente da CIBio - CNPEM