

# UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA



**Silvia Kazue Missawa**

**“Modificação de linhagens industriais de *Saccharomyces cerevisiae*  
para o aumento da produtividade de álcool e floculação condicional”**

Este exemplar corresponde à redação final  
da tese defendida pelo(a) candidato (a)  
Silvia Kazue Missawa  
e aprovada pela Comissão julgadora.

Tese apresentada ao Instituto de Biologia para obtenção do Título de Doutor em Genética e Biologia Molecular, na área de genética de microorganismo.

Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira.

Co-Orientador: Prof. Dr. Anderson Ferreira da Cunha

Campinas, 2009

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP

<b>M691m</b>	<p>Missawa, Silvia Kazue Modificação de linhagens industriais de <i>Saccharomyces cerevisiae</i> para o aumento da produtividade de álcool e flocculação condicional / Silvia Kazue Missawa. – Campinas, SP: [s.n.], 2009.</p> <p>Orientadores: Gonçalo Amarante Guimarães Pereira, Anderson Ferreira da Cunha. Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.</p> <p>1. Fermentação. 2. <i>Saccharomyces cerevisiae</i>. 3. Álcool. 4. Genética molecular. 5. Linhagem industrial. I. Pereira, Gonçalo Amarante Guimarães. II. Cunha, Anderson Ferreira da. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.</p> <p>(rcdt/ib)</p>
--------------	--

**Título em inglês:** Genetic modification of industrial *Saccharomyces cerevisiae* strain to improve ethanol production and conditional flocculation.

**Palavras-chave em inglês:** Fermentation; *Saccharomyces cerevisiae*; Alcohol; Molecular genetics; Industrial strain.

**Área de concentração:** Genética de Microorganismos.

**Titulação:** Doutora em Genética e Biologia Molecular.

**Banca examinadora:** Gonçalo Amarante Guimarães Pereira, Jörg Kobarg, Flávio Henrique da Silva, Jorge Maurício Costa Mondego, André Ricardo Alcarde.

**Data da defesa:** 30/07/2009.

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

Campinas, 30 de julho de 2009.

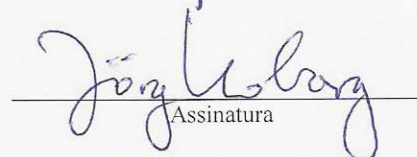
## BANCA EXAMINADORA

Prof. Dr. Gonçalo Amarante Guimarães Pereira  
(Orientador)



Assinatura

Prof. Dr. Jörg Kobarg



Assinatura

Prof. Dr. Flávio Henrique da Silva



Assinatura

Prof. Dr. Silvio Roberto Andrietta

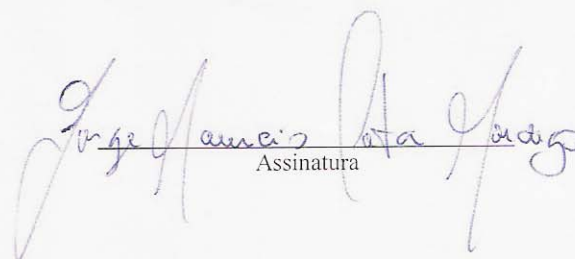


Assinatura

Prof. Dr. André Ricardo Alcarde

Assinatura

Dr. Jorge Maurício Costa Mondego



Assinatura

Prof. Dra. Sandra Regina Ceccato Antonini

Assinatura

Prof. Dra. Telma Teixeira Franco

Assinatura

*"Você nunca será a pessoa que pode ser se pressão,  
tensão e disciplina forem tiradas da sua vida"*  
(James G. Bilkey)

*Dedico aos meus pais  
e meu irmão*

## AGRADECIMENTOS

Ao meu orientador Gonçalo, pela oportunidade, confiança e apoio durante todo o meu doutorado;

Ao meu co-orientador Anderson, por me ensinar as técnicas básicas de biologia molecular. E por toda a ajuda nos experimentos, discussão de resultados, paciência e amizade;

À Eliane, pela amizade, paciência e disposição, sempre pronta a nos ajudar;

Ao Dr. Juan Lucas Argueso, pesquisador da Duke University Medical Center, USA. Pelas importantes sugestões e auxílio;

Ao Luiz Humberto Gomes, o Beto, do laboratório de Genética de Leveduras da Esalq – USP (Piracicaba), pelas dicas e ajuda essencial nos ensaios fermentativos;

A todos os amigos que fiz durante todo o período do meu Doutorado no LGE. Às amigas que me acompanharam desde o começo: Carla e Ana Paula. A todo pessoal do LGE, pela ajuda e amizade construída no laboratório;

Às minhas amigas de graduação, Patrícia, Maria Paula, Lye, com que sempre divido os momentos bons e ruins;

À minha família, sempre me apoiando em todos os momentos da minha vida;

À Pós Graduação do Instituto de Biologia da Unicamp;

À Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro.

## SUMÁRIO

<b>RESUMO.....</b>	<b>1</b>
<b>ABSTRACT .....</b>	<b>2</b>
<b>REVISÃO BIBLIOGRÁFICA .....</b>	<b>3</b>
<b>1. AQUECIMENTO GLOBAL E FONTES ALTERNATIVAS DE ENERGIA.....</b>	<b>3</b>
<b>2. PRODUÇÃO DE ETANOL NO BRASIL .....</b>	<b>7</b>
<b>3. PROCESSO FERMENTATIVO.....</b>	<b>9</b>
<b>4. A LEVEDURA.....</b>	<b>11</b>
<b>5. PROCESSO DE FLOCULAÇÃO .....</b>	<b>14</b>
<b>6. ATIVIDADE MITOCONDRIAL.....</b>	<b>21</b>
<b>JUSTIFICATIVA .....</b>	<b>23</b>
<b>OBJETIVOS .....</b>	<b>24</b>
<b>CAPITULO 1.....</b>	<b>25</b>
<b>CAPITULO 2.....</b>	<b>35</b>
<b>CAPITULO 3.....</b>	<b>54</b>
<b>CONCLUSÕES FINAIS .....</b>	<b>70</b>
<b>ANEXO I.....</b>	<b>78</b>
<b>ANEXO II.....</b>	<b>104</b>

## RESUMO

A crescente necessidade da substituição de fontes de energia derivadas do petróleo por fontes renováveis de energia impulsionou pesquisas em todo o mundo. O etanol se enquadra fortemente como candidato para substituir combustíveis como, por exemplo, a gasolina. A produção de etanol brasileiro ocorre pela fermentação de açúcares proveniente do caldo de cana, realizada por leveduras, fundamentalmente da espécie *Saccharomyces cerevisiae*. No processo fermentativo para a produção de etanol combustível no Brasil, a separação das leveduras após o esgotamento do açúcar é realizado por centrífugas de alta capacidade, o que torna o sistema complexo e dispendioso. Neste trabalho, linhagens de leveduras de laboratório e industriais foram transformadas por engenharia genética de modo que as leveduras reconhecessem o esgotamento da glicose do meio separando-se naturalmente e rapidamente. Para isso, plasmídeos e posteriormente, cassetes integrativos, foram construídos utilizando-se promotores, sensíveis a presença de glicose no meio, regulando a expressão de genes da floculação. Estes promovem a formação de agregados celulares que se separam do meio por decantação. Além disso, está relatado na literatura que mutantes respiratórios produzem uma maior quantidade de etanol. Então foi feita a deleção, em leveduras industriais, do gene *PET191*, cujo produto participa na montagem do citocromo c, que por sua vez é essencial para o funcionamento da cadeia respiratória nestes microorganismos. As leveduras deletadas apresentaram um aumento no rendimento da produção de etanol; na linhagem industrial JAY292, esse aumento foi de 4%. Isto pode refletir ao final de um ano, um aumento de 980 milhões de litros. Os resultados obtidos contribuem para o desenvolvimento tecnológico e industrial desta área estratégica para o país.

## **ABSTRACT**

The increased need for replacing energy sources from fossil oil to renewable sources has prompted research initiatives throughout the world. Bioethanol is a highly competitive candidate to replace fuels such as gasoline. The production of Brazilian ethanol occurs by the fermentation of sugars from sugar cane juice by the yeast *Saccharomyces cerevisiae*. The industrial fermentation process used to produce ethanol fuel in Brazil includes a separation step by high-capacity centrifugation of the yeast cells after the exhaustion of the sugar, which makes the system complex and costly. In this work, laboratory and industrial yeast strains were transformed by genetic engineering techniques so that they are able to recognize the depletion of glucose from the medium and quickly separate by sedimentation, without the need of centrifugation. In order to reach this goal, plasmids and integrative cassettes were constructed using promoters sensitive to the presence of glucose in the environment regulating the expression of the genes that control flocculation. This genetic device promotes the formation of cellular aggregates that are separated from the medium by sedimentation only when glucose is depleted. Furthermore, it is reported in the literature that respiratory mutants produce a higher quantity of ethanol. Consequently, the *PET191* gene was deleted in an industrial yeast strain; the gene product *pet191* participates in the assembly of cytochrome c, which is an essential component of the respiratory chain in these microorganisms. The deleted industrial strain JAY292 showed an increased production of ethanol by 4%. This percentage can be translated in an increase of more than 980 million liters of ethanol produced per year. These results are of great technological and industrial importance, and may contribute to the development of this strategic area for the country.



## REVISÃO BIBLIOGRÁFICA

### 1. Aquecimento global e fontes alternativas de energia

O aquecimento global é o aumento da temperatura média da superfície terrestre provocado principalmente pelo efeito estufa. Desde o início do século 19, a temperatura aumentou cerca de 0,4 a 0,8 °C. A estimativa, que preocupa cientistas de todo o mundo, é que até 2100 a temperatura tenha um aumento de 1,4 a 5,8 °C, (Mastrandrea & Stephen, 2005).

Os maiores responsáveis pelo aumento do efeito estufa na Terra são as atividades humanas, principalmente o desmatamento e a queima de combustíveis fósseis (provocada por veículos automotores e indústrias). O excesso de gases que são liberados na atmosfera, principalmente o dióxido de carbono (CO<sub>2</sub>), potencializam a ação do efeito estufa. Uma opção para minimizar isso é a substituição de combustíveis fósseis por outras fontes de energia (Goldemberg *et al.*, 2008).

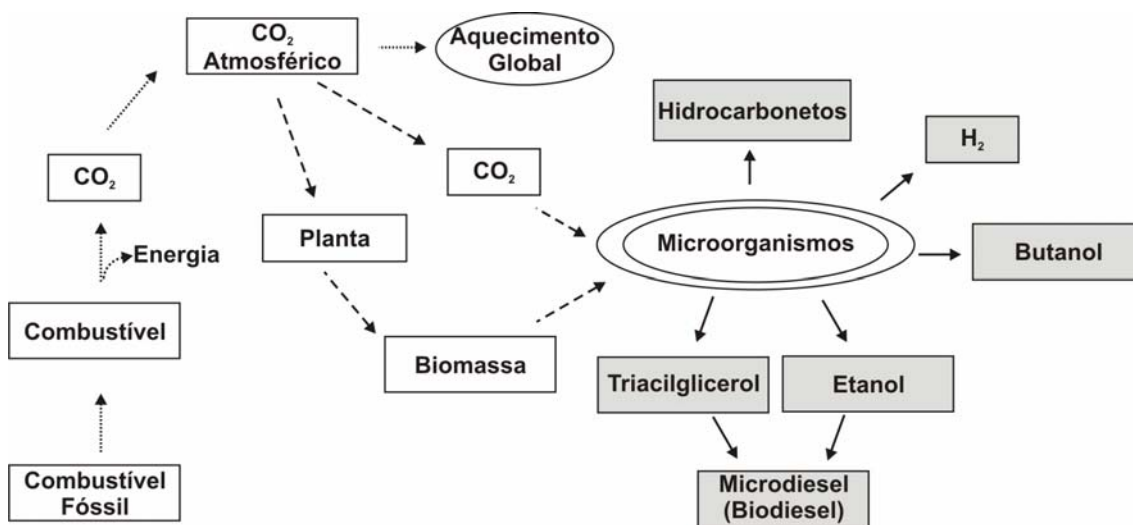
A busca por energia alternativa renovável e menos poluente, visando substituir combustíveis derivados do petróleo, tem sido tema de discussões e pesquisas em todo o mundo (Farrell *et al.*, 2006).

Em 1997, durante uma reunião realizada no Japão, com a presença de delegações de mais de 160 países, foi estabelecido um acordo, conhecido como Protocolo de Kyoto. O objetivo foi a restrição da emissão de CO<sub>2</sub> e outros 5 gases responsáveis pelo efeito estufa. A meta estabelecida é a redução, até 2012, a níveis 5% inferiores aos de 1990, principalmente pela substituição de combustíveis fósseis por fontes alternativas de energia (Mastrandrea & Stephen, 2005).

Uma nova reunião será realizada em Copenhague no final de 2009, a Conferência das Nações Unidas sobre Mudanças Climáticas (COP15). O governo dinamarquês tem uma proposta ambiciosa de estabelecer um acordo global incorporando todos os países do mundo. Serão estabelecidas as metas de redução de emissões de gases para as próximas décadas, reforçando as estabelecidas no protocolo de Kyoto e ainda lembrando o fato de que combustíveis fósseis não são apenas poluentes, mas fontes que se esgotam a cada ano

(<http://en.cop15.dk/about+cop15>). Estima-se que o petróleo se esgote nos próximos 50 anos, o gás natural em 65 anos e o carvão em 200 anos (Branduardi *et al.*, 2008).

Nos últimos anos, aumentaram os investimentos das nações em pesquisas e no desenvolvimento de novas tecnologias para a obtenção de fontes alternativas e renováveis de energia, como por exemplo: bioetanol, biodiesel, 1-butanol, biogás. Além disso, microorganismos estão sendo estudados, de forma que eles utilizem a biomassa para sintetizar compostos que substituam produtos derivados de petróleo (Figura 1) (Rojo, 2008).



Referência: Adaptado de Rojo (2008).

**FIGURA 1:** Biocombustíveis gerados por microorganismos. A queima de combustíveis fósseis libera CO<sub>2</sub> na atmosfera, contribuindo para o aumento da concentração desse gás gerando o aquecimento global (linhas pontilhadas). Parte desse gás pode ser fixada pelas plantas ou por microorganismos fotossintéticos (linhas tracejadas). Alguns microorganismos utilizam o CO<sub>2</sub> ou biomassa como fonte de carbono produzindo componentes que podem ser utilizados como combustível (linhas contínuas).

O desenvolvimento de biocombustíveis derivados de biomassa requer integração de conhecimentos de biologia, química e engenharia. E o mais importante para a indústria é que esses novos combustíveis apresentem características como: (1) ser relativamente barato - a molécula deve ter alta razão energia/massa; (2) ser altamente combustível, mas não explosivo; (3) ser estável sob longos períodos de armazenamento e (4) ser líquido e transportável (Wackett, 2008).

Atualmente, o biocombustível mais utilizado e produzido em grande escala é o etanol. O uso do etanol como combustível no lugar de derivados do petróleo diminui significativamente a taxa de emissão de monóxido de carbono (CO), hidrocarbonetos e enxofre na atmosfera, sendo, portanto menos tóxico (Goldemberg, 2008).

O Brasil é o segundo maior produtor mundial de etanol, sendo superado apenas pelos EUA, e juntos representam aproximadamente 90% da produção mundial (<http://www.ethanolrfa.org/industry/statistics/>).

O etanol brasileiro é produzido pela fermentação dos açúcares presentes na cana-de-açúcar enquanto que o americano é produzido pela fermentação de açúcares (amido - previamente hidrolisado enzimaticamente) do milho. Há claras vantagens na produção de etanol a partir do caldo de cana, como pode ser visto na tabela 1.

**TABELA 1:** Comparação da produção de etanol no Brasil e nos EUA (Sanchez & Cardona, 2008).

	<b>Brasil</b>	<b>EUA</b>
Matéria Prima	Cana-de-açúcar	Milho
Produtividade (ton/ha)	70 - 123	6 - 10
Custo da matéria prima (US\$/Kg)	0,010	0,076
Rendimento etanol (L/ton)	70	370
Rendimento anual etanol (L/ha ano)	5345 - 9381	6600
Balanco Energético (razão saída/entrada de energia)	8,0	1,34 - 1,53
Custos de produção*	14,48	39,47
Subsidio governo*	0	7,93
Venda subprodutos*	0	6,71
Custo líquido da produção*	14,48	24,83

\* Referência: (Goldemberg, 2008). Valores (€/1000L).

O rendimento de etanol (L/ton) proveniente do milho é maior do que a cana-de-açúcar pelo fato de que há uma grande quantidade de açúcares fermentáveis (glicose) que é liberada do material com amido original. Entretanto, a competitividade do etanol de cana do Brasil pode ser demonstrada pelo balanço energético que é aproximadamente 6 vezes

maior que no etanol de milho. O balanço energético é definido como a razão entre a energia contida em um determinado volume de etanol dividido pela energia fóssil gasta para sua produção (uso de fertilizantes, pesticidas, óleo diesel gasto na colheita mecanizada e o transporte de matéria-prima para os locais de processamento). Além disso, o custo de produção de etanol no Brasil é 2,7 vezes menor que nos EUA (Tabela1) (Goldemberg, 2008; Sanchez & Cardona, 2008).

O Brasil é um dos maiores produtores de cana-de-açúcar e álcool do mundo, e esse setor canavieiro, além de estar se transformando no negócio mais produtivo da agroindústria brasileira, é responsável por mais de 1 milhão de trabalhos diretos (Pessoa-Jr *et al.*, 2005). O pioneirismo no setor alcooleiro se iniciou em meados da década de 70, após o primeiro choque do petróleo com a implantação do Proálcool, programa de incentivo a produção de álcool visando o atendimento das necessidades do mercado interno e externo e da política de combustíveis automotivos (Goldemberg *et al.*, 2008).

Atualmente há em operação no Brasil cerca de 330 usinas, processando 425 milhões de toneladas de cana-de-açúcar por ano, sendo que aproximadamente metade é utilizada para a produção de açúcar e metade para a produção de etanol. A área utilizada para o cultivo de cana ocupa menos de 10% da área cultivável do país. O Programa do etanol no Brasil está muito bem estabelecido e tem substituído 1,5% da gasolina utilizada no mundo. A estimativa é que nos próximos 15 - 20 anos seja possível substituir mais de 10% da gasolina mundial (Goldemberg, 2008).

Apesar da supremacia brasileira no mercado de álcool, algumas deficiências precisam ser superadas, como a falta de investimentos em ciência e tecnologia e infraestrutura. Esses investimentos são importantes para conseguir manter a qualidade e os custos competitivos (Salomão & Onaga, 2006). Além disso, a biorefinaria (local do processamento de biomassa), não é local para a fabricação de apenas um ou dois produtos, mas uma ampla gama de produtos de maior valor agregado. E isso ainda é deficiente na indústria canavieira (Cortez *et al.*, 2008).

Recentemente, o governo americano decidiu investir na expansão da produção de etanol no mundo sendo o Brasil o principal foco neste acordo. Expectativas são de que a

produção nacional de etanol aumente de 17,4 bilhões de litros produzidos em 2006 para 35,4 bilhões de litros até 2012 (Jank, 2007).

## **2. Produção de etanol no Brasil**

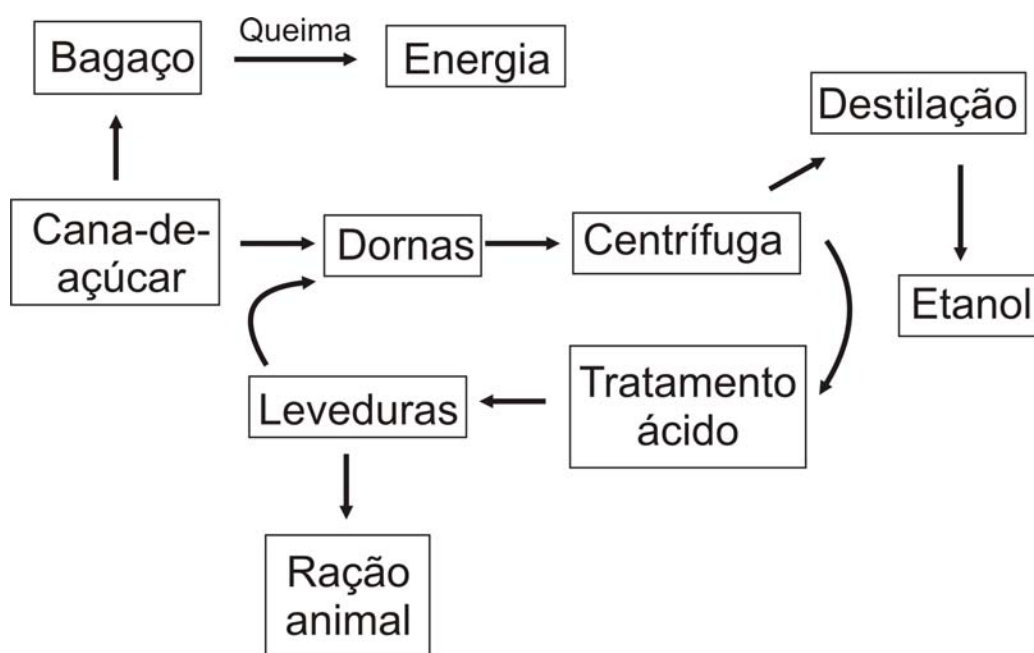
A obtenção do etanol se dá principalmente pela via fermentativa (Wheals *et al.*, 1999), sendo que este é um processo utilizado há séculos para a produção de bebidas, tais como vinhos e cervejas (Stratford, 1996b). O que torna a fermentação a forma mais importante de obtenção de etanol é a grande quantidade de matérias-primas naturais existentes em todo o país, sendo que a cana de açúcar é, sem dúvida, a mais importante (Lima *et al.*, 1975).

Para uma produção industrial de etanol viável, o processo de fermentação deve ser extremamente robusto e pouco afetado por pequenas alterações da matéria prima, além da minimização dos custos que podem ser obtidos pela (Goldemberg *et al.*, 2008):

- obtenção do máximo rendimento de etanol;
- minimização na síntese de outros produtos (glicerol, ácido láctico);
- minimização do tempo de fermentação;
- baixa contaminação bacteriana (ações preventivas, como por exemplo, tratamento ácido do fermento);
- manutenção da alta viabilidade do fermento (70% - 80%);
- minimização de produtos químicos (ácidos, antibióticos, antiespumante);
- minimização dos gastos com manutenção (limpeza, vazamentos);
- minimização de gastos de energia e água;
- automatização das operações manuais (temperatura, brix, pH).

O processo fermentativo de produção industrial de etanol brasileiro (Figura 2) ocorre em dornas nas quais são adicionadas a mistura de mosto (caldo de cana-de-açúcar ou melaço de cana diluído ou a mistura dos dois) e as leveduras. Após o término da fermentação, todo o conteúdo é transferido para centrífugas. O vinho, fermentado de cana sem as leveduras, segue para as torres de destilação, onde se recupera o etanol e o leite de levedura, nome dado ao caldo de leveduras que sai da centrifugação, passa por um

tratamento com ácido sulfúrico e retorna novamente à dorna, na qual será adicionada uma nova carga de mosto para um novo ciclo fermentativo (Wheals *et al.*, 1999; Meleiro & Maciel, 2000). Cerca de 90% das leveduras são reaproveitadas de uma fermentação para outra e esse processo de reciclo permanece durante todo o período da safra que compreende 200 a 250 dias, ocorrendo entre dois a três ciclos fermentativos por dia (Meleiro & Maciel, 2000; Basso *et al.*, 2008). Além disso, todo o resíduo gerado na usina é reaproveitado. As leveduras, por exemplo, são utilizadas como ração animal após o período da safra; e o bagaço da cana-de-açúcar é queimado para geração de energia, que é suficiente para suprir as necessidades de toda usina, e o excedente é encaminhado para ser vendido.



**FIGURA 2:** Esquema do processo, tipo batelada, de produção de etanol no Brasil. Importante destacar também, o reciclo das leveduras que ocorre durante toda a safra de cana-de-açúcar e o reaproveitamento do bagaço para a produção de energia.

Na produção de bebidas, como por exemplo, o vinho, os açúcares fermentados são a frutose e a glicose, obtida do suco de uva (mosto). Este suco é transferido para tanques de fermentação onde recebem as leveduras. Depois de estabelecido o término da fermentação, não há etapa de centrifugação. É necessário esperar que a levedura decante naturalmente e posteriormente ela é retirada do tanque podendo ser utilizada na produção de ração animal

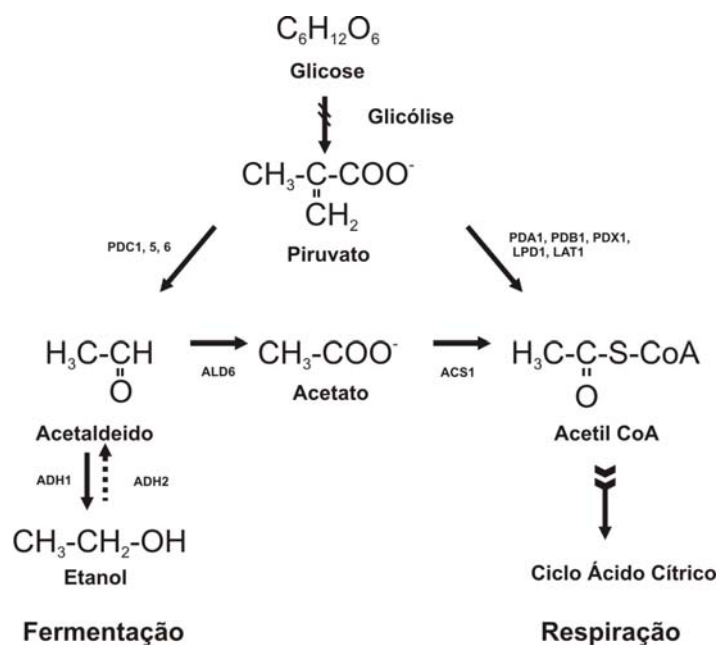
ou adubo. Tempos muito longos de decantação levam ao rompimento das paredes celulares das leveduras que ainda estiverem em suspensão, liberando compostos secundários no meio, alterando o sabor e a qualidade do vinho. Separada a borra, o vinho é enviado para o processo de clarificação através de uma série de filtrações, sendo posteriormente engarrafado e envelhecido (Martinelli, 1983).

Na produção de etanol e bebidas fermentadas (vinhos e cervejas), um fenômeno conhecido como floculação das leveduras desempenha um importante papel, pois acelera a separação das leveduras do meio fermentado. Esse é um processo de formação de agregados multicelulares, reversível (os agregados não são permanentes) e dependente de cálcio. Os agregados celulares formados podem se separar do meio de cultura por sedimentação, quando elas vão para o fundo da dorna, ou flotação, quando o agregado acumula ar levando as células à superfície (Stratford, 1996b; Stan & Despa, 2000).

### **3. Processo fermentativo**

Os microorganismos mais comumente usados na fermentação alcoólica são as leveduras, fundamentalmente a espécie *Saccharomyces cerevisiae*.

Bioquimicamente, a fermentação é a oxidação incompleta do açúcar, gerando como subproduto um composto orgânico oxidável. Considerando-se o processamento da cana pela levedura, inicialmente a sacarose, que é o seu açúcar de reserva, sofre hidrólise pela enzima invertase (*SUC2*) sendo convertida em glicose e frutose. Ambas entram na via glicolítica e, através de uma seqüência de reações, são convertidas a piruvato. Este primeiramente é descarboxilado pela enzima piruvato descarboxilase (*PDC1*, 5, 6), formando acetaldeído e liberando CO<sub>2</sub>. Posteriormente o acetaldeído é reduzido a etanol, sendo essa reação catalisada pela álcool desidrogenase 1 (*ADH1*). O esgotamento da glicose ativa, nas leveduras, a transcrição do gene da álcool desidrogenase 2 (*ADH2*). A enzima *Adh2* oxida o etanol a acetaldeído, o qual posteriormente é convertido a acetato, passando então para as etapas posteriores do ciclo de Krebs (Devlin, 1997) (Figura 3).



**FIGURA 3:** Esquema das reações de fermentação e respiração. As setas pontilhadas indicam a via que é ativada quando a glicose é totalmente consumida (oxidação etanol). *PDC1, 5, 6*: piruvato descarboxilase; *PDA1, PDB1, PDX1, LPD1, LAT1*: enzimas que fazem parte do complexo da piruvato desidrogenase; *ALD6*: aldeído desidrogenase; *ACS1*: acetil CoA sintetase; *ADH1*: álcool desidrogenase 1; *ADH2*: álcool desidrogenase 2.

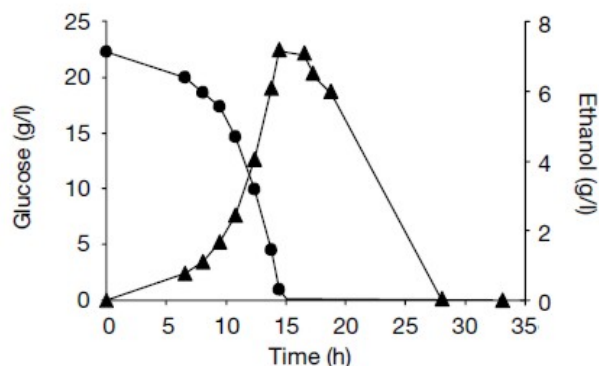
O *ADH2* é um gene muito bem estudado, cujo promotor é reprimido centenas de vezes na presença de glicose e é dereprimido na ausência de glicose (Price *et al.*, 1990). O gene *ADR1* foi originalmente identificado em *S. cerevisiae* como regulador positivo do gene *ADH2* e apresenta sua transcrição inibida na presença de glicose (Ciriacy, 1979; Denis & Young, 1983). A ausência ou a baixa ligação desse ativador transcricional (*Adr1*) ao promotor do gene *ADH2* faz com que a expressão deste seja mínima. Quando a glicose se esgota, o nível de *Adr1* aumenta, se ligando ao promotor do gene *ADH2*, iniciando a transcrição (Lee & DaSilva, 2005).

As leveduras quando são inoculadas em meio rico em açúcar, preferencialmente, seguem a via fermentativa produzindo etanol. Quando os açúcares fermentáveis se esgotam, as leveduras começam a consumir o etanol como fonte de carbono para o crescimento aeróbico. Essa mudança de crescimento anaeróbico para respiração aeróbica devido ao esgotamento da glicose, referida como *diauxic shift* (Figura 4), é correlacionada com alterações na expressão de genes envolvidos em processos celulares fundamentais



como metabolismo de carbono, síntese de proteínas e armazenamento de carboidratos (DeRisi *et al.*, 1997).

O *diauxic shift* é um ponto importante que deve ser considerado durante o processo fermentativo industrial. Após o término da fermentação (esgotamento da glicose), o ideal seria que o fermento (leveduras) seja separado imediatamente do meio para evitar que as células comecem a consumir o etanol, o que acarretaria em perdas de rendimento.



Referência: (Otterstedt *et al.*, 2004)

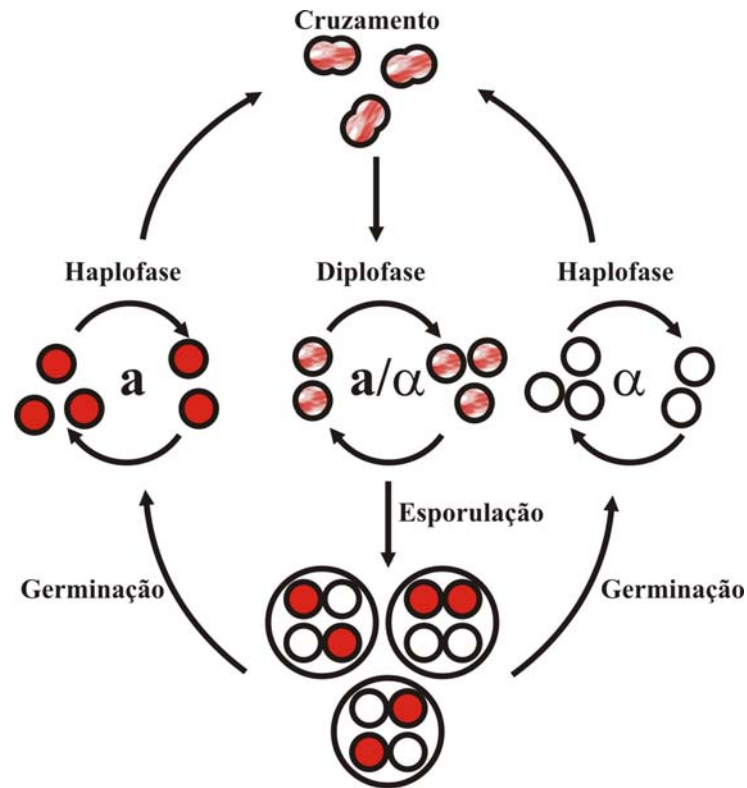
**FIGURA 4:** Gráfico mostrando a curva de consumo de glicose (linha com círculos) e etanol produzido (linha com os triângulos) pelas leveduras. Em 15 horas é ponto em que ocorre o *Diauxic Shift*, ou seja, a transição fermentação/respiração. Depois desse tempo o etanol começa a ser consumido.

#### 4. A levedura

*Saccharomyces cerevisiae* é um fungo leveduriforme, unicelular, tipicamente esférico ou oval, não filamentosos, pertencente ao filo Ascomycota. Possui reprodução assexuada, por brotamento, e reprodução sexuada (Tortora *et al.*, 2002). Na reprodução sexuada, o cruzamento ocorre somente entre indivíduos de tipos sexuais diferentes (*mating type a* ou  $\alpha$ ) (Figura 5).

*S. cerevisiae* foi o primeiro organismo eucarionte a ter o seu genoma totalmente sequenciado (Goffeau *et al.*, 1996). Possui inúmeros bancos de dados onde é possível obter informações sobre seus genes e produtos ([www.yeastgenome.org](http://www.yeastgenome.org)). É considerado o organismo modelo entre os eucariotos por apresentar mecanismos celulares altamente conservados. Por esse motivo há uma série de estudos sobre processos biológicos como envelhecimento e ciclo celular, estudos de doenças humanas como o câncer, e é usada como ferramenta em estudo de drogas, príons, pesquisa básica e aplicada de vírus e ecotoxicologia (Nevoigt, 2008). É popularmente conhecida na indústria alimentícia e de

bebidas, sendo classificada como GRAS (Generally Regarded as Safe) pela FDA (U.S. Food and Drug Administration).



**FIGURA 5:** Esquema das formas de reprodução das leveduras. A forma assexuada (na haplofase), em que uma levedura dá origem a outra por brotamento (mitose). E a reprodução sexuada, na qual um esporo *a* cruza com um esporo  $\alpha$ , formando um diplóide. A esporulação desse diplóide origina uma tétrade que dá origem a 4 esporos.

O grande sucesso da *S. cerevisiae* na indústria de biotecnologia é devido à alta capacidade de produzir etanol e dióxido de carbono a partir de açúcares. Possui importantes características, como: tolerância a baixos valores de pH e alta concentração de açúcar e etanol, propriedades como alta competitividade perante contaminação (bacteriana ou por outras leveduras) na fermentação industrial e alta resistência a inibidores presentes na biomassa hidrolisada e capacidade de crescer anaerobicamente (Branduardi *et al.*, 2008; Nevoigt, 2008).

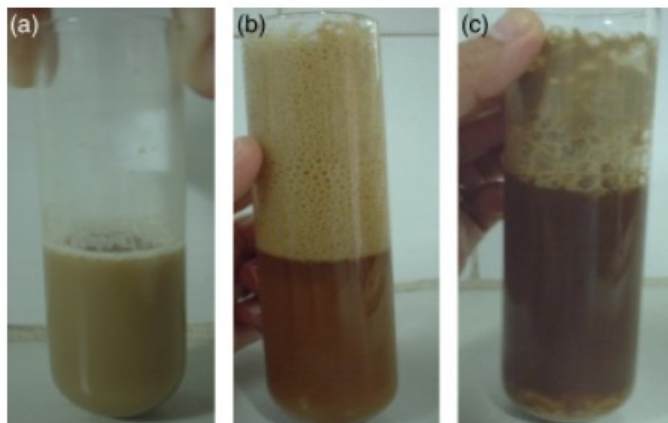
Apesar dos grandes avanços tecnológicos e do vasto conhecimento a respeito do potencial produtivo das leveduras, há ainda muitos aspectos desconhecidos, principalmente

quando se trata da genética e bioquímica das leveduras utilizadas nas usinas para fermentação alcoólica (Basso *et al.*, 2008).

Estudos mais recentes são relativos a análise genômica comparativa. As cepas mais estudadas são linhagens industriais fermentadoras principalmente para a produção de vinho (Borneman *et al.*, 2008; Carreto *et al.*, 2008). Muito pouco se conhece sobre as linhagens produtoras de etanol, especialmente quando se trata de caracterização fenotípica e genotípica e expressão gênica. Estas linhagens precisam ser exploradas e estudadas mais detalhadamente para o entendimento de características e fenômenos que podem ser essenciais para o uso biotecnológico.

Foi verificado em algumas usinas brasileiras, que as “leveduras de pão” (baker’s yeast) que eram inoculadas para iniciar o processo fermentativo, desapareciam do processo, sendo substituídas por outras leveduras durante o período da safra. Isso ocorria devido às condições estressantes e competitivas presentes na fermentação industrial: alta concentração de etanol, alta temperatura, estresse osmótico devido a presença de açúcares e sais, acidez e contaminação bacteriana que favoreciam o desenvolvimento de cepas mais resistentes, que no caso vinham junto com a cana-de-açúcar do campo (Basso *et al.*, 2008).

Por isso, leveduras isoladas do próprio processo fermentativo são hoje utilizadas nas usinas brasileiras. A linhagem mais conhecida e utilizada em aproximadamente 30% das usinas brasileiras é a Pedra 2 (PE-2), que leva esse nome pois foi isolada da Usina da Pedra. Estudos demonstraram que ela permanece na dorna durante todo o período da safra. Além disso, a linhagem PE-2 apresenta características fundamentais para aperfeiçoar o processo fermentativo, como permanecer em suspensão, não apresentando características floculantes e não produzindo espuma (Figura 6) (Basso *et al.*, 2008).



Referência: (Basso *et al.*, 2008)

**FIGURA 6:** Meio industrial (caldo de cana) fermentado por: (a) uma linhagem de levedura selecionada, PE-2; (b) uma linhagem que produz espuma; e (C) uma linhagem que produz espuma e ainda flocula. A presença de células floculadas resulta no baixo rendimento fermentativo.

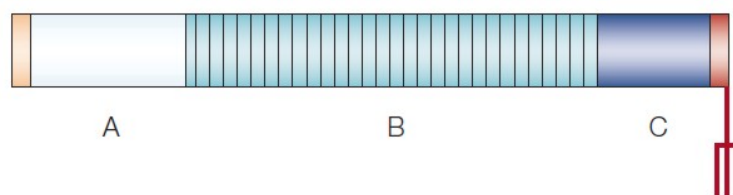
Modificações genéticas são bastante comuns em leveduras, principalmente nas linhagens de laboratório, que são de fácil manipulação e com marcas auxotróficas. Entretanto, para se melhorar o processo fermentativo, as alterações genéticas devem ser feitas em linhagens industriais, que permaneçam e dominem as dornas.

O conhecimento dos fatores responsáveis pela alta eficiência fermentativa e robustez das leveduras industriais é essencial para se estabelecer uma estratégia para melhorar ainda mais estas linhagens.

## 5. Processo de floculação

A floculação é uma das características mais intrigantes e importantes industrialmente manifestadas pelas leveduras. O fenômeno é dependente da linhagem, dificultando o controle do processo (Verstrepen *et al.*, 2003). Esse fenômeno causa grandes prejuízos nas usinas produtoras de etanol brasileiras, uma vez que a floculação se inicia repentinamente durante o processo, sem uma causa aparente conhecida. A floculação aumenta o tempo de fermentação, pois diminui a área de contato célula-meio, além disso, o processo para promover a desfloculação é difícil, sendo que muitas vezes as usinas são obrigadas a trocar o fermento.

Os fenótipos de adesão dependem de uma família de proteínas de superfície celular, também chamada de adesinas ou floculinas que são codificados pelos genes *FLO* na *S. cerevisiae* (Verstrepen & Klis, 2006). As proteínas dessa família consistem de 3 domínios (Figura 7). O domínio N-terminal (A) estaria relacionado com a afinidade dessa proteína à superfície; o domínio central (B) é uma região que apresenta um tamanho variável, é rico em resíduos de serina e treonina e contém muitas repetições em tandem; e o domínio carboxi-terminal (C) que é o local onde se liga covalentemente, a âncora glicosil fosfatidilinositol (GPI), responsável por prender a proteína à membrana. Esta ligação com o GPI ocorre durante as modificações pós-traducional (Verstrepen *et al.*, 2004).



Referência: (Verstrepen *et al.*, 2004)

**FIGURA 7:** Domínios presentes na adesina ou floculina. (A) Domínio N-terminal que está relacionado com a adesão. (B) Domínio central, região rica em serina/treonina que é codificada por sequências de nucleotídeos repetitivos. (C) Domínio carboxi-terminal, contém uma região de ligação covalente ao GPI (mostrado em vermelho).

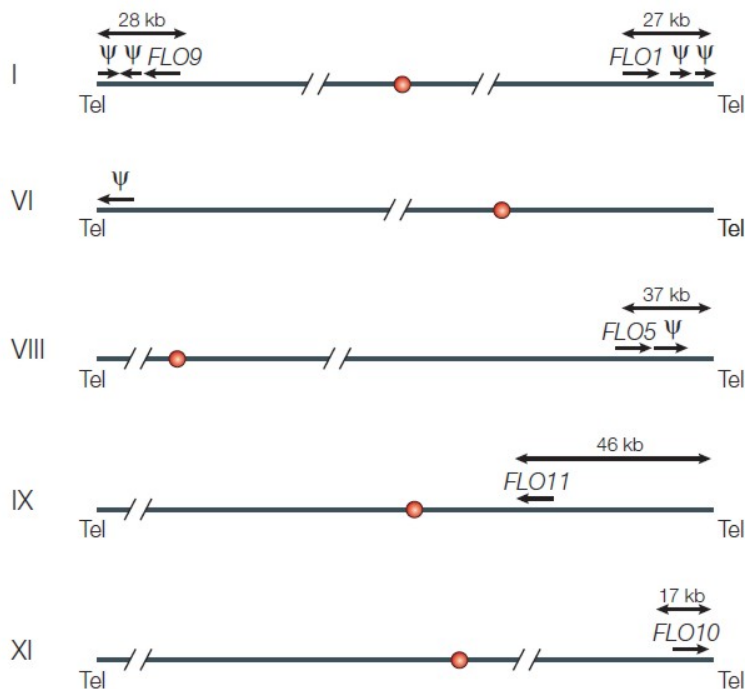
As proteínas Flo conferem adesão por dois mecanismos distintos. A primeira é a adesão célula-superfície, que se acredita depender de interações hidrofóbicas entre certos domínios na adesina e a superfície de ligação. O gene *FLO11* é o único da família que não é responsável pelo fenômeno da floculação em si, mas pelo crescimento invasivo, formação de pseudohifas, e aderência à superfície. (Van Mulders *et al.*, 2009).

O segundo mecanismo é a adesão célula-célula ou também chamado de floculação. Esse é um mecanismo físico, reversível, dependente de íons cálcio e foi primeiramente denominado *Lectin-like theory of flocculation* (Miki *et al.*, 1982). As proteínas de membrana das leveduras floculantes funcionam como lectinas e se ligam a carboidratos na parede celular das células vizinhas (Eddy & Rudin, 1958). As células floculantes possuem em sua membrana celular floculinas (lectinas) e carboidratos, enquanto que as células não floculantes possuem apenas carboidratos (Stratford, 1996a). Além do mecanismo físico de

interação entre as células de levedura, a floculação é influenciada por fatores genéticos de cada linhagem e por fatores externos, como por exemplo, concentração de íons cálcio, etanol, nutrientes, pH, temperatura (Touhami *et al.*, 2003; Verstrepen *et al.*, 2003).

A floculação em *S. cerevisiae* é controlada geneticamente e vários genes foram relatados como sendo dominantes e responsáveis por este fenômeno. Entre eles estão: *FLO1*, *FLO5*, *FLO9* e *FLO10* (Johnston & Reader, 1983; Teunissen & Steensma, 1995).

O genoma de *S. cerevisiae* mostrou que dos 5 genes *FLO* relacionados com a família das adesinas (*FLO1*, *FLO5*, *FLO9*, *FLO10*, *FLO11*) apenas 4 deles (*FLO1*, *FLO5*, *FLO9*, *FLO10*) estão localizados adjacentes ao telômero (Figura 8) (Verstrepen *et al.*, 2004). Os genes *FLO* subtelméricos, na linhagem de laboratório (S288c) de *S. cerevisiae*, são transcricionalmente silenciados, mas suas funções podem ser avaliadas substituindo o promotor nativo por um promotor induzível, como por exemplo, promotor de *GALI* (induzido por galactose) (Guo *et al.*, 2000).



Referência: (Verstrepen *et al.*, 2004)

**FIGURA 8:** Localização cromossômica dos genes e pseudogenes da adesina na linhagem S288c de *S. cerevisiae*. Os números à esquerda indicam os cromossomos das leveduras, os pontos em vermelho são os centrômeros e Ψ indica o local dos pseudogenes *FLO*. O gene *FLO11* é o único que não é nem centromérico, nem telomérico.

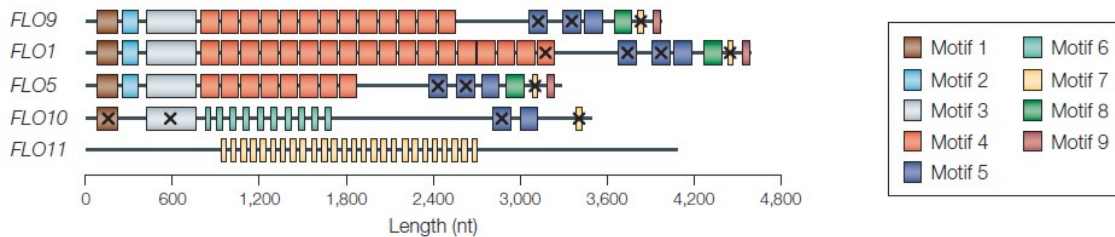
Com a clonagem do gene *FLO1* (Watari *et al.*, 1989) e do gene *FLO5* (Bidard *et al.*, 1994) ficou evidenciado uma semelhança muito grande entre eles. O gene *FLO1*, localizado no cromossomo I, possui o produto de seu gene 96% similar ao produto do gene *FLO5* (cromossomo VIII) e 94% similar ao produto do gene *FLO9* (cromossomo I). Já o produto do gene *FLO10* (cromossomo XI) apresenta 58% de similaridade com o gene *FLO1*. Apesar dessa baixa similaridade, quando se alinhou a região N-terminal da proteína referente ao gene *FLO10* com as proteínas dos genes *FLO1*, *FLO5* e *FLO9* obtiveram-se uma similaridade de 82%. Essa informação foi importante para classificá-lo como um gene de floculação (Figura 9) (Teunissen & Steensma, 1995).

		*				* *		*																																							
Flo1	200	L	P	N	I	E	G	T	V	Y		M	Y	A	G	Y	Y	P	M	K		V	V	Y	S	N	A	V	S	W	G		T	L	P	I	S	V	T	L	P	D		G	240		
Flo5	200	P	P	D	N	I	T	G	T	V	Y		M	Y	A	G	F	Y	P	M	K		I	V	Y	S	N	A	V	A	W	G		T	L	P	I	S	V	T	L	P	D		G	240	
Flo9	200	L	P	D	N	I	T	G	T	V	Y		M	Y	A	G	Y	Y	P	L	K		V	V	Y	S	N	A	V	S	W	G		T	L	P	I	S	V	E	L	P	D		G	240	
Flo10	222	A	P	T	D	I	K	G	S	T	Y		M	Y	A	G	Y	Y	P	I	K		I	V	Y	S	N	A	V	S	W	G		T	L	P	V	S	V	V	L	P	D		G	262	
Flo11	140	S	C	Q	V	W	M	P	N	F	Q		I	Q	F	E	Y	L	Q	G	S	A		A	Q	Y	A	S	S	W	Q	W	G		T	T	S	F	D	L	S	T	G	C		N	180

Referência: (Van Mulders *et al.*, 2009)

**FIGURA 9:** Alinhamento múltiplo da sequência de aminoácidos da região N-terminal das adesinas envolvidas no reconhecimento de açúcares. Os resíduos de aminoácidos não conservados estão sombreados de cinza. Os aminoácidos marcados com um (X) contribuem para o reconhecimento de açúcares (Kobayashi *et al.*, 1998).

Todos os quatro genes dominantes responsáveis pelo efeito da floculação, *FLO1*, *FLO5*, *FLO9* e *FLO10*, possuem repetições intragênicas em *tandem* (Figura 10) contendo mais que 40 nucleotídeos. Experimentos realizados com alelos do gene *FLO1* que possuem regiões repetitivas de diferentes tamanhos mostraram que quanto maior o alelo, mais forte é adesão entre as células, evidenciando que a intensidade de floculação possa estar intimamente relacionada com o tamanho (quantidade de repetições) destas regiões (Verstrepen *et al.*, 2005; Rando & Verstrepen, 2007).



Referência: (Verstrepen *et al.*, 2004)

**FIGURA 10:** Motivos repetitivos de nucleotídeos nos genes *FLO*. A figura mostra os 5 genes *FLO* indicando apenas as sequências similares. Os quadrados mostram os motivos de sequência de DNA altamente conservados. Os quadros marcados com (X) representam motivos conservados em aminoácidos que não são conservados em sequência de DNA. Os “motif 4” e “motif 6” indicam as repetições intragênicas em tandem, que é maior no gene *FLO1*.

Leveduras patogênicas como *Candida albicans*, possuem genes de adesinas com repetições intragênicas em *tandem* (igualmente aos genes da floculação). Esses genes podem sofrer eventos de recombinação em determinadas situações de estresses, formando novas adesinas. Dessa forma a levedura pode ter uma grande plasticidade de fenótipos se adaptando rapidamente a novos ambientes. Tanto os genes da adesina como os da floculação estão posicionados na região subteloamérica do cromossomo. Essa posição isolada permite que ocorram eventos de recombinação sem que outros genes sejam afetados. Não se sabe se esse mecanismo de indução de variabilidade ocorra em *S. cerevisiae* com os genes da floculação da mesma forma que ocorre com os genes da adesina. Mas a instabilidade do fenômeno da floculação tem sido um grande problema para a indústria (Verstrepen & Klis, 2006).

O gene *FLO10* possui uma característica muito interessante que o diferencia dos outros genes da floculação. Os agregados de células formados pela expressão desse gene são inibidos pela presença de glicose, sacarose e maltose, enquanto que agregados formados pela expressão do gene *FLO1* são inibidos somente por manose (Guo *et al.*, 2000). Estudos realizados super expressando o gene *FLO10* mostram que ele apresenta uma forte floculação bem semelhante ao produto do gene *FLO1* (Guo *et al.*, 2000). A regulação da expressão do gene *FLO10*, assim como do gene *FLO11* são influenciadas por fatores genéticos e epigenéticos (Halme *et al.*, 2004).

Além dos genes responsáveis pela síntese da floclina, há outros genes como o gene *FLO8*, que juntamente com *TUP1* e *CYC8/SSN6*, participam da regulação transcricional de



*FLO1* (Kobayashi *et al.*, 1996). Em um estudo mais recente, foi verificado que o gene *MSS11*, que anteriormente era conhecido participando da regulação da degradação do amido, da formação de pseudohifas e do crescimento invasivo haplóide, também age como um forte indutor da floculação (Bester *et al.*, 2006). Outro gene, o *GTS-1*, foi caracterizado como um coativador transcricional. O produto da super expressão desse gene se liga ao repressor transcricional Sfl1p induzindo a expressão do gene *FLO1* e conseqüentemente promovendo a floculação na linhagem de laboratório testada. Isso foi verificado em linhagens mutantes para o gene *FLO8* (Shen *et al.*, 2006).

O controle da expressão dos genes de floculação, principalmente para os processos industriais de fermentação, pode trazer grandes benefícios (Domingues *et al.*, 2000). Um modelo muito interessante para a obtenção desta característica seria através da inserção de um promotor, ou parte dele, que atuasse no sentido de permitir a transcrição da proteína responsável pela floculação apenas no final da fermentação.

Os primeiros estudos que tentaram alterar e controlar a floculação em leveduras através de modificação genética foram descritos por Barney *et al.* (Barney *et al.*, 1980), que introduziu um fragmento do DNA cromossomal derivado de uma linhagem floculante, que incluía os genes *ADE1* e *FLO1*, em uma linhagem não floculante. Os mutantes se mostraram floculantes constitutivos, sendo assim indesejáveis aos propósitos da indústria. O mesmo resultado foi obtido por Ishida-Fujii (Ishida-Fujii *et al.*, 1998) usando um plasmídeo em que o gene *FLO1* estava sob o controle do promotor constitutivo da álcool desidrogenase 1 (*ADH1*). Em outro trabalho, o gene *FLO1* foi colocado sob o controle do promotor *HSP30*. Este é um promotor *heat shock* que é ativado em condições de estresse, o que acontece no final da fermentação devido a alta concentração de etanol e limitação de nutrientes (Verstrepen *et al.*, 2001).

A utilização do promotor do gene *ADH2* regulando o gene da floculação em linhagens leveduras de laboratório se mostrou muito promissor, floculando apenas quando a glicose havia se esgotado do meio, ao final da fermentação, tornando-se floculantes condicionais. A mesma construção foi utilizada para transformar uma linhagem industrial (Cunha *et al.*, 2006). O processo foi patenteado (INPI n<sup>o</sup> 0001122) e despertou grande atenção da comunidade especializada, de empresas de fermentação e biotecnologia e da imprensa de uma forma geral.

Em um estudo mais recente, promotores nativos dos genes *FLO1*, *FLO5* e *FLO11* foram substituídos por promotores do gene *ADH2* e *HSP30* (Govender *et al.*, 2008). Com estes experimentos, leveduras não floculantes tiveram sua sedimentação controlada expressando o gene apenas no final da fermentação. Porém estes experimentos foram feitos apenas em laboratório, onde as condições são controladas, e com cepas de levedura “domesticadas”. Nas dornas de fermentação, as condições são altamente estressantes (por exemplo, presença de contaminantes, variações bruscas de temperatura e pH, altas concentrações de etanol, entre outras), e o entendimento sobre a resposta das células a este tipo de ambiente ainda é limitado (Gibson *et al.*, 2007). Os genes de resposta a estresse podem também ser ativados em outras fases da fermentação e não somente no final. Isto pode ser visto em um trabalho de expressão gênica realizado durante o processo fermentativo de produção de vinho. Diversos genes foram analisados, entre eles o gene de *heat shock HSP30*, que se encontra ativado durante todo o processo fermentativo (Varela *et al.*, 2005).

A construção de uma linhagem de levedura que seja floculante para a produção de etanol combustível também foi realizada por um grupo de pesquisa chinês (Wang *et al.*, 2007). Apesar de a linhagem transformante apresentar as características como, produtividade de etanol, similares à linhagem selvagem, e ainda apresentar uma floculação forte e estável, o início da floculação era anterior ao esgotamento da glicose no meio de cultura.

Além da possibilidade de regulação dos genes da floculação, outros fatores podem ser interessantes para melhorar uma linhagem. Alper (Alper *et al.*, 2006) encontrou em bibliotecas de linhagens de laboratório de *S. cerevisiae* expostas a agentes mutagênicos, uma levedura que além de ser tolerante, produzia uma quantidade maior de etanol. Alper verificou que uma alteração de três bases (promovendo a substituição de três aminoácidos), sem a alteração da funcionalidade do gene *SPT15*, que codifica a proteína que se liga ao TATA-box, eram responsáveis por essas características. A mutagênese foi responsável pela substituição de fenilalanina por serina (Phe<sup>177</sup>Ser), tirosina por histidina (Tyr<sup>195</sup>His) e lisina por arginina (Lys<sup>218</sup>Arg). Em outro experimento, foi demonstrado que a interrupção da via do glicerol, pela deleção dos genes *GPD1* e *GPD2*, também promoveu um maior rendimento na produção de etanol. Foi observado um aumento, relativo à quantidade de

substrato consumido, de 6,83% e 7,41%, respectivamente, comparado com o selvagem (Guo *et al.*, 2009).

## 6. Atividade Mitocondrial

A atividade mitocondrial é essencial em diversas funções nas células de leveduras. A perda do DNA mitocondrial (petite citoplasmático) provoca sérias alterações na superfície celular, como a redução na absorção de açúcares, alteração na floculação (Calleja, 1973; Lung *et al.*, 1999) ou aglutinação e alteração nas proteínas da membrana plasmática reduzindo a tolerância ao etanol (Wilkie & Evans, 1982).

Células tratadas com drogas que inibem a função mitocondrial (Slonimski *et al.*, 1968; Egilsson *et al.*, 1979) ou com deleções dos genes *OL11* (codifica subunidade 9 da ATP sintase) ou *OXI2* (*COX3* – subunidade 3 da citocromo oxidase) (Hinrichs *et al.*, 1988) têm sua capacidade de floculação, crescimento e tolerância ao etanol reduzida ou anulada.

A comparação de células de *S. cerevisiae* deficiente-respiratório e competente-respiratório mostrou que a disfunção mitocondrial promove a ativação de um conjunto de genes associados com atividade peroxissomal e vias de restauração metabólica atenuando perdas do ciclo do ácido tricarboxílico (Epstein *et al.*, 2001).

A deleção de um gene específico, *PET191*, do DNA genômico gera um mutante de *S. cerevisiae* 100% deficiente respiratório denominado petite nuclear. Tal gene codifica a proteína responsável pela montagem do citocromo C na cadeia respiratória e foi verificado que seu mutante tem a capacidade de fermentação aumentada, produzindo 43% mais etanol que a linhagem parental, mostrando-se ainda tolerantes à presença de etanol e não apresentando comprometimento na taxa de crescimento (Hutter & Oliver, 1998). É importante salientar que neste experimento foi feito com linhagem de laboratório (FY23) e utilizando meio sintético. A deleção desse mesmo gene (*PET191*) em linhagens industriais, produtoras de vinho, promoveu um aumento de 9% na produção de etanol quando comparada com a linhagem selvagem (Panoutsopoulou *et al.*, 2001).

Petites citoplasmáticos ou mitocondriais, produzidos com o tratamento das células com brometo de etídio (agente mutagênico), em ensaios utilizando linhagens de vinícolas, produziram mais etanol quando comparada com a linhagem parental (Ooi & Lankford, 2009). Entretanto, petites citoplasmáticos sofrem diminuição na sua taxa de crescimento e

tendem a ser mais vulneráveis às condições adversas das dornas como a baixa tolerância ao etanol (Wilkie & Evans, 1982), tornando a linhagem imprópria para a utilização na indústria.

## JUSTIFICATIVA

Apesar de todos estes estudos, existem ainda uma ampla área a serem explorados com relação às linhagens industriais. Com o maior conhecimento a respeito das características destas leveduras, novos estudos poderão ser feitos e que permitirão o melhoramento do processo aumentando a capacidade industrial de produção de etanol. Nesse trabalho, iniciamos o processo de manipulação de linhagens industriais com o objetivo de transferir essa tecnologia para escala industrial. Analisamos a possibilidade de regular a expressão do gene da floculação de leveduras, tornando-as condicionais aos níveis de glicose do meio e avaliamos a produção de leveduras industriais deletadas para o gene *PET191* objetivando aumentar a produção de etanol. Os resultados deste trabalho apresentam perspectivas biotecnológicas e que poderão contribuir nesta área estratégica para o Brasil.

## **OBJETIVOS**

Os objetivos desse trabalho visam o estudo e modificações genéticas da levedura *Saccharomyces cerevisiae* para:

- Melhoria do processo fermentativo

- # Obtenção de floculantes condicionais – Avaliar linhagens transformadas com genes da floculação regulados por promotores sensíveis a níveis de glicose, possibilitando eliminar a etapa de centrifugação do processo.

- Melhoria do desempenho celular

- # Caracterização molecular de uma linhagem industrial.

- # Obtenção de petites (deficientes respiratório) buscando aumentar o rendimento na produção de etanol.

## **CAPITULO 1**

### **CONTROL BY SUGAR OF SACCHAROMYCES CEREVISIAE FLOCCULATION FOR INDUSTRIAL ETHANOL PRODUCTION**

Anderson F. Cunha, Silvia K. Missawa, Luiz H. Gomes, Sérgio F. Reis, Gonçalo A. G. Pereira

FEMS Yeast Res 6 (2006): 280–287

A idéia inicial do trabalho que originou esse artigo foi, através de modificações genéticas na levedura *Saccharomyces cerevisiae*, melhorar ou simplificar o processo fermentativo nas usinas brasileiras produtoras de etanol e/ou cachaça.

Foi verificado em artigos científicos a existência de genes da floculação, sendo que a ativação dessas proteínas promovem a separação das células do meio. A possibilidade de controlar a expressão desse gene poderia revolucionar o setor com a eliminação da etapa de centrifugação do processo fermentativo. O momento desejado para a ativação do gene seria quanto a fermentação acabasse, ou seja, com o esgotamento da glicose do meio.

Sendo assim, o trabalho foi iniciado utilizado como regulador para o gene da floculação o promotor do gene *MOX* (codifica a metanol oxidase) de *Hansenula polymorpha*. O estudo desse promotor foi tema do doutorado do Prof. Dr. Gonçalo (orientador desse trabalho), demonstrando que ele é devidamente expresso em *S. cerevisiae*, uma levedura que não metaboliza metanol. Além disso, este estudo demonstrou que a expressão do gene *MOX* é mediada por *ADRI*.

A utilização do promotor do gene *MOX* se mostrou promissor, mas apresentou alguns problemas, entre eles, o fato de se introduzir uma sequência exógena na levedura *S. cerevisiae*, gerando um transgênico. Sendo assim, procuramos na própria *S. cerevisiae*, um promotor que fosse regulado pelo gene *ADRI*. O promotor do gene *ADH2* se enquadrou perfeitamente nesse requisito.

Os principais resultados então presente neste artigo que foi publicado na FEMS Yeast Research no início do meu doutorado. Participei de alguns experimentos para a conclusão do artigo.

Além disso, o processo tecnológico foi submetido à patente (INPI nº 0001122). Esse trabalho também recebeu a menção honrosa no XXVI Prêmio Governador do Estado de São Paulo e foi tema de extenso artigo na revista *Scientific American Brasil*.

Segue abaixo o manuscrito.



# Control by sugar of *Saccharomyces cerevisiae* flocculation for industrial ethanol production

Anderson F. Cunha<sup>1</sup>, Sílvia K. Missawa<sup>1</sup>, Luiz H. Gomes<sup>2</sup>, Sérgio F. Reis<sup>3</sup> & Gonçalo A. G. Pereira<sup>1</sup>

<sup>1</sup>Laboratório de Genômica e Expressão, Departamento de Genética e Evolução Instituto de Biologia, Universidade Estadual de Campinas, São Paulo, Brazil; <sup>2</sup>Laboratório de Genética de Leveduras, Universidade de São Paulo, São Paulo, Brazil; and <sup>3</sup>Departamento de Parasitologia, Instituto de Biologia, Universidade Estadual de Campinas, São Paulo, Brazil

**Correspondence:** Gonçalo Amarante Guimarães Pereira, Laboratório de Genômica e Expressão, Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brazil. Tel.: +55 19 3788 6237; fax: +55 19 3788 6235; e-mail: goncalo@unicamp.br

Received 20 April 2005; revised 4 August 2005; accepted 25 October 2005.  
First published online 10 January 2006.

doi:10.1111/j.1567-1364.2006.00038.x

Editor: Monique Bolotin-Fukuhara

## Keywords

*FLO5*; flocculation; *Saccharomyces cerevisiae*; ethanol production; fermentation; biotechnology.

## Introduction

Brazil is the largest sugarcane producer in the world, followed by India and Australia. In 2003, 25 million tons of sugar were produced, together with 15 billion litres of ethanol ([www.portalunica.com.br/referencia/estatisticas.jsp](http://www.portalunica.com.br/referencia/estatisticas.jsp)), which was mostly utilized for fuel and industrial applications. The ethanol industry in Brazil typically uses the batch process (Wheals *et al.*, 1999); yeast is separated from the medium by centrifugation, which is the most expensive and complex step in ethanol production (Watari *et al.*, 1990). Centrifuge operation starts when fermentation stops, and this point is monitored by the evaluation of the sugar concentration (Brix) in the medium. For the production of beverages, such as beer, wine and sugar cane 'spirits' (Brazilian 'cachaça'), cell separation is normally obtained by natural cell sedimentation (Schwan *et al.*, 2001).

Yeast sedimentation can be accelerated by a process called flocculation, a reversible asexual process in which cells adhere to each other, resulting in the development of aggregates (Stratford, 1989; Bony *et al.*, 1997). In *Saccharomyces cerevisiae*, the phenomenon involves a calcium-

## Abstract

The goal of this study was to develop conditional flocculant yeasts for use in the alcohol fermentation industry. Promoters were constructed to completely repress gene transcription in the presence of sugar and to support strong expression after the exhaustion of this compound. A fusion of regulatory regions of the *ADH2* promoter with the *FLO5* core promoter was constructed to regulate the *FLO5* gene. This construct was inserted into multicopy plasmids and transformed into laboratory strains of *Saccharomyces cerevisiae*, whereby the transformed cells were selected by sedimentation from the bulk medium after sugar exhaustion, without decreasing ethanol production. The *ADH2-FLO5* region was converted into an integrative cassette to disrupt the *CAN1* gene in industrial yeast strains. Transformed cells became resistant to canavanine and demonstrated conditional flocculation. Although ethanol production was significantly decreased in the industrial transformants, this development reveals a promising technology for the substitution of centrifugation in industrial ethanol production.

dependent protein–sugar interaction between specific lectin-like surface proteins, called flocculins, and the cell wall mannan of adjacent cells (Stratford, 1989; Teunissen *et al.*, 1993). The best-known flocculin gene is *FLO1*, a dominant gene situated on the right arm of chromosome I (Teunissen *et al.*, 1993), and the locus is also occupied by the alleles *FLO2* and *FLO4* (Stewart & Russell, 1977; Sieiro *et al.*, 1997). Other important *FLO* genes are *FLO5*, which is a *FLO1* paralogue located on chromosome VIII (Teunissen & Steensma, 1995), *FLO9* and *FLO10* (Teunissen & Steensma, 1995; Guo *et al.*, 2000). Interestingly, in the yeast *Kluyveromyces marxianus*, flocculation has been associated with the overexpression of an isoform of glyceraldehyde-3-phosphate dehydrogenase (Almeida *et al.*, 2003).

Some processes have tried to replace centrifugation by the natural ability of certain yeast strains to flocculate (Wieczorek & Michalski, 1994; Abate *et al.*, 1996; Paiva *et al.*, 1996; Kondo *et al.*, 2002), or by the introduction of flocculation genes into nonflocculant strains (Barney *et al.*, 1980; Watari *et al.*, 1994; Ishida-Fujii *et al.*, 1998). In the study by Watari *et al.* (1994), the resulting yeasts flocculated constitutively, reducing the performance of the process, although the

ethanol concentration at the end of the fermentation was similar to that realized with nonflocculant cells. These authors suggested that the process could be improved by the use of more appropriate promoters. Ishida-Fujii *et al.* (1998) observed that the fermentation ability of the transformed strain was almost the same as that of the parent strain, except for the fermentation rate. However, the experiment was performed under constant agitation, which may mask the disadvantage of flocculation during natural fermentation.

Constitutive flocculation seems to produce sluggish fermentation (Verstrepen *et al.*, 2003), the cause of which is still unknown. Nevertheless, constitutive flocculation has been successfully employed to increase the cell density of yeast in the fermentation of cheese whey permeate, a by-product of dairy industries, which presents rather high pollutant characteristics and is produced in large amounts (Domingues *et al.*, 2001). Similarly, Remize *et al.* (1998) employed flocculant yeasts, obtained by transformation of nonflocculant strains with the *FLO1* gene under the control of the *ADH1* promoter, for the enrichment of fructose from Jerusalem artichoke tuber extracts, hydrolysed by H<sub>2</sub>SO<sub>4</sub>. In this case, yeast sedimentation, controlled by air input, facilitated an incomplete fermentation process, generating a fructose-rich extract; ethanol and yeast biomass were valuable by-products.

At an industrial scale, systems using flocculating yeast cells have been employed only in the later phase of primary beer fermentation to separate biomass from the fermented broth (Domingues *et al.*, 2001). In this case, the process employs natural yeast strains presenting the NewFlo phenotype, in which the flocculins are inhibited by sugars, allowing sedimentation at the end of fermentation (Stratford & Assinder, 1991).

For biotechnology, the possibility of inducing controlled flocculation in a naturally nonflocculant strain provides great potential for improving brewing, wine-making, baking and ethanol-producing yeast (Dequin, 2001). A recent report placed the *FLO1* gene under the transcriptional control of the late-fermentation *HSP30* promoter (Verstrepen *et al.*, 2001). The transformed yeast cells showed strong flocculation at the end of fermentation, indicating that flocculation could be controlled by genetic modification.

*Saccharomyces cerevisiae* preferentially utilizes glucose via fermentation to produce ethanol. When glucose is available, regulatory systems repress the expression of the genes involved in respiration and in the utilization of other carbon sources, in a phenomenon called carbon catabolite repression (Gancedo, 1998). After glucose exhaustion, a number of genes are activated, completely changing the cell transcriptome (DeRisi *et al.*, 1997; Ferea *et al.*, 1999). Interestingly, the molecular mechanisms involved in glucose repression are widely conserved. For example, the regulation of the gene encoding the peroxisomal methanol oxidase

(*MOX*) from *Hansenula polymorpha* has been demonstrated to be conserved in *S. cerevisiae*, a species that does not metabolize methanol (Pereira & Hollenberg, 1996).

In this study, a nonflocculant laboratory strain of *S. cerevisiae* was transformed with a *FLO5* gene under the control of glucose-repressible promoters, resulting in strains that aggregated only when the sugar was completely exhausted. The same process was successfully used to transform industrial yeasts. The results show the potential of this promising new technology to replace centrifugation in industrial ethanol production.

## Materials and methods

### Strains, media and glucose measurement

The *Saccharomyces cerevisiae* strains used in this study were W303-1a (MATa ade2-1 trp1-1 can1-100 leu2-3, 112 his3-11 ura3) and Fleischmann (commercial yeast used in bakeries). Cells were grown at 30 °C on YEP medium (1% yeast extract, 2% Bacto-peptone), with 2% glucose (YPD), or SD medium containing 0.67% Yeast Nitrogen Base without amino acids and 2% glucose supplemented with 20 mg L<sup>-1</sup> histidine, 30 mg L<sup>-1</sup> leucine, 20 mg L<sup>-1</sup> tryptophan (for laboratory yeasts transformed with pYADE4) or 20 mg L<sup>-1</sup> uracil (for laboratory yeasts transformed with YEp plasmids), and 60 mg L<sup>-1</sup> L-canavanine (Sigma, St Louis, MO) (for industrial yeasts). Agar at 2% was used for solid media (Ausubel *et al.*, 1998). The concentration of glucose was enzymatically determined using the glucose oxidase/peroxidase-4-aminophenazone-phenol system (Trinder, 1969).

### Plasmids and DNA manipulation

Plasmids YEp352-*FLO5*, donated by Bidard [Laboratoire de Microbiologie et Technologie des Fermentations, Institut des Produits de la Vigne (IPV), INRA, Montpellier, France] (Bidard *et al.*, 1994), pYADE4 (Brunelli & Pall, 1993) and pUC18*MOX-BB* (Pereira & Hollenberg, 1996) were used for the construction of recombinant plasmids. Plasmid and genomic DNA preparation, gel electrophoresis and purification, as well as yeast and bacterial transformation, were all performed using standard methods, as described elsewhere (Ausubel *et al.*, 1998).

### RNA isolation and analysis

Total yeast RNA was extracted by the hot acid phenol method, and Northern blotting was performed as described previously (Ausubel *et al.*, 1998). For probe preparation, a 1900 bp *HincII* fragment containing the region from +633 to +2547 of the *FLO5* coding sequence was isolated from the YEp352-*FLO5* plasmid (Bidard *et al.*, 1994) and radioactively labelled by random primers using  $\alpha$  (Bidard *et al.*,

1994) P-dATP (Ausubel *et al.*, 1998). Probed membranes were exposed to Kodak film (X-OMAT). Ribosomal RNA stained with ethidium bromide, which showed a fairly constant abundance under different growth conditions and amongst the constructed strains, was used to calibrate the amount of total RNA loaded onto the gels per sample.

### Construction of the *CAN::ADH2/FLO5* cassette

To construct the *CANADH2/FLO5* cassette to transform industrial yeast, a PCR was performed using the primers CANFLO5R (5'CATAGGTGATGAAGATGAAGGAGAAGTACAGAACGTTAAATAATTGCCAGCAATAAGG3') and CANADH2F (5'AACTTTGTACCACCAGTAGATGTCTCCATGTAAGAAAGGGGCAAAACGTAGGGGC3'). The reaction was performed in a final volume of 50  $\mu$ L using the following amplification mixture: 40 ng  $\mu$ L<sup>-1</sup> of pYADE4-*FLO5* plasmid as DNA template, 5  $\mu$ L of 10  $\times$  reaction buffer (50 mM of KCl; 10 mM of Tris-HCl, pH 9; 0.1% Triton X-100), 2  $\mu$ L of 50 mM MgSO<sub>4</sub>, 12  $\mu$ L of 1.25 mM of each dNTP (dATP, dCTP, dGTP and dTTP), 15 pmol each of the two amplification primers and 1  $\mu$ L of Platinum *Taq* DNA Polymerase (Invitrogen, Carlsbad, CA; 5 U  $\mu$ L<sup>-1</sup>). The cycling parameters were as follows: 94 °C for 2 min, followed by 35 cycles at 94 °C for 15 s and 68 °C for 5 min.

### Ethanol quantification

Ethanol production was evaluated for five fermentations and three repetitions per assay with cells growing in

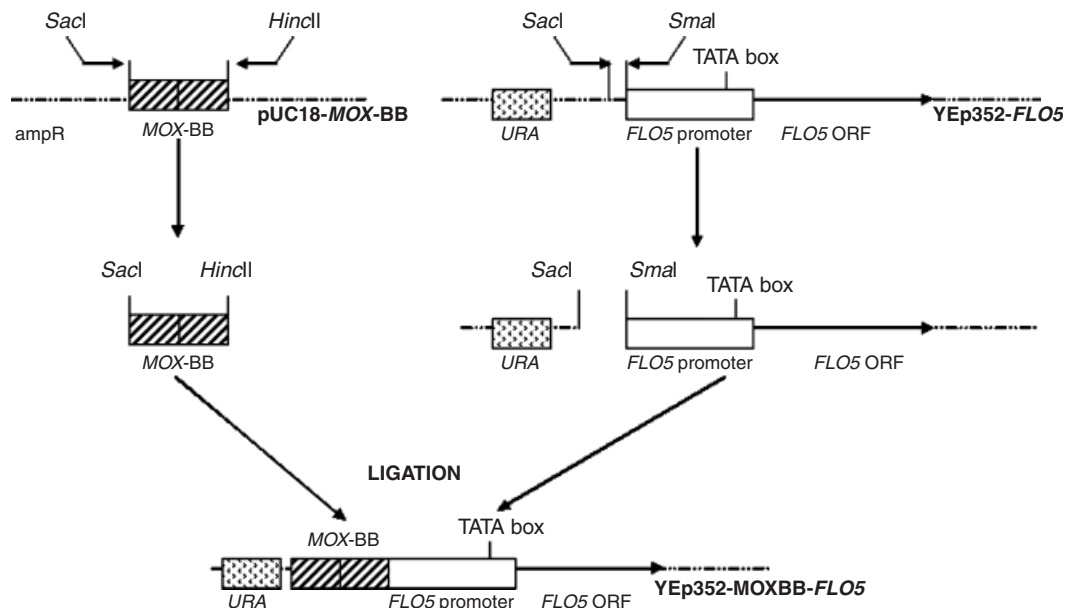
YEP+8% sucrose. The alcohol concentration was measured after distillation by densitometry.

## Results

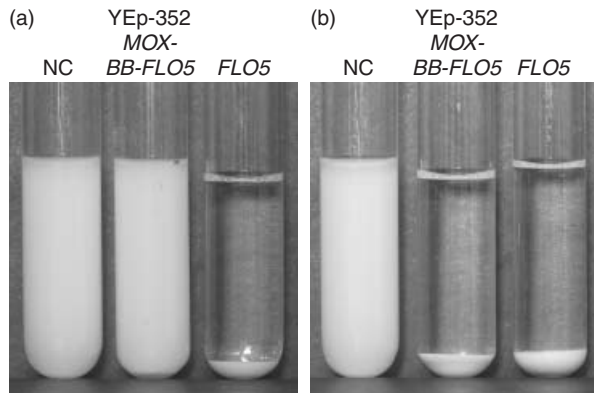
### Induction of flocculation based on plasmids

To place the expression of the gene *FLO5* under glucose repression, a regulatory region of the *MOX* promoter, named *MOX-B* (Pereira & Hollenberg, 1996), was inserted in duplicate, upstream of the putative *FLO5* promoter present in YEp352-*FLO5* (Bidard *et al.*, 1994), creating YEp-*BB-FLO5* (Fig. 1). The laboratory strain W303-1a was transformed with YEp-*BB-FLO5*, YEp352 (negative control) and YEp352-*FLO5* (positive control). The transformants were inoculated into 20 mL test tubes containing 5 mL of YPD with 2% glucose and incubated with vigorous agitation for 14 h. These cultures were used to inoculate 500 mL flasks containing 50 mL of the same medium (OD<sub>600</sub> = 0.1). After 12 h, no flocculation could be observed in the negative control or in the transformant bearing the YEp352-*MOX-BB-FLO5* plasmid. Conversely, the transformants carrying YEp352-*FLO5* flocculated continuously. After 20 h, the cells carrying the YEp352-*MOX-BB-FLO5* plasmid clearly started to flocculate. The flasks were then taken from the shaker and the cells precipitated instantly, in contrast with the cells of the negative control (Fig. 2).

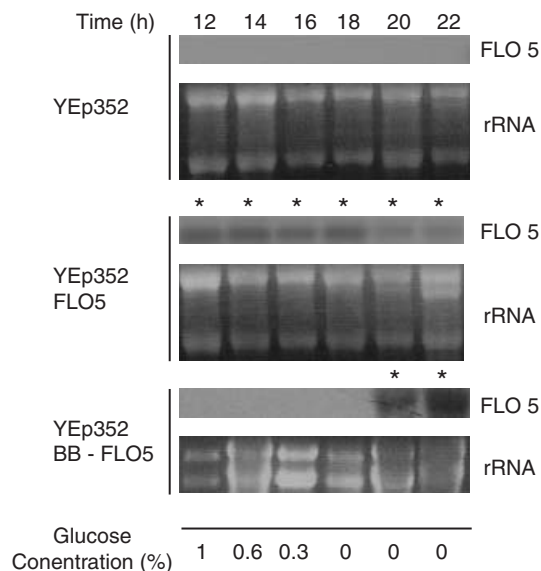
To investigate the pattern of plasmid-based *FLO5* expression, total RNA of these cultures was taken after 12, 14, 16, 18, 20 and 22 h, and Northern blotting experiments were



**Fig. 1.** Construction of YEp-*BB-FLO5*. The pUC18*MOX-BB* plasmid was digested by *SacI* and *HincII* and the resulting fragment (*MOX-BB*) was cloned in the YEp352-*FLO5* plasmid, previously digested with *SmaI* and *SacI*. This construct was used for yeast transformation to place the *FLO5* putative promoter and gene under the transcriptional control of the *MOX* promoter regulatory region (*MOX-BB*).



**Fig. 2.** *FLO5* expression under the control of the *MOX-BB* region. The laboratory strain W303-1a was transformed with YEp352 (negative control, NC), YEp352-*FLO5* (positive control, PC) and YEp352-*MOX-BB-FLO5*. The figure shows the flocculation patterns of the yeast strains transformed with YEp352-*MOX-BB-FLO5*, which present a nonflocculant growth in a medium still containing glucose (a) and rapidly aggregate after glucose depletion (b).



**Fig. 3.** *FLO5* expression during fermentation. Northern blot analysis of RNA isolated from yeast transformed with YEp352 (negative control), YEp352-*FLO5* (positive control) and YEp352-*BB-FLO5* during the times (h) indicated in the figure. The glucose concentration was evaluated during fermentation. Asterisk indicates yeast flocculation.

performed. The results showed a clear correlation between flocculation and *FLO5* expression (Fig. 3). In the strain transformed with the YEp-352-*BB-FLO5* construct, the expression of *FLO5* was observed only after 20 h (Fig. 3), 2 h after the exhaustion of glucose in the medium.

### Control of *FLO5* expression by the *ADH2* promoter

*MOX-B* is not a *Saccharomyces cerevisiae* native sequence and its future use in industrial processes could raise regula-

tory concerns. To overcome this problem, we tried the same strategy by replacing *MOX-BB* with the *ADH2* promoter, an *S. cerevisiae* native sequence that, similar to *MOX-B*, is regulated by Adr1p (Pereira & Hollenberg, 1996).

To test the expression pattern of this promoter, the green fluorescent protein (*GFP*) reporter gene was placed under its control (plasmid pYADE4). The results showed that, although *GFP* expression was higher after glucose exhaustion, the promoter was not silent at higher glucose concentration, presenting a significant basal expression (result not shown). Thus, instead of using the complete *ADH2* promoter, the regulatory sequences of this region encompassing the poly-T region, *UAS1* and *UAS2* (Yu *et al.*, 1989), were fused to the *FLO5* core promoter (Fig. 4).

### Yeast transformation and inspection of *FLO5* expression during fermentation

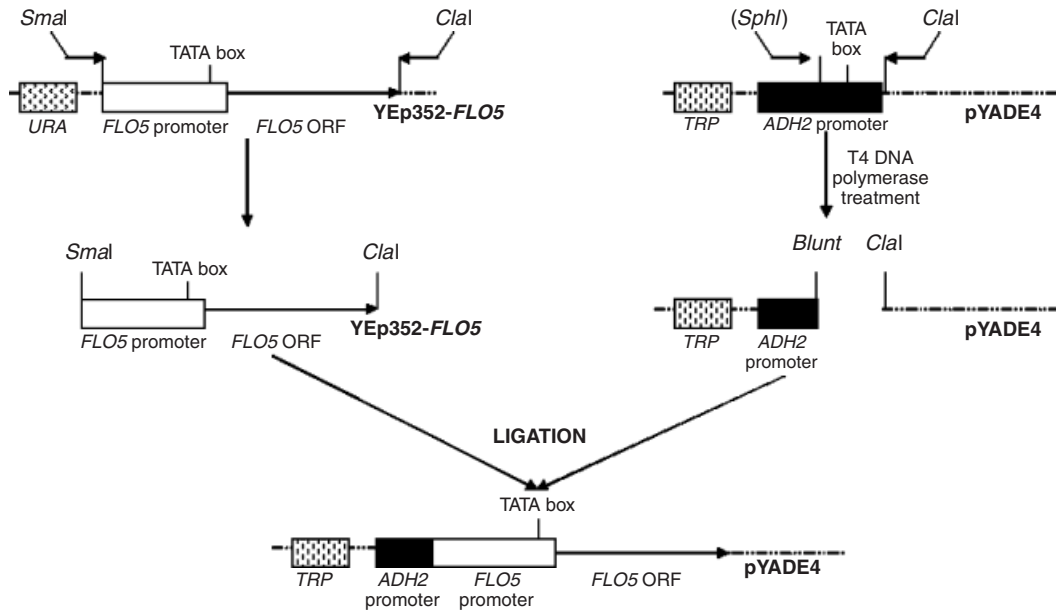
Strain W303-1a was transformed with the plasmid pYADE4-*FLO5*. After 18 h, the cells carrying the plasmid clearly started to flocculate (data not shown). The results of a Northern blot analysis showed a direct correlation between the exhaustion of glucose and the initiation of *FLO5* expression (Fig. 5a).

The next step was to evaluate whether ethanol production would be affected by flocculation. Figure 5b shows that the ethanol production after the end of fermentation (22 h) was practically the same in both transformed and untransformed strains.

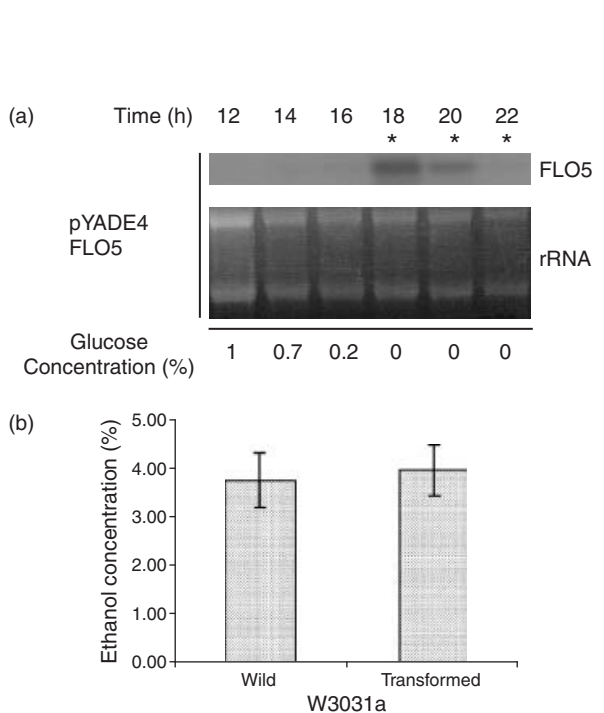
### Modification of industrial yeast

In Brazil, each ethanol industry utilizes a specific strain, adapted to the local conditions. Therefore, in order to become a widely accepted technology, the constructs described above should be adapted to transform, in an integrative manner, industrial strains that are normally diploid and have no nutritional markers. To do this, we used the strategy of integrating the expression cassette in the middle of the *CAN1* gene, which encodes an arginine permease (Wach *et al.*, 1994; Camargo, 2000). Positive integration events led to the inactivation of *CAN1* and, thereby, the cells became resistant to canavanine, an antibiotic analogous to arginine. A diagram of this construct is summarized in Fig. 6a. Basically, the cassette *ADH2-FLO5* from plasmid pYADE4-*FLO5* was amplified using primers that contained additional sequences of the *CAN1* gene. The PCR products were then utilized to transform the commercial Fleischmann strain. Transformants were selected by their ability to grow in the presence of canavanine. The *CAN1* deletion and the integration of the cassette were confirmed by PCR (Fig. 6b).

Transformed cells were tested for flocculation and alcohol production. The transformed strain was grown in YPD with



**Fig. 4.** Construction of pYADE4-FLO5. The YEp352-FLO5 plasmid was digested with *Sma*I and *Cl*at to excise the *FLO5* ORF together with its putative promoter (sequence from – 1 to – 156). This sequence was then cloned into pYADE4 at the *Cl*at and *Sph*I sites, which were later converted into a blunt-end site by *T4* DNA polymerase treatment. This plasmid was used to transform the laboratory strain W303-1a.



**Fig. 5.** *FLO5* expression under the control of the *ADH2* promoter during fermentation, and evaluation of ethanol production. (a) Northern blot analysis of RNA isolated from yeast transformed with pYADE4-FLO5 during the times (h) indicated in the figure. The glucose concentration was evaluated during fermentation. Asterisk indicates yeast flocculation. (b) No differences in ethanol production were detected between the transformed and wild-type yeasts.

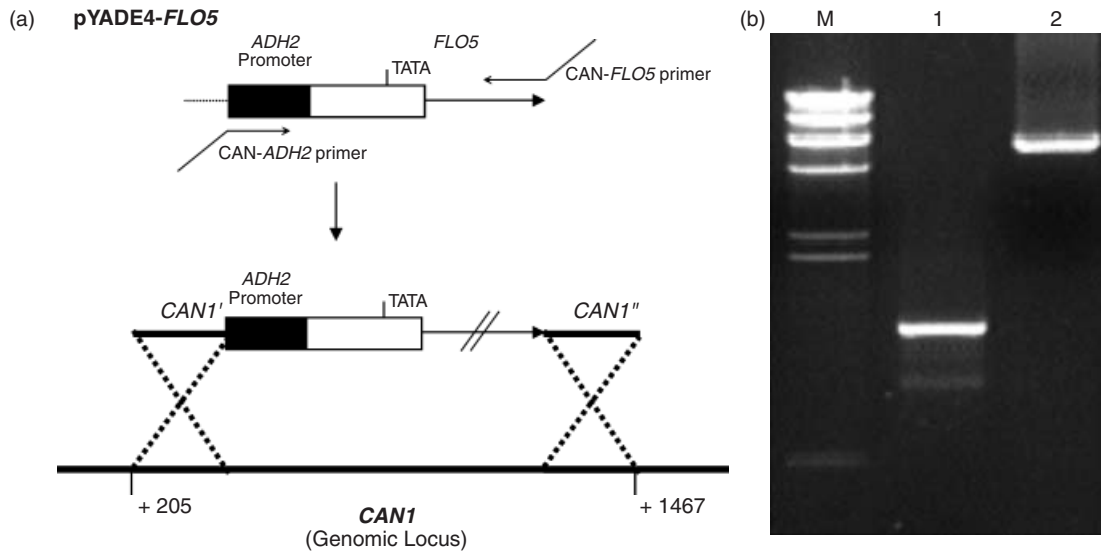
2% glucose and fermented in the same medium, but with 20% glucose, until this sugar was exhausted, approximately 15 h. The results showed a strong flocculation in the transformed strain after this time in comparison with the nontransformed strain (Fig. 7a). However, the ethanol production by the transformed strain was reduced by approximately 38% (Fig. 7b).

## Discussion

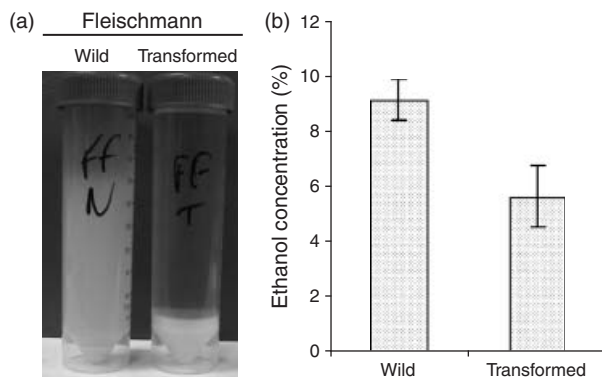
The goal of this study was to develop a genetic system to produce yeast cells capable of sensing the glucose concentration in the medium and flocculating after glucose exhaustion. The resulting industrial processes using these strains could then eliminate centrifugation in the case of ethanol production, or shorten the period of cell sedimentation for spirits or wine production. In the latter case, the decrease in the sedimentation time could reduce the breakage of cells, which is responsible for off-flavours that lower the quality of the products.

We started by transforming the laboratory strain, W303-1a, with the *FLO5* gene containing 156 bp upstream of the ATG, thus achieving constitutive expression of the gene, which led to constitutive flocculation. To control the expression of this gene, a regulatory region of the *MOX* gene, controlled by the transcriptional activator Adr1p (Pereira & Hollenberg, 1996), was inserted as a doublet upstream of the *FLO5* promoter. This construct facilitated the control of flocculation, rendering the *FLO5* promoter and flocculation





**Fig. 6.** Construction of *CAN::ADH2/FLO5* cassette and confirmation of cassette integration in the *CAN1* locus. (a) The *ADH2-FLO5* cassette from plasmid pYADE4-*FLO5* was amplified using primers with additional sequences of the *CAN1* gene. The amplified cassette was used for industrial yeast transformation (Fleischmann) and replaced the wild-type *CAN1* by a construct in which the putative *FLO5* promoter and gene were under the transcriptional control of the *ADH2* promoter region. (b) The *CAN1* deletion and the integration of the cassette were confirmed by PCR. The amplification of the construct *ADH2-FLO5* (5246 bp) was observed only in the transformed yeast (column 2).



**Fig. 7.** Flocculation and ethanol production of the transformed Fleischmann strain. (a) Difference in flocculation between cultures of wild-type Fleischmann and transformed Fleischmann after 15 h of fermentation in medium containing 1% yeast extract, 2% Bacto-peptone and 20% glucose. (b) Evaluation of ethanol production by the transformed strain shows a decrease of approximately 38% when compared with the wild-type. The fermentation was performed in medium containing 20% glucose. The cells for this assay were grown in 2% glucose.

subject to glucose repression. This prototype was patented in 2000 by the National Institute of Intellectual Property-Brazil under the patent number 0001122.

However, this regulatory region comes from another yeast species, *Hansenula polymorpha*, and potential industrial *Saccharomyces cerevisiae* strains bearing this region would be considered genetically modified organisms requiring more strict regulatory rules. Moreover, flocculation was observed to occur only 2 h after glucose exhaustion.

In view of this, it was decided to test native promoters to improve the process and overcome the regulatory concerns. Initially, the *ADH2* promoter, which is also controlled by *Adr1p* (Yu *et al.*, 1989), was tested. However, initial trials using the GFP reporter gene showed that the transcription in the presence of glucose was not completely abolished (data not shown). To avoid this problem, we used the native *FLO5* core promoter fused to only the regulatory regions of *ADH2*. With this modification, we effectively eliminated transcription in the presence of glucose and co-ordinated *FLO5* expression precisely with the depletion of glucose. Therefore, this seems to be the ideal promoter to satisfy our goals.

Other authors have reported the conversion of nonfloculant yeast strains into flocculant strains (Barney *et al.*, 1980; Watari *et al.*, 1990, 1994; Ishida-Fujii *et al.*, 1998). In particular, Verstrepren *et al.* (2001) have successfully produced conditional flocculant laboratory strains using the late-fermentation *HSP30* promoter to control *FLO1* expression. However, the *HSP30* promoter can be activated by a series of other stress factors in addition to ethanol concentration, such as temperature or pH (Seymour & Piper, 1999; Varela *et al.*, 2005). Thus, in the case of the ethanol industry, in which the cells are under tremendous stresses of temperature, pH and competition with bacteria and other yeasts, stress-sensitive promoters can present transcription escapes before the end of fermentation. In this sense, the exhaustion of sugar seems to be a much more appropriate environmental signal.

One important point to be considered is that large-scale industrial ethanol fermentation is performed under nonsterile conditions. Therefore, there is no guarantee that the yeast strain inoculated at the beginning of a continuous fermentation cycle will remain until the end of the process, which normally takes several months. To address this point, we followed the fate of an inoculated strain in a standard alcohol industrial plant. Using microsatellite PCR, we detected that the cells of the inoculated strains were completely replaced by a wild-type strain 30 days after inoculation (result not shown). Similar results have been observed by other authors (Sabate *et al.*, 1998; Wheals *et al.*, 1999; Howell *et al.*, 2004).

Therefore, for industrial applications, a system must be developed which is able to convert specific yeast strains, which predominate at each industrial plant, to conditional flocculant strains. For this, we integrated the flocculation cassette into the *CAN1* locus (Camargo, 2000). This system has the advantage of avoiding the use of positive antibiotic genes, such as for kanamycin or gentamicin resistance (Wach *et al.*, 1994; Longtine *et al.*, 1998; Goldstein *et al.*, 1999), which would be very harmful as the excess yeast cells are normally used for animal feed after the end of the fermentation cycle. Therefore, strains bearing such genes would not be approved for industrial applications.

In the initial experiments, we were able to select canavanine-resistant cells, in which the cassette was effectively integrated in the cells and which presented flocculation after the exhaustion of sugar from the growth medium. However, in contrast with the observations with the laboratory strain, the industrial strain presented a significant decrease in ethanol production in comparison with nontransformed cells. We are currently investigating the causes of this decrease, testing new constructions combining different promoters and *FLO* genes and evaluating other industrial yeast strains.

In summary, this study shows the first results of the development of industrial yeast strains capable of flocculation co-ordinated with sugar consumption, allowing the separation of cells at the end of fermentation in a natural manner, without the need for mechanical processing.

## Acknowledgements

This work was supported by grants from GENESEARCH Biotech and Fundo de Apoio ao Ensino e Pesquisa – Universidade Estadual de Campinas. Anderson Cunha was supported by a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior fellowship. We thank Eduardo F. Formighieri for image processing.

## References

Abate C, Callieri D, Rodriguez E & Garro O (1996) Ethanol production by a mixed culture of flocculent strains of

*Zymomonas mobilis* and *Saccharomyces* sp. *Appl Microbiol Biotechnol* **45**: 580–583.

Almeida C, Queirós O, Wheals A, Teixeira J & Moradas-Ferreira P (2003) Acquisition of flocculation phenotype by *Kluyveromyces marxianus* when overexpressing *GAP1* gene encoding an isoform of glyceraldehyde-3-phosphate dehydrogenase. *J Microbiol Methods* **55**: 433–440.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith AJ & Struhl K (1998) *Current Protocols in Molecular Biology*. Wiley, New York.

Barney MC, Jansen GP & Helber GR (1980) Use of genetic transformation for the introduction of flocculence into yeast. *J Am Soc Brewing Chemists* **38**: 71–74.

Bidard F, Blondin B, Dequin S, Vezinhet F & Barre P (1994) Cloning and analysis of a *FLO5* flocculation gene from *S. cerevisiae*. *Curr Genet* **25**: 196–201.

Bony M, Thines-Sempoux D, Barre P & Blondin B (1997) Localization and cell surface anchoring of the *Saccharomyces cerevisiae* flocculation protein Flo1p. *J Bacteriol* **179**: 4929–4936.

Brunelli JP & Pall ML (1993) A series of yeast shuttle vectors for expression of cDNAs and other DNA sequences. *Yeast* **9**: 1299–1308.

Camargo ME (2000) Sistema para transformação de leveduras industriais e detecção de atividade recombinogênica. Thesis/ Instituto de Ciências Biológicas, Universidade de São Paulo, São Paulo.

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

Dequin S (2001) The potential of genetic engineering for improving brewing, wine-making and baking yeasts. *Appl Microbiol Biotechnol* **56**: 577–588.

Domingues L, Lima N & Teixeira JA (2001) Alcohol production from cheese whey permeate using genetically modified flocculent yeast cells. *Biotechnol Bioeng* **72**: 507–514.

Ferea TL, Botstein D, Brown PO & Rosenzweig RF (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc Natl Acad Sci* **96**: 9721–9726.

Gancedo JM (1998) Yeast carbon catabolite repression. *Microbiol Mol Biol Rev* **62**: 334–361.

Goldstein AL, Pan X & McCusker JH (1999) Heterologous URA3MX cassettes for gene replacement in *Saccharomyces cerevisiae*. *Yeast* **15**: 507–511.

Guo B, Styles CA, Feng Q & Fink GR (2000) A *Saccharomyces* gene family involved in invasive growth, cell–cell adhesion, and mating. *Proc Natl Acad Sci USA* **97**: 12158–12163.

Howell KS, Bartowsky EJ, Fleet GH & Henschke PA (2004) Microsatellite PCR profiling of *Saccharomyces cerevisiae* strains during wine fermentation. *Lett Appl Microbiol* **38**: 315–320.

Ishida-Fujii K, Goto S, Sugiyama H, Takagi Y, Saiki T & Takagi M (1998) Breeding of flocculent industrial alcohol yeast strains by self-cloning of the flocculation gene *FLO1* and repeated-batch fermentation by transformants. *J Gen Appl Microbiol* **44**: 347–353.

- Kondo A, Shigechi H, Abe M, *et al.* (2002) High-level ethanol production from starch by a flocculent *Saccharomyces cerevisiae* strain displaying cell-surface glucoamylase. *Appl Microbiol Biotechnol* **58**: 291–296.
- Longtine MS, McKenzie A III & Demarini DJ (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**: 953–961.
- Paiva TC, Sato S, Visconti AE & Castro LA (1996) Continuous alcoholic fermentation process in a tower reactor with recycling of flocculating yeast. *Appl Biochem Biotechnol* **57–58**: 535–541.
- Pereira GG & Hollenberg CP (1996) Conserved regulation of the *Hansenula polymorpha* MOX promoter in *Saccharomyces cerevisiae* reveals insights in the transcriptional activation by Adr1p. *Eur J Biochem* **238**: 181–191.
- Remize F, Schorr-Galindo S, Guiraud JP, Dequin S & Blondin B (1998) Construction of a flocculating yeast for fructose production from inulin. *Biotechnol Lett* **20**: 313–318.
- Sabate J, Cano J, Querol A & Guillamón JM (1998) Diversity of *Saccharomyces* strains in wine fermentations: analysis for two consecutive years. *Lett Appl Microbiol* **26**: 452–455.
- Schwan RF, Mendonça AT, da Silva JJ Jr, Rodrigues V & Wheals AE (2001) Microbiology and physiology of cachaça (aguardente) fermentations. *Antonie Van Leeuwenhoek* **79**: 89–96.
- Seymour IJ & Piper PW (1999) Stress induction of HSP30, the plasma membrane heat shock protein gene of *Saccharomyces cerevisiae*, appears not to use known stress-regulated transcription factors. *Microbiology* **145**: 231–239.
- Sieiro C, Reboredo NM, Blanco P & Villa TG (1997) Cloning of a new FLO gene from the flocculating *Saccharomyces cerevisiae* IM1-8b strain. *FEMS Microbiol Lett* **146**: 109–115.
- Stewart GG & Russell I (1977) The identification, characterization, and mapping of a gene for flocculation in *Saccharomyces* sp. *Can J Microbiol* **23**: 441–447.
- Stratford M (1989) Evidence for two mechanisms of flocculation in *Saccharomyces cerevisiae*. *Yeast* **5**: S441–S445.
- Stratford M & Assinder S (1991) Yeast flocculation: flo1 and NewFlo phenotypes and receptor structure. *Yeast* **7**: 559–574.
- Teunissen AW, Holub E, van der Hucht J, van den Berg JA & Steensma HY (1993) Sequence of the open reading frame of the FLO1 gene from *Saccharomyces cerevisiae*. *Yeast* **9**: 423–427.
- Teunissen AW & Steensma HY (1995) Review: the dominant flocculation genes of *Saccharomyces cerevisiae* constitute a new subtelomeric gene family. *Yeast* **11**: 1001–1013.
- Trinder P (1969) Determination of blood glucose using an oxidase-peroxidase system with a non-carcinogenic chromogen. *J Clin Pathol* **22**: 158–161.
- Varela C, Cardenas J, Melo F & Agosin E (2005) Quantitative analysis of wine yeast gene expression profiles under winemaking conditions. *Yeast* **22**: 369–383.
- Verstrepen KJ, Derdelinckx G, Delvaux FR, *et al.* (2001) Late fermentation expression of FLO1 in *Saccharomyces cerevisiae*. *J Am Soc Brewing Chemists* **59**: 69–76.
- Verstrepen KJ, Derdelinckx G, Verachtert H & Delvaux FR (2003) Yeast flocculation: what brewers should know. *Appl Microbiol Biotechnol* **61**: 197–205.
- Wach A, Brachat A, Pohlmann R & Philippsen P (1994) New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* **10**: 1793–1808.
- Watari J, Kudo M, Nishikawa N & Kamimura M (1990) Construction of flocculant yeast cells (*Saccharomyces cerevisiae*) by mating or protoplast fusion using a yeast cell containing the flocculation gene FLO5. *Agric Biol Chem* **54**: 1677–1681.
- Watari J, Nomura M, Sahara H, Koshino S & Keranen S (1994) Construction of flocculant brewers yeast by chromosomal integration of the yeast flocculation gene FLO1. *J Inst Brewing* **100**: 73–77.
- Wheals AE, Basso LC, Alves DMG & Amorim HV (1999) Fuel ethanol after 25 years. *Trends Biotechnol* **17**: 482–487.
- Wieczorek A & Michalski H (1994) Continuous ethanol production by flocculating yeast in the fluidized bed bioreactor. *FEMS Microbiol Rev* **14**: 69–74.
- Yu J, Donoviel MS & Young ET (1989) Adjacent upstream activation sequence elements synergistically regulate transcription of ADH2 in *Saccharomyces cerevisiae*. *Mol Cell Biol* **9**: 34–42.



## **CAPITULO 2**

### **REGULAÇÃO CONDICIONAL DO GENE DA FLOCULAÇÃO *FLO10***

## INTRODUÇÃO

Os resultados obtidos com a regulação do gene *FLO5* mostraram a possibilidade da “criação” de uma levedura inteligente. Porém, o gene *FLO5* apresenta uma desvantagem, as células não desfloculam após a adição de um novo mosto. Isso causaria um problema na indústria, pois não seria possível a reutilização das leveduras.

Segundo a literatura, os agregados de células formada pela expressão do gene *FLO5* são inibidos apenas por manose, entretanto, foi relatado que agregados de células formados pela expressão de um outro gene da floculação, *FLO10*, são inibidos pela presença de glicose, sacarose e maltose (Guo *et al.*, 2000). Essa informação é importante e motivou a continuidade da idéia de produzir linhagens floculantes condicionais, pois haveria grandes possibilidades de que as células desfloculassem imediatamente ao serem colocadas nas dornas de fermentação com um novo mosto rico em glicose e sacarose.

O objetivo desta etapa foi usar a mesma estratégia descrita no capítulo anterior para controlar a expressão do gene *FLO10*, em linhagens de laboratório, avaliando sua expressão gênica durante a fermentação e também construindo um cassete integrativo para transformação de linhagens industriais que são utilizadas na produção de etanol combustível.

## METODOLOGIA

### I - Linhagens utilizadas

Linhagens da levedura *Saccharomyces cerevisiae*:

- W303-1a: Mat **a** ade2 can1-100 his3-11, 15 leu2-13, 112 trp1-1 ura3-1;
- FY23: Linhagem padrão S288c: Mat **a**, ura3-52, leu2D1, trp1D63;
- JAY270: isolado da linhagem industrial Pedra 2, sem marca auxotrófica;
- JAY291: Mat **a**, haplóide de um esporo proveniente da JAY270;
- JAY292: Mat **alfa**, haplóide de um esporo proveniente da JAY270.

Linhagem da bactéria *Escherichia coli* empregada para a clonagem e multiplicação dos plasmídeos:

- *DH10b*: F<sup>-</sup> *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\Phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74* *recA1* *endA1* *ara* $\Delta$ 139  $\Delta$ (*ara, leu*)7697 *galU* *galK*  $\lambda$ - *rpsL* (Str<sup>R</sup>) *nupG*;

## II - Meios de cultura

Para crescimento de leveduras:

- YEPD: meio completo composto por 1% de extrato de levedura, 2% de peptona e 2% de glicose. Para meio sólido adicionar 2% de ágar;

Para a seleção dos transformantes integrativos o meio foi complementado com o antibiótico geneticina (G418) a uma concentração de 200 mg/L e sorbitol na concentração de 1 M.

- YNB: meio mínimo composto por 0,67% YNB e 3% de glicose, sendo necessário suplementar com aminoácidos para a seleção de transformantes. Os aminoácidos triptofano, histidina e leucina são preparados à parte em uma solução estoque com concentração de 100X e a uracila, 50X. É adicionada também uma mistura de aminoácidos (*drop out*) ao meio (Ausubel *et al.*, 1998);

Para crescimento de bactérias:

- LB: meio composto por 1% tripton, 0,5% extrato de levedura e 0,5% cloreto de sódio (NaCl) complementado com os antibióticos apropriados para seleção de transformantes (ampicilina 100 µg/ml).

Para a preparação dos meios sólidos foram adicionados 2% de ágar.

## III – Plasmídeos

- pYADE4 – plasmídeo de expressão em levedura, contendo o promotor do gene da álcool desidrogenase 2 (*ADH2*) a frente de um sítio de clonagem múltipla (Brunelli & Pall, 1993);
- pYADE4-*FLO5* – plasmídeo que contém o gene *FLO5* sob controle do promotor *ADH2* (Cunha *et al.*, 2006);
- YEp352-*FLO10* – plasmídeo YEp352 contendo o gene *FLO10*. Contém marca de seleção *URA3* (uracila) e *Amp<sup>R</sup>* (resistência a ampicilina).

#### IV – Extração de DNA genômico de levedura

Para a extração do DNA genômico, as leveduras foram crescidas em meio YEPD durante 16 horas a 30 °C sob agitação de 300 rpm. As células foram separadas do meio de cultura e extração foi feita utilizando um tampão de quebra com bolinhas de vidro (*glass beads*) e uma solução de fenol/clorofórmio/álcool isoamílico. Essa mistura foi “vortexada” e centrifugada. O DNA foi precipitado com 1/10 de NaAc (3 M, pH 5,3) e 2 x volume de etanol 100% (Ausubel *et al.*, 1998).

#### V – Amplificação e clonagem do gene FLO10

O gene *FLO10* foi amplificado por PCR diretamente do DNA genômico de *Saccharomyces cerevisiae* usando *primers* específicos:

- (1) FLO10BamF: 5' CGCGGA TCCATGCCTGTGGCTGCTCG 3' (62°C)
- (2) FLO10SalR: 5' ACGCGTCGACCTATTAAACGATTGCCAG 3' (62°C)
- (3) FLO10PROMSAMA1: 5' TCCCCCGGGGGAGGTTGTTTGGTATGTATC 3' (50°C)
- (4) FLO10BMHI: 5' CGCGGATCCGCGTTAAACGATTGCCAGTAATAG 3' (50°C)
- (5) FLO10SMAIprcurto: 5' TCCCCCGGGGGATTCTGTTGTAACGAATGATTC 3' (52°C)
- (6) FLO10\_117pr\_sma: 5' TCCCCCGGGGGACATGTTTCGAGTGCG 3'(52°C)

A seqüência sublinhada indica os sítios de restrição das enzimas para posterior digestão, e a temperatura que se encontra entre parênteses indica a temperatura de anelamento para a PCR.

A PCR foi feita da seguinte forma: 3 µl de tampão (10 X), 2,5 µl MgSO<sub>4</sub> (50 mM), 1,5 µl de cada *primer* (15 pmol), 8 µl de dNTPs (1,25 mM) e 1,5 U de taq de alta fidelidade (Invitrogen) para um volume final de 30 µl. O programa utilizado foi: um ciclo inicial de desnaturação a 94 °C por 2 minutos, seguido por 40 ciclos de 94 °C por 30 segundos, temperatura anelamento por 30 segundos, 68 °C por 2 minutos e uma extensão final de 68 °C por 7 minutos.

Os fragmentos amplificados foram clonados no vetor pGEM-T Easy (Promega), seguindo as especificações do kit.

## **VI – Reação de sequenciamento**

As reações de sequenciamento dos produtos de PCR clonados foram realizadas em volume final de 10 µl, sendo: 1,0 µl de Bigdye terminator Mix (Applied Biosystems), 5 pmol/µl de *primer* M13. Foi utilizado 400 ng de plasmídeo, que era o molde que estava clonado em pGEM-T. Completando-se o volume final com H<sub>2</sub>O milliQ. Os ciclos de sequenciamento foram: um ciclo inicial de desnaturação a 95 °C por 2 minutos, seguido por 50 ciclos de 95 °C por 20 segundos, 57 °C por 20 segundos e 60 °C por 4 minutos. A precipitação das amostras seguiu a orientação do fabricante, usando-se o Sequenciador Applied Biosystems 377-96.

## **VII - Construção dos plasmídeos**

A técnica de purificação e isolamento de DNA do gel de agarose foi feita utilizando-se o kit da Qiagen (Qiaex II Gel Extraction Kit) ou kit da Invitrogen (PureLink). Todas as outras técnicas, como a extração de DNA plasmidial (mini preparação), o uso de enzimas de restrição e ligação foram realizadas de acordo com os protocolos básicos (Ausubel *et al.*, 1998).

## **VIII – Preparação de células (bactérias e leveduras) competentes e transformação**

Todas essas técnicas seguiram os protocolos descritos em Ausubel (Ausubel *et al.*, 1998).

Para o preparo das bactérias competentes, é necessário o acompanhamento do crescimento celular em 1 L de meio LB até a OD<sub>600</sub> = 0,6 (fase exponencial de crescimento). Nesse momento o crescimento celular é parado colocando as células em banho de gelo. São feitas 3 lavagens sucessivas com glicerol gelado (10%) para retirar todo resíduo de meio de cultura. As células são ressuspensas em 2 mL de glicerol (10%), aliqüotadas e armazenadas em freezer -80 °C para uso posterior.

No preparo de leveduras competentes, as células são crescidas até OD<sub>600</sub> = 1,3. São feitas sucessivas lavagens em Sorbitol (1 M) para retirar resíduos de meio. No final, as células são ressuspensas em 200 µL de sorbitol e aliqüotadas para uso. Nesse caso as

células foram utilizadas imediatamente após o preparo. O armazenamento, nesse caso, diminui o rendimento da transformação.

A transformação, tanto de leveduras e bactérias foi realizada pelo método de eletroporação.

### **IX - Extração e análise de RNA**

O RNA total foi extraído pelo método do fenol ácido quente e o *Northern blotting* foi feito segundo protocolos básicos (Ausubel *et al.*, 1998).

Na extração de RNA, as células coletadas durante os experimentos foram tratadas com fenol ácido, tampão e aquecidas a 65 °C por 1 hora. As etapas adicionais de extração foram feitas novamente com fenol ácido e clorofórmio. O RNA foi precipitado com NaAc (3 M, pH 5,3) e etanol 100% gelado.

Para os experimentos de *Northern blotting*: Foi feita eletroforese do RNA extraído em gel de agarose denaturante para que as fitas de RNA permanecessem simples e não se ligassem entre si. O RNA presente no gel foi transferido para uma membrana Hybond-N+ (Amersham Pharmacia Biotech) e posteriormente hibridado com uma sonda específica marcada radioativamente. Esse material foi exposto em um filme e revelado.

### **X – Determinação da concentração de glicose**

A quantificação da concentração de glicose foi feita com um kit colorimétrico enzimático da Laborlab. Os resultados foram expressos em porcentagem de glicose.

Os cálculos seguem abaixo:

$$F = \frac{100\text{mg/dL}}{P} \qquad \% \text{ de glicose} = \frac{D \cdot F \cdot \text{diluição}}{1000}$$

Sendo que: F = constante dependente do padrão

P = OD<sub>505</sub> do reagente padrão

D = OD<sub>505</sub> da amostra desconhecida

## RESULTADOS E DISCUSSÃO

### Clonagem do gene *FLO10* variando tamanho do promotor e construção do plasmídeo expressão

O gene *FLO10* é um gene pouco estudado e por isso não havia na literatura nenhuma informação sobre a sua região promotora. Para verificar se o tamanho da região promotora teria interferência na regulação, o promotor de *FLO10* foi sendo reduzido através da amplificação utilizando *primers* cada vez mais internos e utilizado na sua forma nativa ou fusionado com o a região regulatória do promotor do gene *ADH2*.

O gene *FLO10* foi amplificado por PCR com 4 variações no tamanho do promotor, utilizando os pares de *primers* descritos abaixo. Destacado entre parênteses estão os nomes dados a cada um dos fragmentos:

- *FLO10*BamF e *FLO10*SalR: amplifica somente a parte codante do gene, ou seja, a partir do ATG gerando um fragmento de 3510 pb (*FLO10*). Nesse caso, o gene é fusionado com o promotor completo do gene *ADH2*;

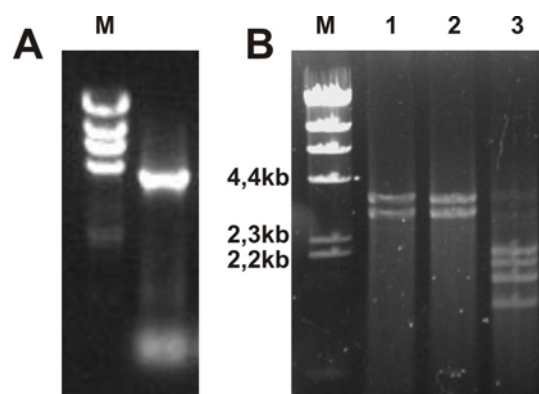
- *FLO10*PROMOSMAI e *FLO10*BMHI: amplifica um fragmento de 3834 pb (*FLO10*pr), sendo que os 324 pb iniciais são referentes ao promotor do gene *FLO10*;

- *FLO10*SMAIprcurto e *FLO10*BMH: amplifica um fragmento de 3717 pb (*FLO10*pr\_curto), sendo 207 pb referente ao promotor do gene *FLO10*;

- *FLO10* 117pr Sma e *FLO10*BMH: amplifica um fragmento de 3627 pb (*FLO10*\_117pr) com o promotor contendo 117 pb.

Os fragmentos amplificados foram clonados no vetor pGEM-T easy (Promega). Em todos os casos a clonagem foi confirmada pelo padrão de bandas por digestão, PCR com *primers* específicos e sequenciamento. A figura 11 mostra a amplificação do gene *FLO10* e a confirmação da construção pGEM-*FLO10*. Nesse caso, a digestão com a enzima *Sall* gera fragmentos de 3,5 kb e 3,0 kb; com *BamHI/Sall*, 3,5 kb e 3,0 kb, e com *Scal/Sall*, 2,0 kb, 1,8 kb, 1,5 kb e 1,1 kb.

Como descrito em materiais e métodos, os *primers* utilizados para amplificação do gene *FLO10* foram desenhados com a seqüência de determinadas enzimas de restrição, que foram utilizadas, posteriormente, para retirar o gene do vetor de clonagem pGEM.



**FIGURA 11.** (A) Amplificação por PCR do gene *FLO10*. (B) Digestão do pGEM-*FLO10* com as enzimas (1) *Sall*, (2) *BamHI/Sall* e (3) *ScaI/Sall*. M: marcador de peso molecular  $\lambda$  hind.

Para a construção dos plasmídeos de expressão, o pGEM com os genes clonados foram digeridos com enzimas de restrição. A banda correspondente ao gene *FLO10* foi purificada (Kit PureLink - Invitrogen) e ligada ao plasmídeo pYADE4 (que contém o promotor do gene *ADH2*), previamente linearizado também por tratamento com enzimas, de acordo com tabela 2:

**TABELA 2:** Plasmídeos com os genes clonados, as enzimas utilizadas para liberar o gene, as bandas obtidas após a digestão, a banda isolada e as enzimas utilizadas para digerir o vetor de expressão.

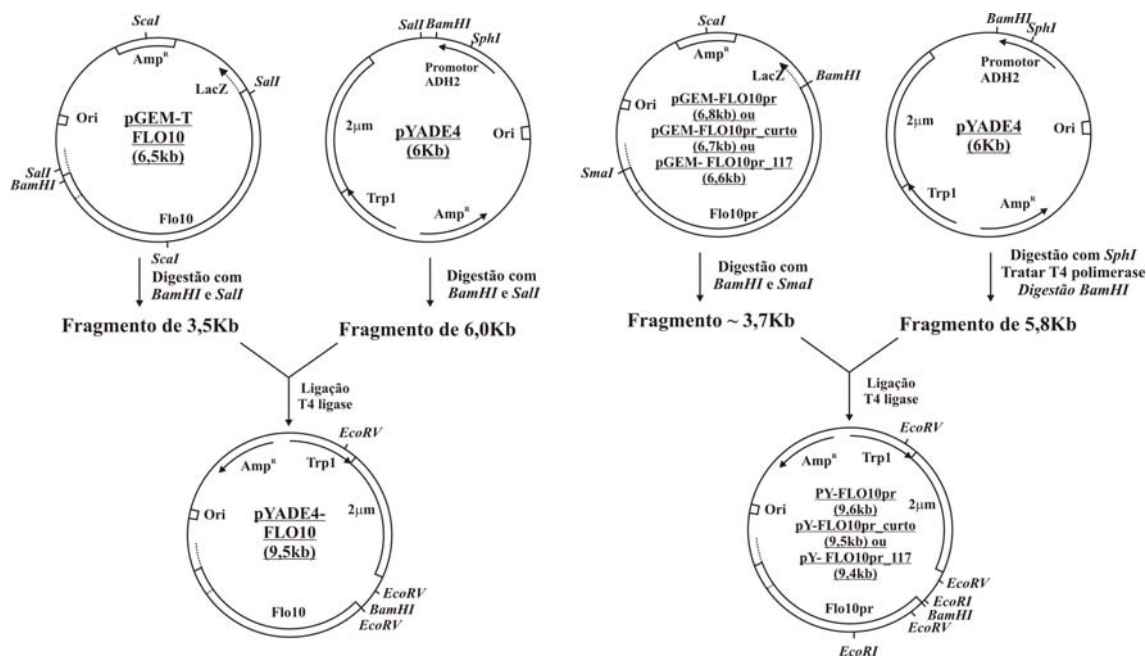
Plasmídeo	Enzimas para digerir pGEM	Bandas Obtidas	Banda Isolada e Purificada	Enzima para digestão PYADE4
pGEM-FLO10	<i>BamHI</i> e <i>Sall</i>	3,0 e 3,5 Kb	3,5 Kb	<i>BamHI</i> e <i>Sall</i>
pGEM-FLO10pr	<i>SmaI</i> e <i>BamHI</i>	3,0 e 3,8 Kb	3,8 Kb	<i>SphI</i> e <i>BamHI</i>
pGEM-FLO10pr curto	<i>SmaI</i> e <i>BamHI</i>	3,0 e 3,7 Kb	3,7 Kb	<i>SphI</i> e <i>BamHI</i>
pGEM-FLO10pr_117	<i>SmaI</i> e <i>BamHI</i>	3,0 e 3,6Kb	3,6 Kb	<i>SphI</i> e <i>BamHI</i>

O plasmídeo pYADE4 para ser ligado a gene *FLO10* sem o promotor foi linearizado com *BamHI* e *Sall*, ou seja, deixando o promotor do gene *ADH2* inteiro. Para a reação de ligação usou-se a enzima T4 ligase (Invitrogen), e a concentração molar do inserto utilizado foi 5 vezes maior que a do plasmídeo.

Para a construção do gene *FLO10* sob o controle da parte regulatória do promotor do gene *ADH2*, o plasmídeo pYADE4 foi linearizado com a enzima *SphI*, gerando extremidades coesivas, cortando o plasmídeo na região do promotor do gene *ADH2*. Esse

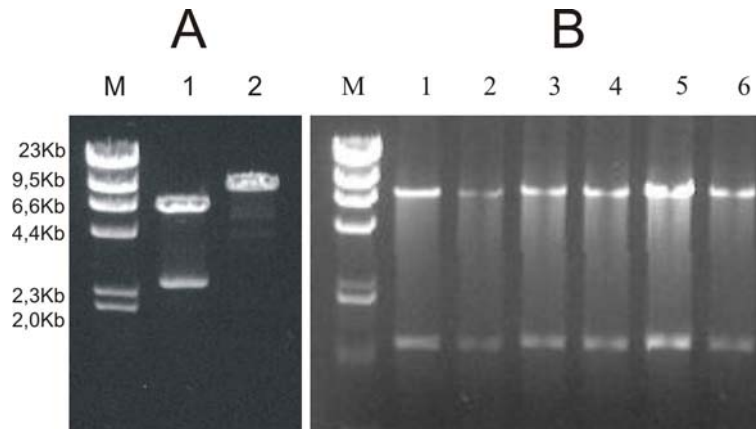


material foi tratado com a enzima T4 DNA polimerase, deixando com que as extremidades se tornem abruptas. Em seguida foi feita uma nova digestão com a enzima *Bam*HI. Esta digestão remove o TATA box do promotor do gene *ADH2*, mantendo no plasmídeo apenas a região regulatória (Beier *et al.*, 1985; Cunha *et al.*, 2006). Posteriormente o gene *FLO10* com o promotor foi ligado ao plasmídeo PYADE4 utilizando a enzima T4 ligase (Invitrogen). O esquema das construções está mostrado na figura 12.

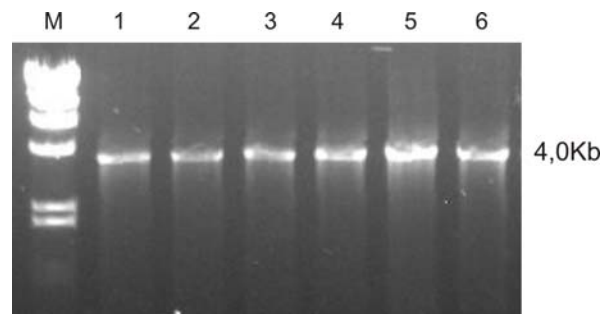


**FIGURA 12.** Esquema da construção dos plasmídeos pYADE4-*FLO10*, pYADE4-*FLO10*pr, pYADE4-*FLO10*pr\_curto e pYADE4-*FLO10*pr\_117. Os plasmídeos (pGEM) foram digeridos com as enzimas restrição, liberando um fragmento de aproximadamente 3,7 Kb, que foi ligado ao plasmídeo pYADE4 linearizado previamente digerido.

A confirmação da construção correta dos plasmídeos de expressão foi feita por digestão com as enzimas de restrição (Figura 13) e por PCR com os *primers* específicos (Figura 14).



**FIGURA 13.** (A) Foto do gel de agarose com a digestão do plasmídeo pYADE4-*FLO10*pr com as enzimas: (1) *EcoRV*, produzindo bandas de 6,9 Kb, 2,5 Kb e 145 pb (banda não visível); (2) *BamHI*, linearizando o plasmídeo (9,6 Kb). (B) Digestão com *EcoRI* para confirmar a construção dos plasmídeos pYADE4-*FLO10*pr\_curto (1-3) e pYADE4-*FLO10*pr\_117 (4-6). O padrão de banda esperado é de ~1,4 kb e ~8,2 kb. M = marcador de peso molecular o  $\lambda$ hind (Invitrogen).



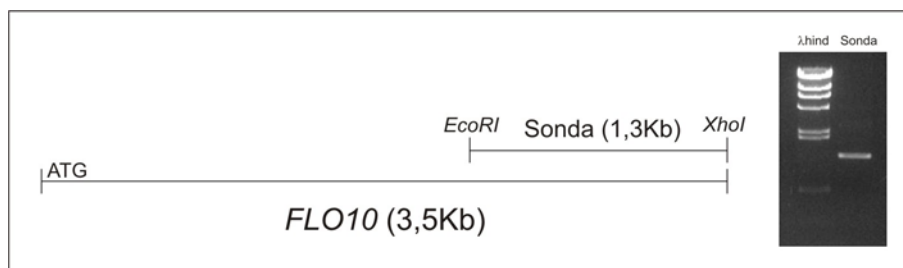
**FIGURA 14:** Produto da reação de PCR para confirmar a construção utilizando-se os *primers* ADH2F/*FLO10*BMHI. Foi obtido o padrão de bandas esperado (4,0Kb). As amostras referentes aos plasmídeos pYADE4-*FLO10*pr\_curto (1-3) e pYADE4-*FLO10*pr\_117 (4-6).

### **Transformação de leveduras de laboratório e testes de floculação**

A linhagem de laboratório W303-1a foi transformada por eletroporação com os plasmídeos: pYADE4 (controle negativo), pYADE4-*FLO10*, pYADE4-*FLO10*pr, pYADE4-*FLO10*pr\_curto, pYADE4-*FLO10*pr\_117.

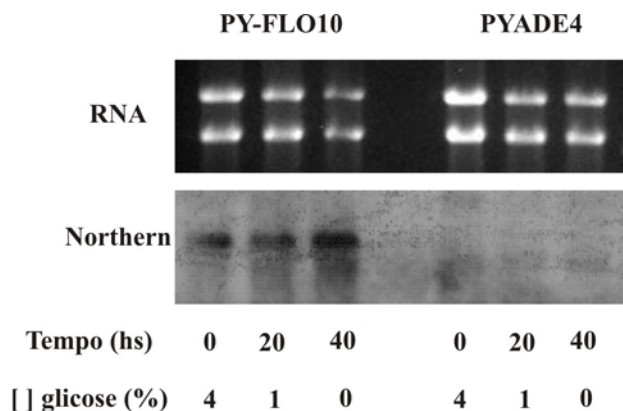
As amostras de células foram coletadas periodicamente até o esgotamento da glicose para posterior extração de RNA, foram feitas medidas de densidade óptica (OD), concentração de glicose, presença ou ausência de floculação (observação visual).

Para os experimentos de *Northern Blot* foi necessária a construção de uma sonda. O fragmento da sonda, contendo 1,3 kb, referente ao gene *FLO10* foi isolado da digestão do plasmídeo pYADE4-*FLO10* com as enzimas *EcoRI* e *XhoI* (Figura 15) e foi marcado radioativamente usando  $\alpha^{[32]}\text{P}$ -dATP (Ausubel *et al.*, 1998).



**FIGURA 15:** Esquema da região de onde a sonda do gene *FLO10* (1,3 Kb) foi retirada após a digestão do plasmídeo pYADE4-*FLO10* com as enzimas *EcoRV* e *XhoI* e foto do gel de agarose contendo a banda isolada.

O primeiro resultado de *Northern blot* com linhagens transformada com o plasmídeo pYFLO10 (*FLO10* clonado a partir do ATG) mostra que o promotor do gene *ADH2* não ativa somente quando a glicose se esgota, mas há uma expressão basal mesmo com quantidades elevadas de glicose (Figura 16). Esse mesmo fenômeno foi observado com o gene *GFP3* (green fluorescent protein) na construção pYADE4-*GFP3* realizado pelo Prof. Luiz Humberto do Laboratório de Genética da ESALQ, Piracicaba (comunicação pessoal).

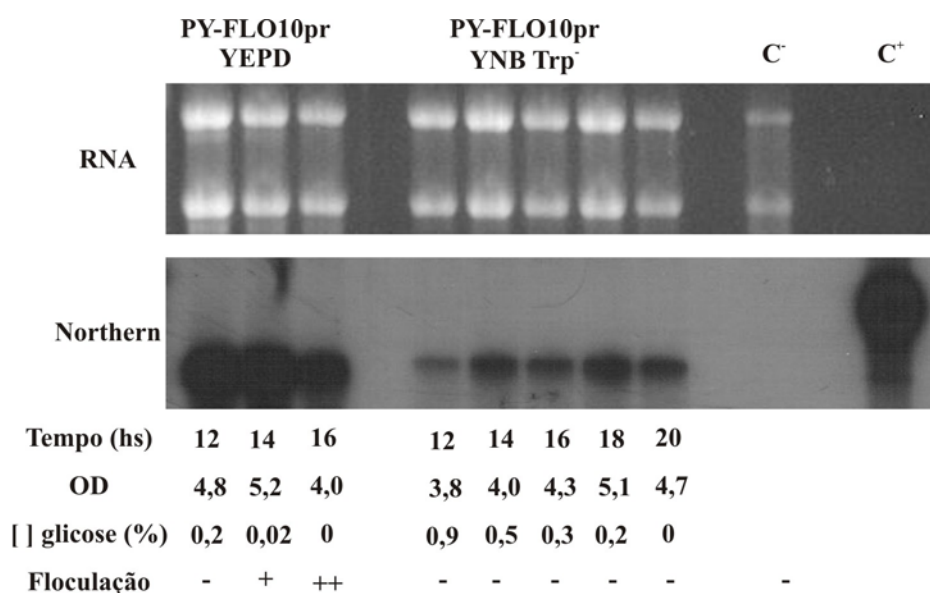


**FIGURA 16.** Experimento de *Northern blotting* mostrando a expressão basal do gene *FLO10* durante ensaio de fermentação. Nas linhagens transformadas com o plasmídeo PYADE4 não se observa a expressão do gene, como esperado.

Com a expressão basal do gene, partimos para as outras construções, então linhagens de laboratório (W303-1a) foram transformadas com o plasmídeo pYFLO10pr, na

qual apenas a região regulatória do promotor do gene *ADH2* foi fusionada com o promotor básico do gene *FLO10*.

O experimento de *Northern* continuou apresentando expressão basal do gene. As células crescidas em meio completo (YEPD) começaram a flocular podendo ser identificadas visualmente após o esgotamento da glicose. A intensidade da floculação aumentava um pouco a medida que a glicose estava sendo consumida, mas isso pode ser decorrente também do aumento da massa celular. Já as células crescidas em meio YNB- $\text{Trp}^-$  não flocularam em nenhum momento, nem após o esgotamento da glicose (Figura 17).



**FIGURA 17:** Experimento de *Northern blotting* de linhagens transformadas com o plasmídeo pYADE4-*FLO10pr* crescidas em meio YEPD e YNB- $\text{Trp}^-$ . (C<sup>+</sup>) controle positivo, gene *FLO10* amplificado por PCR; (C<sup>-</sup>) controle negativo, RNA de leveduras transformadas com o plasmídeo pYADE4.

As células crescidas em meio YNB- $\text{Trp}^-$  apresentaram uma intensidade de expressão do gene menor do que as linhagens crescidas em YEPD (Figura 17). Govender (Govender *et al.*, 2008) observou também que transformantes para o gene *FLO1* e *FLO5* crescidos em meio mínimo apresentam capacidade de floculação diminuída. Nesse caso, o RNAm para pode estar sendo expresso, mas a proteína pode não estar ativa. Pelo fato do meio de cultura ser pobre, algum elemento essencial para ativar a proteína, como por exemplo, o cálcio, pode funcionar como um fator limitante.

As células floculadas resultante da expressão do gene *FLO10* foram lavadas com uma solução de EDTA (quelante de íons) e observou-se a desfloculação de todas as células, ou seja, ficaram todas em suspensão. Isso indica que a floculação é dependente de íons  $\text{Ca}^{++}$  como no gene *FLO1*.

Para verificar se o tamanho do promotor teria influencia na expressão gênica, as leveduras foram transformadas com os plasmídeos pYADE4-*FLO10*pr\_curto, pYADE4-*FLO10*pr\_117. Os dados de tempo de fermentação, OD, concentração de glicose e floculação se encontram na tabela 3. Foram feitos com os 3 plasmídeos construídos pY-*FLO10*pr, pY-*FLO10*pr\_curto, pY-*FLO10*pr\_117, e também com um controle negativo, pYADE4, e positivo, YEP*FLO10*pr.

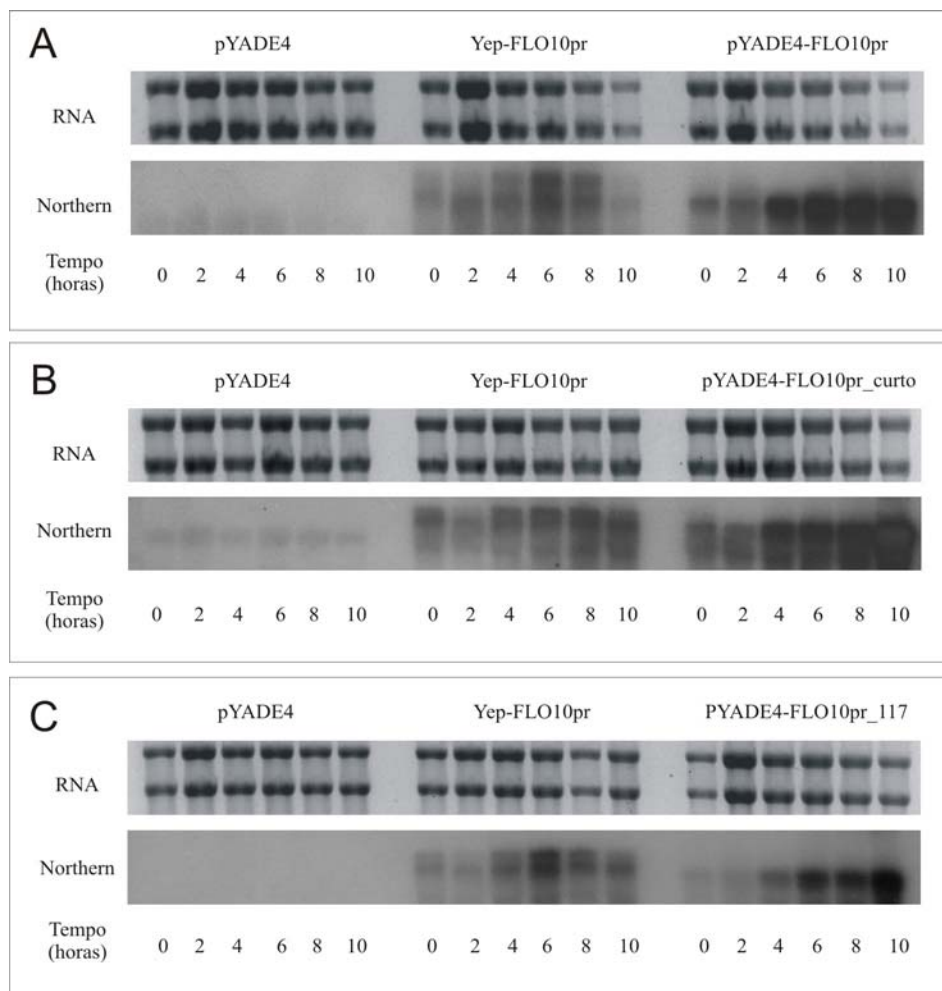
Como observados no experimento de *Northern* (Figura 18), as leveduras transformadas com os três plasmídeos construídos apresentaram floculação basal durante todo o processo fermentativo. A intensidade da expressão gênica aumenta significativamente à medida que a glicose vai sendo consumida, principalmente no último ponto de coleta (10hs), quando já não há mais glicose presente no meio de cultura.

É interessante observar que as leveduras transformadas com o plasmídeo pYADE4-*FLO10*pr\_117 apresentaram o mesmo padrão de expressão gênica das transformadas com as demais construções, ou seja, a expressão se mostrou basal aumentando nas etapas finais da fermentação (Figura 18 C). Mas quando se observaram as células, elas não apresentaram característica floculante. Visualmente, o fenótipo dessas leveduras era muito semelhantes ao fenótipo das leveduras transformadas com o plasmídeo pYADE4 (controle negativo), com as células todas em suspensão. Nessa construção, o tamanho do promotor do gene *FLO10* apresenta 117pb (que contém a região de TATA box) e o restante do promotor referente à região regulatória se encontra no pedaço do promotor do gene *ADH2*. O RNA mensageiro referente ao gene *FLO10* está sendo transcrito, como pode ser visto no experimento de *Northern*, mas a proteína provavelmente não está sendo traduzida corretamente ou após a tradução ela não esteja sendo ativada. Outra hipótese seria que nessa construção as enzimas responsáveis pela transcrição tenham encontrado outra região iniciadora para a proteína, mudando o frame, transcrevendo outra sequência.

**TABELA 3.** Testes de floculação com as leveduras da linhagem W303-1a transformadas com os diferentes plasmídeos e crescidas em meio completo YEPD.

Plasmídeo	Tempo (hrs)	OD	[ ] Glicose (%)	Floculação*	Observação
pYADE4	0	2,7	1,4	0	Plasmídeo sem o gene da floculação (C-)
	2	4,3	1,3	0	
	4	5,7	1,0	0	
	6	7,0	0,7	0	
	8	8,1	0,2	0	
	10	8,3	0,04	0	
YEPFLO10p	0	2,7	1,3	1	Gene FLO10 constitutivo (C+)
	2	4,1	1,3	1	
	4	5,5	1,1	1	
	6	6,7	0,8	1	
	8	7,9	0,2	1	
	10	8,3	0,04	1	
pYFLO10pr	0	2,4	1,3	1	Promotor do gene FLO10 com 324pb
	2	3,8	1,3	1	
	4	5,4	1,1	1	
	6	6,3	0,8	1	
	8	7,9	0,3	1	
	10	8,2	0,04	2	
pYFLO10pr_ curto	0	2,7	1,3	1	Promotor do gene FLO10 com 207pb
	2	4,2	1,2	1	
	4	5,4	1,0	1	
	6	6,5	0,8	1	
	8	8,0	0,2	1	
	10	8,1	0,03	2	
pYFLO10pr_ 117	0	2,7	1,3	0	Promotor do gene FLO10 com 117pb
	2	4,0	1,3	0	
	4	5,9	1,1	0	
	6	6,4	0,8	0	
	8	8,1	0,3	0	
	10	8,3	0,04	0	

\* A intensidade da floculação: (0) ausência de floculação; (1) Floculação fraca; (2) Floculação intensa.



**FIGURA 18.** Experimentos de *Northern blotting* da linhagem W303-1a transformada com o plasmídeo (A) pYADE4-*FLO10pr*, (B) pYADE4-*FLO10pr\_curto* e (C) pYADE4-*FLO10pr\_117*. Linhagens transformadas com pYADE4 representam o controle negativo, e as transformadas com YEP*FLO10pr*, controle positivo.

Em nenhuma das outras duas construções (pYADE4-*FLO10pr* e pYADE4-*FLO10pr\_curto*) foi obtida uma regulação controlada pelos níveis de glicose no meio como era esperado. Observando na figura 18 A e B, nos experimentos de *Northern*, os transcritos obtidos apresentam o tamanho um pouco menor do que o controle (Yep-*FLO10pr* – flocula o tempo todo). Isso indica que pode estar havendo uma regulação pós-transcricional ou que o promotor do gene *FLO10* tenha algum elemento que esteja impedindo o controle exato da floculação no momento que a glicose se esgote.

### Construção dos cassetes integrativos e integração no genoma

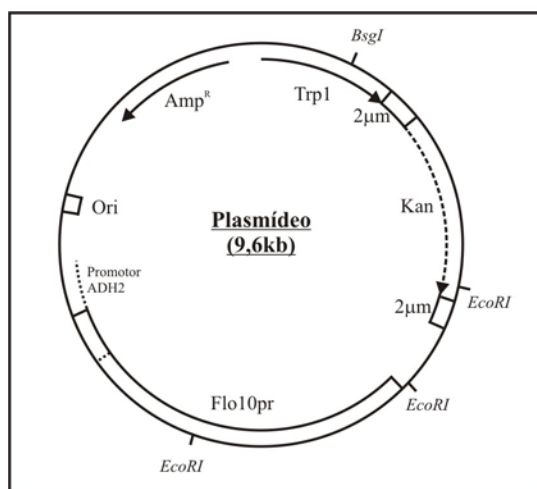
A partir dos plasmídeos pYADE4-*FLO10*pr, pYADE4-*FLO10*pr\_curto e pYADE4-*FLO10*pr\_117, foram construídos cassetes integrativos, transformando linhagens industriais (que efetivamente realizam a fermentação nas usinas) e avaliando o comportamento da expressão do gene *FLO10*.

O plasmídeo pYADE4, que é utilizado para expressão de genes em leveduras, possui uma região de 2  $\mu$ m, que confere a habilidade desse plasmídeo de se replicar com autonomia dentro da célula de levedura (Griffiths *et al.*, 1993). Para a construção dos cassetes integrativos, a região de 2  $\mu$ m foi substituída pelo gene *KAN* (Figura 19), que confere resistência ao antibiótico G418.

A retirada da região de 2  $\mu$ m foi necessária para garantir que o plasmídeo não continuasse funcional no caso de sua recircularização. Houve também a necessidade de inserir o gene *KAN* para selecionar as linhagens industriais transformadas, pois elas não possuem marcas auxotróficas como as linhagens de laboratório.

Para que ocorresse a recombinação homóloga, o plasmídeo foi linearizado com a enzima *BsgI*. Esta corta a região central do gene triptofano. É nesse gene que o cassete irá se integrar no genoma da levedura.

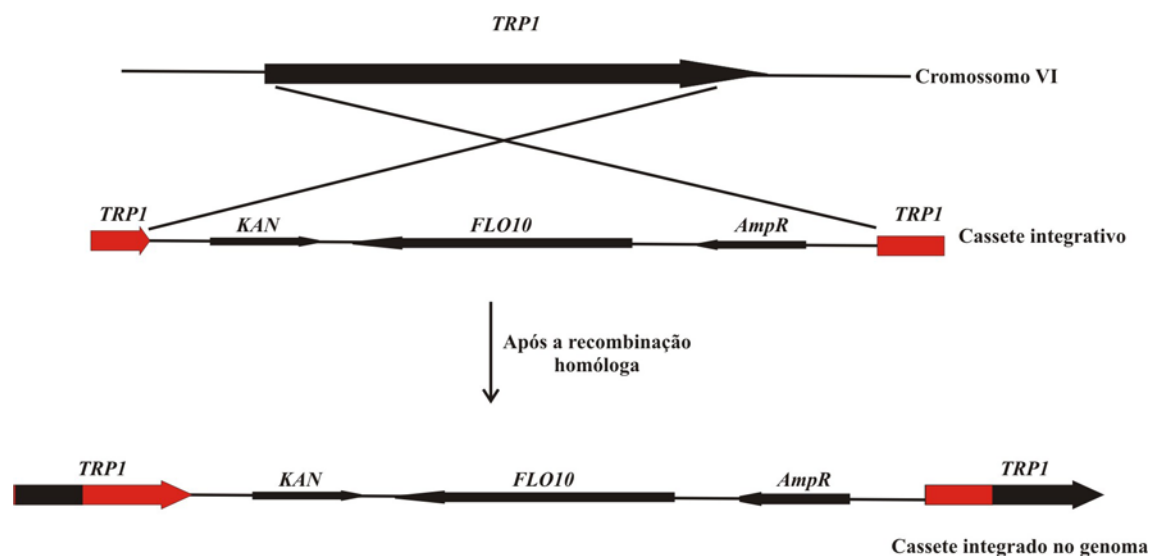
As leveduras têm a capacidade de reconhecer sequências homólogas de DNA livres dentro da célula e promover a recombinação homóloga. Esse é um fenômeno muito comum entre as leveduras.



**FIGURA 19:** Esquema padrão do plasmídeo integrativo contendo o gene *FLO10* sob regulação do promotor do gene *ADH2* e com parte da região de 2  $\mu$ m substituída pelo gene *KAN*. O plasmídeo foi linearizado com a enzima *BsgI* para transformar leveduras.



Quando se insere o fragmento de DNA linearizado dentro da célula, as extremidades desse fragmento, que apresentam a sequência do gene triptofano, se pareiam com esse mesmo gene presente no genoma da levedura no cromossomo VI onde ocorre a recombinação homóloga. Nesse caso, como mostrado no esquema da figura 20, observa-se que após a integração do cassete, o gene *TRP1* não é deletado ou, ele ficará duplicado sendo que as duas cópias permanecem funcionais. Além disso, as leveduras ficam resistentes à ampicilina e à geneticina (G418), genes presentes no cassete.



**FIGURA 20:** Esquema da integração do cassete no genoma da levedura. Observa-se a duplicação do gene triptofano.

As transformações com os cassetes integrativos (pYADE4-*FLO10*pr\_Kan, pYADE4-*FLO10*pr\_curto\_Kan e pYADE4-*FLO10*pr\_117\_Kan) nas linhagens FY23 (laboratório), JAY270, JAY291 e JAY292 foram feitas por eletroporação.

Após a eletroporação, as leveduras foram ressuspensas em meio YEPD com sorbitol e deixadas crescer durante toda a noite para serem plaqueadas em meio seletivo (YEPD + G418). O sorbitol ajuda na estabilização da membrana que sofreu uma desestabilização após a eletroporação. Antes de se plaquear é necessário esse tempo para que as células possam se multiplicar e aumentar a probabilidade de que ocorra a recombinação homóloga.

Após 3 dias na estufa a 30°C, observou-se crescimento de colônias em todas as linhagens transformadas. Confirmou-se a integração do cassete, na região do gene triptofano, por PCR.

Algumas colônias selecionadas foram crescidas em meio YEPD e crescidas até o total esgotamento da glicose. A tabela 4 mostra os resultados preliminares encontrados nestas linhagens. Foi feita apenas uma triagem inicial para verificar se todas as colônias expressavam o fenótipo da floculação.

A maioria das colônias obtidas não flocularam mesmo apresentando resistência ao antibiótico G418. Isso pode ter acontecido devido a algum erro na recombinação que eliminou a funcionalidade do gene da floculação. A construção pYADE4-FLO10pr\_117 não consegue expressar o fenótipo floculante.

**Tabela 4:** Linhagens de leveduras transformadas com os respectivos cassetes integrativos e observações da capacidade floculante das colônias obtidas.

<b>Linhagem</b>	<b>Plasmídeo</b>	<b>*Colônia 1</b>	<b>*Colônia 2</b>	<b>*Colônia 3</b>
JAY270	Kan-FLO10pr	+	-	0
	Kan-FLO10pr curto	-	0	0
	Kan-FLO10pr_117	-	-	0
JAY291	Kan-FLO10pr	-	-	0
	Kan-FLO10pr_117	-	-	0
JAY292	Kan-FLO10pr	+	-	0
	Kan-FLO10pr curto	+	+	+
	Kan-FLO10pr_117	-	-	0
FY23	Kan-FLO10pr	+	-	0
	Kan-FLO10pr curto	-	-	0
	Kan-FLO10pr_117	-	-	-

\*intensidade de floculação de cada colônia selecionada. (-) ausência de floculação, (+) floculação fraca, não separa completamente do meio, (0) Colônia inexistente.

Os resultados de integração do cassete no genoma da levedura ainda não são conclusivos. Uma nova estratégia de integração de cassetes no genoma da levedura industrial está em andamento. A “domesticação” de linhagens industriais é o trabalho do mestrando Felipe Galzerani, que se baseia na técnica do *Delitto Perfetto*, na qual modificações genéticas são feitas de forma que, ao final de todo processo de transformação, não permaneça nenhuma marca auxotrófica ou marca de seleção. Isso é muito importante, pois diferentemente da técnica utilizada aqui, que insere o gene *KAN*, deixando na levedura

a marca de resistência ao antibiótico G418, é possível se obter uma levedura “limpa”, evitando toda a polêmica e discussão sobre organismos transgênicos.

Depois de se obter transformantes a partir dessa nova técnica de integração será feito um acompanhamento mais detalhado durante a fermentação até o esgotamento da glicose.

Há ainda a necessidade de se investigar o motivo da expressão basal de todas as construções do gene *FLO10*. Nas linhagens de laboratório, apesar de ter apresentado expressão basal nos experimentos de *Northen*, a expressão aumenta após o esgotamento da glicose e é possível se observar o aumento da floculação.

Uma outra possibilidade seria a busca de um outro promotor que seja expresso somente ao final da fermentação. Essa informação poderá ser obtida no trabalho do doutorando Osmar Netto, que está avaliando a expressão gênica de linhagens industriais durante o processo fermentativo na condições reais de dornas. Essa expressão gênica será comparada com linhagens de laboratório, investigando também diferenças que possam dar indícios da maior robustez e eficiência da linhagem industrial.

A possibilidade de se controlar a floculação e a desfloculação é de grande importância biotecnológica. Novos promotores serão testados, para se tentar melhorar o controle da expressão do gene da floculação.

## **CONCLUSÃO**

O gene *FLO10*, nos experimentos de *Northen*, mostrou uma expressão basal desde o início da fermentação, entretanto essa expressão é aumentada após o esgotamento da glicose.

Não foi observado o fenômeno da desfloculação.

## **CAPITULO 3**

### **AUMENTO DA PRODUÇÃO DE ETANOL EM LINHAGENS DE LEVEDURAS INDUSTRIAIS BRASILEIRAS**

## INTRODUÇÃO

As pesquisas para o melhoramento genético de linhagens de leveduras industriais brasileiras ainda são bem escassas. As linhagens que efetivamente realizam fermentação nas usinas têm características de alta eficiência e robustez. Alterações que promovam o melhoramento da linhagem baseado no aumento na produtividade ou diminuição de problemas como, formação de espuma ou floculação, são de extremo interesse para os usineiros.

Estudos com linhagens de *Saccharomyces cerevisiae* mostraram que mutantes respiratórios apresentam um aumento na produtividade de etanol comparada com leveduras que respiram (Hutter & Oliver, 1998; Panoutsopoulou *et al.*, 2001). A deleção de um gene específico, *PET191*, responsável por uma das etapas da montagem do citocromo C, produz um mutante 100% deficiente respiratório, neste caso, um mutante nuclear. Foi verificado que esse mutante promove um aumento de 43% e 9% na produção de etanol em linhagens de laboratório e industriais de vinho, respectivamente.

Outro tipo de deficiente respiratório, o mutante citoplasmático, não possui genoma mitocondrial funcional. Esse mutante é obtido pelo tratamento das células com brometo de etídeo. O grande problema desse mutante é que o genoma mitocondrial, que está degradado, possui genes importantes para a linhagem como, por exemplo, tolerância ao etanol.

O objetivo dessa parte do trabalho é a utilização de linhagens industriais brasileiras (haplóides derivados da linhagem PE-2) para a produção de deficientes respiratórios (petites nucleares e mitocondriais) e a análise do rendimento fermentativo comparando-se com as linhagens selvagens.

Os resultados mostram uma tendência de aumento da quantidade de etanol pelas linhagens petites quando comparadas com as controle.

## MATERIAL E MÉTODOS

### Linhagens de levedura e meios de cultura

As linhagens de *Saccharomyces cerevisiae* utilizadas neste trabalho estão listadas na tabela 5. As células foram incubadas a 30 °C para crescimento em meio completo YEPD (1% extrato de levedura, 2% peptona, 2% glicose). Para o meio seletivo, foi adicionado 200 mg/L geneticina (G418). Os petites mitocondriais ( $\rho^0$ ) foram selecionados em meio YPG (1% extrato de levedura, 2% peptona, 2% glicerol). Os ensaios fermentativos foram feitos com extrato de cana-de-açúcar esterilizado, meio utilizado nas indústrias Brasileiras de etanol.

**TABELA 5.** Linhagens de *Saccharomyces cerevisiae* usadas neste trabalho.

<b>Linhagem</b>	<b>Genótipo</b>	<b>Referência</b>
FY23	Linhagem padrão S288c: Mat <i>a</i> , <i>ura3-52</i> , <i>leu2D1</i> , <i>trp1D63</i>	(Winston <i>et al.</i> , 1995)
PE-2	Linhagem industrial brasileira	LNF-Latino Americana
JAY270	Colônia isolada da PE-2	Argueso <i>et al.</i> , 2009 (Anexo I)
JAY291	Esporo da JAY270, Mat <i>a</i> , haplóide	
JAY292	Esporo da JAY270, Mat <i>alpha</i> , haplóide	

#### Produção de mutantes respiratórios (100% deficiente respiratório)

**Mutante nuclear (estratégia 1):** O cassete para deletar o gene *PET191* foi amplificado por PCR utilizando como DNA molde o plasmídeo pFA6aKanMX4 (plasmídeo contendo o gene de resistência à ampicilina e o gene KAN que confere resistência ao antibiótico G418) com os seguintes *primers*:

DsPET191-2 (rev)

5' TATTTATGTATATATTTACAGGCCAATTTTCATAAATTCGACAGCAGTATAGCGACCAG 3'

DsPET191-1 (fw)

5' ACAAATAAATTTGCAGTGCGTTTGTAGATATCAAAGAGCTTGCCTCGTCCCCGCCGGGT 3'

A região não sublinhada é a que apresenta homologia com o DNA molde e a região sublinhada é homóloga ao começo a ao fim do gene *PET191*, local onde ocorrerá a recombinação homóloga. Essa PCR amplifica um fragmento de 1472 pb (Figura 16).

As condições da PCR foram: 5 µl de tampão (10 X), 1,5 µl MgCl<sub>2</sub> (50 mM), 1,25 µl de cada *primer* (15 pmol), 9 µl de dNTPs (1,25 mM) e 2,5 U de Taq DNA polimerase (Invitrogen) para um volume final de 50 µl. O programa utilizado foi: um ciclo inicial de desnaturação a 94 °C por 3 minutos, seguido por 35 ciclos de 94 °C por 30 segundos, 60 °C

por 30 segundos e 72 °C por 1 minuto e 30 segundos, além da extensão final de 72 °C por 4 minutos.

**Mutante nuclear (estratégia 2):** A produção desses petites foi realizada deletando um gene específico, o gene *PET191*, através do método do double-joint PCR (Yu *et al.*, 2004). Com essa metodologia, é possível aumentar a região de homologia facilitando a recombinação na região em que se pretende deletar o gene. Para isso, 3 etapas de PCR são necessárias para a construção do cassete integrativo. A figura 21 mostra os diversos passos envolvidos nesta técnica. Os *primers* utilizados foram: **KAN-F\_dj:** 5' CGTTAGAACGCGGCTACAAT 3', **KAN-R\_dj:** 5' GCCGATTCATTAATG CAGGT 3', **5'PET-F:** 5' AGGCAGTTTACTGGCGTGTC 3', **5'PET-R-KAN:** 5' TGTATTAATTGTAGCCGCGTTCTAACGTGCTACCCCTCAGACAGCTT 3', **3'PET-F-KAN:** 5' CAGGTAACTGCATTAATGAATCGGCAGGGCGATGCAAT AAAAGTG 3', **3'PET-R:** 5' ACGTTGCAGAGCGGTCTTAC 3', **PET-KAN\_nst-F:** 5' CGACCGCAGTACAGTCTGAT 3', **PET-KAN\_nst-R** 5' GCGATTGTAGGTTGC TAAAG 3'.

Na primeira etapa (I), três PCRs independentes são feitas. Uma amplifica o gene de resistência (1), no caso, o gene *KAN*; a outra amplifica a região 5' do gene que se pretende deletar (2); e a última amplifica a região 3' do gene que se quer deletar (3). Estas duas últimas PCRs, são feitas com *primers* contendo uma cauda homóloga ao gene *KAN*.

Na segunda etapa, os três produtos da primeira etapa (1, 2 e 3) são adicionados em uma mesma reação de PCR. Os fragmentos 2 e 3 serão os “*primers*”, estendendo os fragmentos.

Para finalizar, a terceira etapa consiste em utilizar *primers* internos (*nested*) para selecionar apenas os fragmentos completos, contendo o gene *KAN* flanqueado com a região de homologia ao gene *PET191*. Os parâmetros de todas as reações de PCR foram as mesmas descritas no trabalho original (Yu *et al.*, 2004).

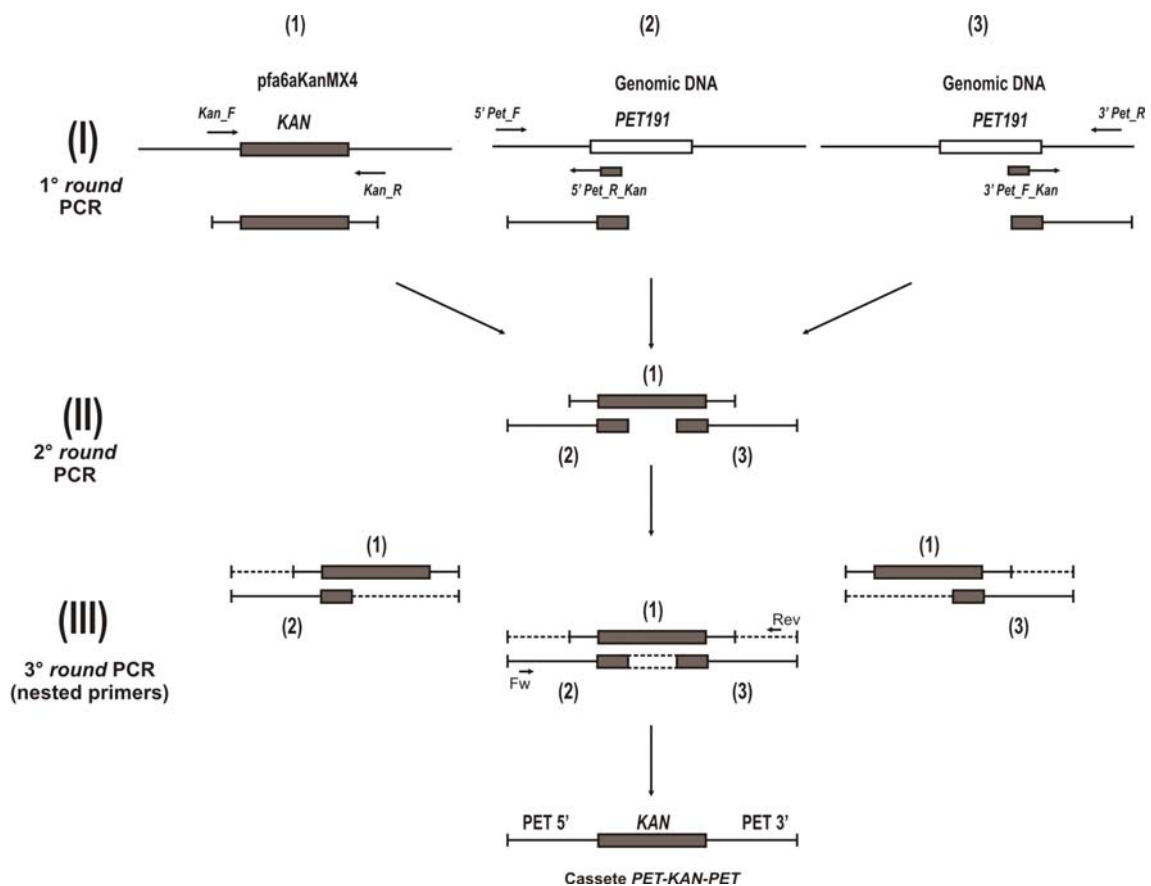
Ambos os cassetes construídos foram utilizados para transformar leveduras por eletroporação.

Para verificar se o cassete (PET-KAN-PET) se recombinou, integrando-se corretamente no cromossomo da levedura, foram desenhados os seguintes *primers* que flanqueiam o gene *PET191*:

- JAO262 - 5' CTTAGGCAGTTTACTGGCGTG 3'

- JAO263 - 5' AAGCACGTTA ACTCCCAGATC 3'

As condições utilizadas foram: 1 µl de tampão (10 X), 0,7 µl MgCl<sub>2</sub> (50 mM), 0,5 µl de cada *primer* (15 pmol), 2 µl de dNTPs (1,25 mM) e 0,5 U de Taq DNA polimerase (Invitrogen) para um volume final de 10 µl. O programa utilizado foi: um ciclo inicial de desnaturação a 94 °C por 2 minutos, seguido por 35 ciclos de 94 °C por 30 segundos, 51 °C por 30 segundos e 72 °C por 1 minuto e 30 segundos, além da extensão final de 72 °C por 5 minutos.



**FIGURA 21.** Metodologia *double-joint* PCR usado para construir os cassetes integrativos. Três etapas de PCR são necessárias para fazer o cassete. (I) Uma PCR para amplificar o gene marcador (*KAN*) e outros dois para amplificar a região do DNA genômico na região 3' e 5' do gene que se pretende deletar usando *primers* com cauda homologa ao gene *KAN*. (II) O produto das três PCRs anteriores são adicionadas numa nova reação para a segunda etapa. (III) Última PCR com *primers* internos.



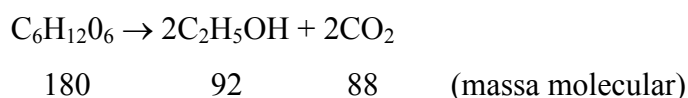
**Mutante citoplasmático ou mitocondrial ( $\rho^0$ ):** A deleção completa de todo mtDNA gera esse tipo de mutante. As células, nesse caso, são tratadas com brometo de etídio, degradando todo o DNA mitocondrial. As leveduras foram incubadas em 10  $\mu\text{g/ml}$  de brometo de etídio por aproximadamente 8 horas (Fox *et al.*, 1991). Posteriormente, as células foram plaqueadas em meio completo (YEPD). Algumas colônias foram colocadas em meio YPG para verificar a deficiência respiratória já que as células *petites* são incapazes de crescer e utilizar glicerol como fonte de carbono.

Para confirmar se todos os mutantes produzidos realmente eram 100% deficientes respiratórios, foi utilizado um reagente, o TTC (cloridrato de 2.3.5 trifeniltetrazolio). O TTC é um sal incolor que forma um precipitado vermelho quando é reduzido. As amostras celulares são espalhadas sobre a superfície de meio sólido na placa (YEPD com ágar) e incubados por 2 dias a 30 °C sob condições aeróbicas. Depois desse tempo, a placa com as colônias crescidas são cobertas com 20 mL de solução TTC à 50 °C (ágar 0,75%; glicose 0,5%; TTC 0,05%) e incubadas por 3 horas a temperatura ambiente. Colônias que apresentarem coloração rosa têm a capacidade de respirar, enquanto as colônias brancas são deficientes respiratórios (Allen, 1994).

### Ensaio fermentativos

Os ensaios fermentativos foram compostos por 4 fermentações consecutivas e em triplicata. Cada amostra foi montada com um volume final de 40 ml, utilizando 10% de massa fresca (úmida) de fermento, ou seja, 4 g de massa fresca e 36 ml de meio de fermentação. O meio utilizado foi caldo de cana esterilizado. A fermentação teve duração de 15 horas a uma temperatura de 30 °C e sem agitação. O pH foi acertado para 4,0. A quantidade inicial de açúcares redutores total (ART%) foi iniciada com aproximadamente 18%.

Para se saber a quantidade máxima teórica de etanol que pode ser obtida a partir da glicose inicial, faz-se o cálculo baseado na equação:



O valor de 180 equivale 100% da glicose que pode ser convertida em etanol. Desses 180, apenas 92 podem ser efetivamente convertidos em etanol, ou seja,  $51,111... \% = 0,51111$ . Esta é a taxa de conversão utilizada nos cálculos de rendimento teórico máximo. Para o cálculo, multiplica-se o valor de ART% (que nesse caso é de 18,0%) pela taxa de conversão (0,5111). Resultando no valor de 9,20.

Para a análise dos resultados e construção dos gráficos, os dados da concentração de etanol foram normalizados em função da quantidade máxima teórica que pode ser convertida em etanol e que está diretamente ligada à quantidade inicial de Açúcares Redutores Totais (ART%).

A medida de ART% foi feita segundo a metodologia de Somogyi-Nelson (Nelson, 1944; Somogyi, 1945). Foram preparados os reativos de Somogyi, o reativo de Nelson e uma solução padrão de glicose. À cada amostra foi adicionada 1 mL de reativo de Somogyi e colocado em banho-maria por 10 min. Depois de frio, adicionou-se 1 mL do reativo de Nelson. As amostras foram medidas em espectrofotômetro a 540 nm.

A medida de álcool foi feita por destilação e densitometria. O produto fermentado foi destilado e posteriormente foi medido em um densímetro.

## **RESULTADOS E DISCUSSÃO**

Há ainda poucos estudos visando melhorar linhagens industriais, seja por genética clássica ou engenharia genética, principalmente, as linhagens brasileiras produtoras de etanol. A grande parte dos estudos está focada no melhoramento da cana-de-açúcar e no processo durante a fermentação. As leveduras, que efetivamente convertem a sacarose em etanol, têm recebido pouca atenção. Existe, portanto, uma ampla área a ser estudada e explorada, para poder aumentar ainda mais a produção de etanol e para a síntese de novos compostos.

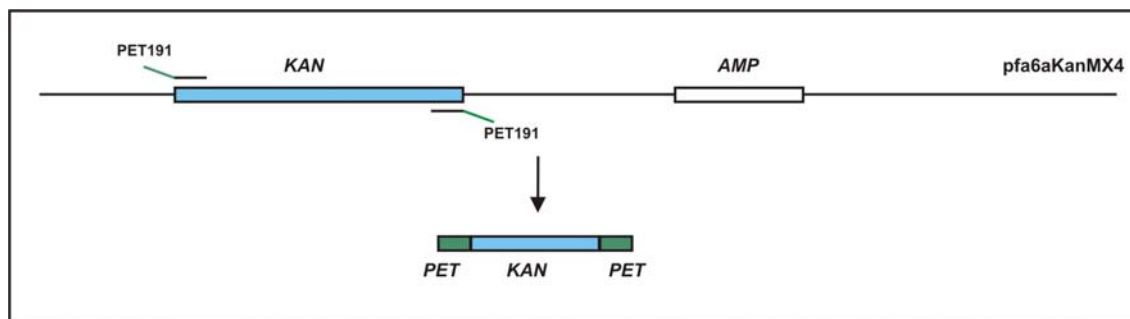
Principalmente no Brasil, cada usina possui uma linhagem específica que predomina durante toda a safra. Estas leveduras foram isoladas do próprio processo fermentativo e são multiplicadas todo ano para iniciarem o processo (Basso *et al.*, 2008). Para a realização de modificações genéticas é importante considerar a permanência das

linhagens nas dornas para que as leveduras possam expressar as características desejáveis e sobreviver ao estresse da fermentação.

Foram usadas para a realização desses experimentos, linhagens industriais haplóides derivadas da PE-2. Essa escolha foi feita, pois elas possuem as mesmas características, como a alta produtividade de etanol, observadas na linhagem diplóide (JAY270 – PE-2) (Argueso *et al.*, 2009 – artigo aceito para publicação Anexo I). A linhagem industrial haplóide pode ter a capacidade de exibir a mesma habilidade de competir e sobreviver nas condições fermentativas das indústrias.

### Levedura petite

Para deletar o gene *PET191*, gerando um petite nuclear (100% deficiente respiratório -  $\Delta$ pet191), e verificar se a ausência funcional desse gene realmente aumenta a produção de etanol, foram desenhados os *primers* DsPET191-2 (rev) e DsPET191-1 (fw). Eles amplificam o gene *KAN* do plasmídeo pFA6aKanMX4 tendo as extremidades homólogas ao gene *PET191* do genoma da levedura (Figura 22).



**FIGURA 22:** Esquema da construção do cassete integrativo por PCR a partir do plasmídeo pfa6aKanMX4. Os *primers* apresentam uma cauda homóloga ao gene *PET191*. No final, o cassete apresenta o gene *KAN* flanqueado por seqüências da região onde se deseja que ocorra a recombinação homóloga.

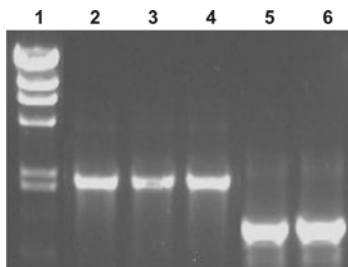
As linhagens FY23, JAY270, JAY291 e JAY292 foram transformadas com esse cassete e após plaquear em meio seletivo (YEPD + sorbitol + G418) observou-se uma grande quantidade de pequenas colônias. O aspecto aparentava crescimento residual de leveduras, mesmo assim, algumas colônias foram selecionadas para verificar se o gene *PET191* havia sido deletado. Isso foi feito por PCR com *primers* adjacentes ao gene *PET191*, o JAO262/JAO263.

Se houver integração, o fragmento amplificado seria de 2,4 Kb. Se a recombinação não tiver ocorrido na região correta a PCR amplificaria somente o gene *PET191*, gerando um fragmento de 1,4 Kb. O resultado da PCR mostrou o cassete PET-KAN-PET não integrou no genoma da levedura. O padrão de bandas encontrado na linhagem transformada foi o mesmo da linhagem não transformada.

Um dos motivos pode ter sido porque a região de homologia do cassete com a região do genoma é de apenas 37 bases. Essa quantidade de bases é o mínimo que se deve ter para que a recombinação ocorra, mas a eficiência aumentaria muito com o aumento de bases homólogas.

Esta estratégia foi baseada em Wach (Wach *et al.*, 1994). Segundo este artigo, é possível obter transformantes por recombinação com um mínimo de 35 pb de homologia com a região que se pretende substituir. Entretanto, esta metodologia não se mostrou eficiente para transformação de linhagens industriais. Para se obter sucesso na transformação e recombinação foi necessária a construção de um cassete com uma maior região de homologia. As dificuldades encontradas em se transformar linhagens industriais não são conhecidas. Elas devem ter algumas características peculiares que dificultam o processo de transformação e integração quando comparada com as linhagens de laboratório. Para isso uma nova metodologia foi escolhida, a do *double-joint* PCR.

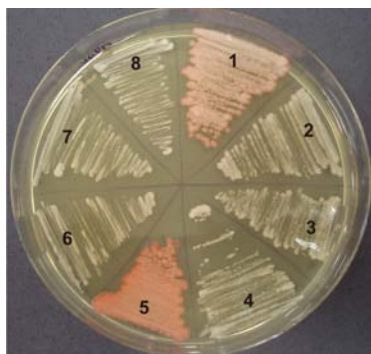
Nesse caso, o cassete foi construído com uma região de homologia de 143 bp na região 5' do gene *PET191* e de 245 bp na região 3' do gene *PET191*. Depois da transformação por eletroporação, diversas colônias apresentaram crescimento no meio com o antibiótico G418 e foram separadas para confirmação da integração por PCR. Como mostra a figura 23, as colônias 2, 3 e 4 confirmaram a inserção e foram usadas nos experimentos posteriores, enquanto as colônias 5 e 6 não apresentaram deleção do gene *PET191*.



**FIGURA 23.** PCR de confirmação da integração do cassete e conseqüente deleção do gene *PET191*. (1) marcador peso molecular,  $\lambda$ hind; (2 a 4) banda de 2,4 Kb indicando a inserção do gene *KAN* substituindo o gene *PET191*; e (5 e 6) banda de 1,4 Kb, indicando apenas o tamanho do gene *PET191*.

Para a obtenção do outro tipo de deficiente respiratório, também 100% deficiente, (petite mitocondrial ou citoplasmático -  $\rho^0$ ), as células foram crescidas e posteriormente tratadas com brometo de etídio. As colônias selecionadas são incapazes de crescer usando glicerol como fonte de carbono, não apresentando crescimento em meio YPG. Esse tipo de petite tem a taxa de crescimento menor quando comparada com as células normais, se apresentado como pequenas colônias quando plaqueadas.

Todos os mutantes respiratórios produzidos, tanto nucleares (testado previamente também por PCR) quanto o mitocondrial, foram confirmados pela técnica de cobertura por TTC. A figura 24 mostra duas linhagens testadas, FY23 e JAY292. As colônias 1 e 5 são os controles (células normais) e a coloração rosa confirma a capacidade de respirar. As demais colônias sem coloração são deficientes respiratórios: petites nucleares (2, 6) e petites mitocondriais (3, 4, 7, 8).



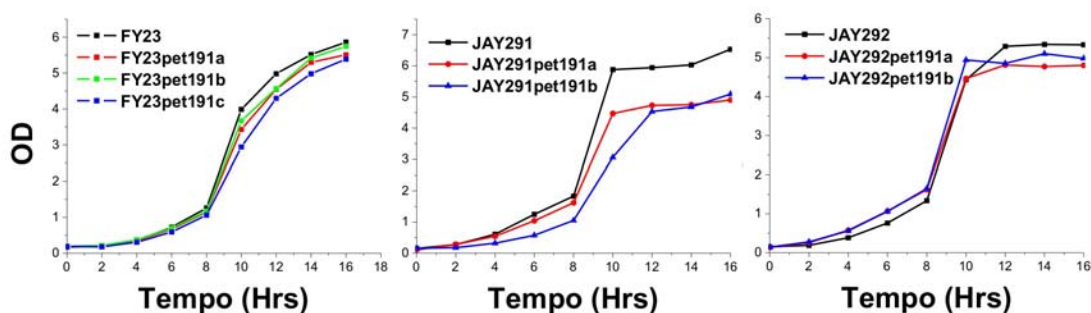
**FIGURA 24.** Técnica de cobertura com TTC. As colônias rosa têm a capacidade de respirar e as brancas são deficientes respiratórias. (1) JAY292; (2) JAY292 $\Delta$ pet191; (3) JAY292-EtBr-a; (4) JAY292-EtBr-b; (5) FY23; (6) FY23 $\Delta$ pet191; (7) FY23-EtBr-a; (8) FY23-EtBr-b.

As linhagens transformadas e os mutantes produzidos foram utilizados para a realização dos ensaios fermentativos.

### Ensaio Fermentativos

Foram feitas 4 fermentações consecutivas, reutilizando as mesmas células, em caldo de cana esterilizado (a porcentagem de açúcares redutores totais foi diluído para aproximadamente 18%), simulando as mesmas condições das usinas.

O mutante *petite* mitocondrial ( $\rho^0$ ) tem a taxa de crescimento menor (dados não mostrados) e isso pode também ser visualizado na placa, pois suas colônias apresentam tamanho bem inferior às colônias de células normais, devido a esse fato elas são chamadas de *petites*. O *petite* nuclear ( $\Delta pet191$ ), interessante, nas linhagens FY23 e JAY292 não apresentaram diferenças muito significativas na taxa de crescimento, quando comparada com as células normais (Figura 25). Esse mesmo fato foi observado por Hutter & Oliver (Hutter & Oliver, 1998). Apenas os mutantes da linhagem JAY291 apresentaram uma taxa de crescimento menor. Possivelmente esta última, por ser um haplóide derivado da esporulação da JAY270, apresenta algumas características diferentes do outro haplóide (JAY292), que acaba ficando mais acentuado após a deleção do gene *PET191*.



**FIGURA 25.** Curva de crescimento das linhagens FY23, JAY291 e JAY292 com seus respectivos mutantes ( $\Delta pet191$ ).

Foi comparada a produção de etanol entre as células normais e os dois tipos de mutantes produzidos.

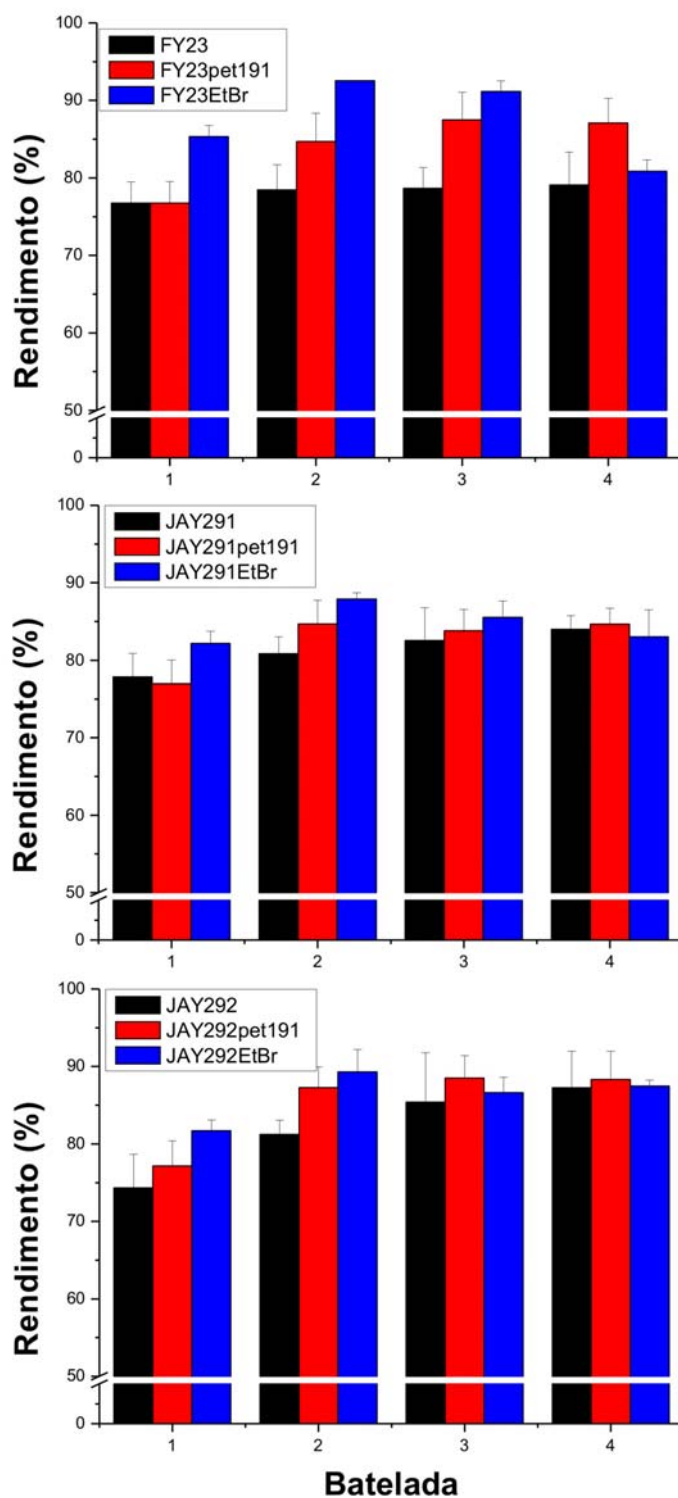
Os gráficos da figura 26 mostram os resultados das 4 bateladas consecutivas. A primeira batelada de todos os ensaios apresenta uma menor produção de etanol devido a

adaptação das leveduras ao meio de fermentação. Pois inicialmente o aumento da massa celular do inoculo foi preparado em meio YEPD.

Na fermentação feita com linhagens de laboratório é possível observar os melhores resultados (Figura 26). Células tratadas com brometo de etídio, FY23EtBr, já iniciaram a fermentação com alta produtividade. O petite nuclear, FY23 $\Delta$ pet191, e as células normais, iniciam a fermentação com uma produção reduzida e depois de alguns ciclos a produtividade se estabiliza. Na última batelada houve uma diminuição da produção de etanol pelo mutante tratado com brometo. Mas isso provavelmente foi decorrente de oscilações normais presentes no processo de fermentação. O controle FY23 naturalmente é uma linhagem ruim para a fermentação.

Essa diferença na produtividade entre células normais e petites é mais evidente, pois a linhagem FY23 apresenta sempre uma baixa taxa de produção de etanol. Isto pode ser visualizada quando se analisam os dados de todas as fermentações juntas (Figura 27). O aumento da produtividade em FY23 $\Delta$ pet191 e FY23EtBr foi de 7,4% e 11,8% respectivamente, comparado com as células normais FY23.

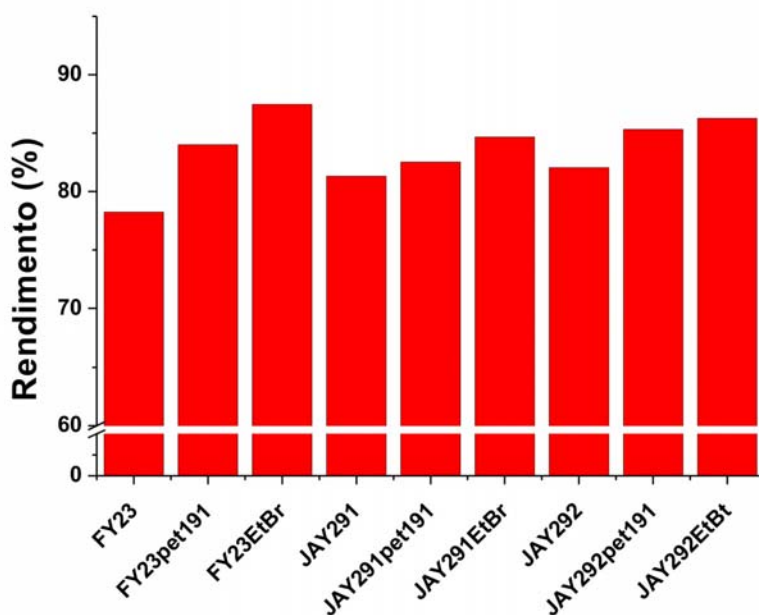
As duas linhagens haplóides petites JAY291 e JAY292 apresentam a mesma tendência durante os ensaios fermentativos. Na primeira batelada que seria o período de adaptação ao meio, é mais rápido ou é menos afetado pelas linhagens petite, especialmente o petite mitocondrial. Já na segunda batelada, os dois petites se destacam mais acentuadamente, principalmente na linhagem JAY292. Após a terceira batelada há uma tendência de se igualar a produtividade entre os mutantes e o controle.



**FIGURA 26.** Fermentação com caldo de cana. Os gráficos representam 4 bateladas de fermentação com o reciclo de células e cada barra representa a média das triplicatas. A porcentagem de etanol produzido mostrado no gráfico foi normalizada com o valor máximo teórico, que depende da quantidade inicial de glicose.



As linhagens haplóides industriais (JAY291 e JAY292) se mostram bons fermentadores da mesma forma que o diplóide que os originou e por isso o aumento na produtividade de etanol pelos dois tipos de mutantes é menor do que as observadas nas linhagens de laboratório. O haplóide, JAY292, tem uma maior capacidade fermentativa, produzindo 4,0% e 5,2% mais etanol em JAY292 $\Delta$ pet191 e JAY292EtBr, respectivamente, comparado com células normais. O outro haplóide, JAY291, teve um aumento de 1,5% e 4% pelos mutantes, JAY291 $\Delta$ pet191 e JAY291EtBr, respectivamente (Figura 27 e tabela 5).



**FIGURA 27.** Total de etanol produzido durante o todo o período do experimento. O gráfico mostra a média de todas as repetições e de todas as bateladas. Na tabela, é mostrada a produção total e o aumento de produção pelos mutantes. A porcentagem de aumento é referente à respectiva linhagem controle.

**TABELA 5.** Rendimento total e o aumento de produção de etanol pelos mutantes. A porcentagem de aumento é referente à respectiva linhagem controle.

<b>Linhagem</b>	<b>Rendimento (%)</b>	<b>% de aumento</b>
FY23	78,24	-
FY23 $\Delta$ pet191	84,00	7,36
FY23EtBr	87,46	11,78
JAY291	81,31	-
JAY291 $\Delta$ pet191	82,52	1,49
JAY291EtBr	84,66	4,12
JAY292	82,04	-
JAY292 $\Delta$ pet191	85,31	3,99
JAY292EtBr	86,27	5,15

A perda funcional do DNA mitocondrial (petite citoplasmático) torna as células incapazes de crescer por meio da respiração, mas isso pode ser suprido pois as leveduras podem crescer utilizando fontes de carbono fermentáveis, e as células não estão sujeitas ao efeito Pasteur (inibição da glicólise pela fosforilação oxidativa ou a inibição da fermentação pela adição de oxigênio) (Slonimski *et al.*, 1968). Apesar disso, a ausência do DNA mitocondrial pode provocar muitas alterações deletérias nas características da superfície da célula, como a redução na captação dos açúcares, alteração da floculação e aglutinação, diminuição da tolerância ao etanol por afetar as proteínas da membrana plasmática (Wilkie & Evans, 1982). Os petites citoplasmáticos provavelmente não sobreviveriam às condições das dornas de fermentação, devido às condições altamente estressantes.

A vantagem do petite nuclear ( $\Delta$ pet191) é que os genes presentes na mitocôndria continuam intactos e todas as outras funções da mitocôndria permanecem funcionais aumentando a probabilidade das células colonizarem as dornas de fermentação. Além disso, a deleção do gene *PET191* não promove alteração na taxa de crescimento. O fato de a linhagem possuir uma taxa de crescimento semelhante ao selvagem é uma vantagem, pois não altera o tempo para a produção do inóculo. Um atraso na taxa de crescimento implica em um maior tempo na produção do inóculo e conseqüentemente um gasto maior pra usina. E ainda, devido a essa deficiência respiratória, o petite não é capaz de metabolizar o etanol como substrato secundário, permitindo assim manter todo o etanol produzido.

## CONCLUSÃO

Linhagens de laboratório e industriais pequenas apresentaram um aumento de rendimento na produção de etanol.

As linhagens industriais modificadas com as características desejáveis têm a possibilidade de dominar o processo durante todo o período da safra. As leveduras industriais sem o gene *PET191*, tiveram um aumento de 4% na produção, baseando-se na linhagem JAY292. No Brasil, onde a produção anual é cerca de 24,5 bilhões de litros de etanol, este aumento de porcentagem representaria um aumento de 980 milhões de litros por ano, que é muito significativo.

Novos experimentos serão feitos para comprovar o efeito provocado pela deleção do gene *PET191*. Um deles é a cinética da produção de etanol durante a fermentação. Provavelmente os mutantes para o gene *PET191* conseguiriam atingir o máximo de etanol produzido mais rápido que as linhagens que respiram. Isso diminuiria ainda mais o tempo necessário para que a fermentação ocorra.

## CONCLUSÕES FINAIS

A melhoria do processo fermentativo pode ser obtida com o controle da expressão do gene da floculação. A regulação condicional do gene *FLO5* foi obtida com sucesso, entretanto a impossibilidade de desfloculação inviabiliza a utilização deste gene no processo industrial. Além disso, a linhagem industrial transformada com a construção do gene *FLO5* teve uma diminuição do rendimento fermentativo. Isso pode ter ocorrido, pois a integração da construção se deu no local do gene da arginina permease. Este gene não é essencial para a levedura, mas a redução ou interrupção na captação de arginina pode ter afetado a eficiência da linhagem na produção de etanol.

A utilização de outro gene da floculação, *FLO10*, também regulado condicionalmente, se mostrava promissor pois a proteína resultante deste gene seria inativado naturalmente pela glicose. Infelizmente, os resultados obtidos não foram os esperados. O gene *FLO10* apresentou uma expressão basal durante toda a fermentação e os agregados produzidos não desfloculavam após nova adição de glicose.

Apesar disso, há a possibilidade de se utilizar linhagens floculantes na produção de cachaça e vinho, sem a necessidade de desfloculação, pois não há reuso de leveduras.

Durante o período do doutorado foi feito um trabalho em parceria com o pesquisador, Dr. Juan Argueso, com o objetivo de se estudar uma linhagem industrial, compreendendo sua fisiologia e genética. O resultado desse estudo gerou um artigo que foi aceito (agosto 2009) para publicação pela Genome Research (AnexoII). Foram encontrados vários fatores que podem estar relacionados com o melhor desempenho das leveduras industriais nas dornas. Este artigo começa a desvendar os “segredos do sucesso” das linhagens selvagens industriais brasileiras, abrindo oportunidades para futuras manipulações desses organismos.

A última etapa da tese, foi a deleção do gene *PET191*, que resultou em um aumento na produção de etanol nas linhagens de laboratório e industriais. A possibilidade da domesticação e manipulação de leveduras industriais visando uma maior produtividade do processo motiva a continuidade das pesquisas em busca de novos desafios.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Allen, F. (1994). The Microbrewery Laboratory Manual - Part III: Wild Yeasts Detection and Remediation, BrewingTechniques. Vol. 2, No 6.
- Alper, H.; Moxley, J.; Nevoigt, E.; Fink, G. R.; Stephanopoulos, G. (2006). "Engineering yeast transcription machinery for improved ethanol tolerance and production." Science **314**(5805): 1565-8.
- Ausubel, F. M.; Brent, R.; Kingston, R. E.; Moore, D. D.; Seidman, J. G.; Smith, A. J.; Struhl, K. (1998). Current Protocols in Molecular Biology. New York - USA, John Wiley & Sons Inc.
- Barney, M. C.; Jansen, G. P.; Helber, G. R. (1980). "Use of genetic transformation for the introduction of flocculence into yeast." J Am Soc Brew Chem **38**: 71-74.
- Basso, L. C.; de Amorim, H. V.; de Oliveira, A. J.; Lopes, M. L. (2008). "Yeast selection for fuel ethanol production in Brazil." FEMS Yeast Res **8**(7): 1155-63.
- Beier, D. R.; Sledziewski, A.; Young, E. T. (1985). "Deletion analysis identifies a region, upstream of the ADH2 gene of *Saccharomyces cerevisiae*, which is required for ADR1-mediated derepression." Mol Cell Biol **5**(7): 1743-9.
- Bester, M. C.; Pretorius, I. S.; Bauer, F. F. (2006). "The regulation of *Saccharomyces cerevisiae* FLO gene expression and Ca<sup>2+</sup> -dependent flocculation by Flo8p and Mss11p." Curr Genet **49**(6): 375-83.
- Bidard, F.; Blondin, B.; Dequin, S.; Vezinhet, F.; Barre, P. (1994). "Cloning and analysis of a *FLO5* flocculation gene from *Saccharomyces cerevisiae*." Curr Genet **25**(3): 196-201.
- Borneman, A. R.; Forgan, A. H.; Pretorius, I. S.; Chambers, P. J. (2008). "Comparative genome analysis of a *Saccharomyces cerevisiae* wine strain." FEMS Yeast Res **8**(7): 1185-95.
- Branduardi, P.; Smeraldi, C.; Porro, D. (2008). "Metabolically engineered yeasts: 'potential' industrial applications." J Mol Microbiol Biotechnol **15**(1): 31-40.
- Brunelli, J. P.; Pall, M. L. (1993). "A series of yeast shuttle vectors for expression of cDNAs and other DNA sequences." Yeast **9**(12): 1299-308.
- Calleja, G. B. (1973). "Role of mitochondria in the sex-directed flocculation of a fission yeast." Arch Biochem Biophys **154**(1): 382-6.
- Carreto, L.; Eiriz, M. F.; Gomes, A. C.; Pereira, P. M.; Schuller, D.; Santos, M. A. (2008). "Comparative genomics of wild type yeast strains unveils important genome diversity." BMC Genomics **9**: 524.

- Ciriacy, M. (1979). "Isolation and characterization of further cis- and trans-acting regulatory elements involved in the synthesis of glucose-repressible alcohol dehydrogenase (ADHII) in *Saccharomyces cerevisiae*." Mol Gen Genet **176**(3): 427-31.
- Cortez, L. A. B.; Lora, E. E. S.; Gómez, E. O. (2008). Biomassa para Energia, Editora Unicamp.
- Cunha, A. F.; Missawa, S. K.; Gomes, L. H.; Reis, S. F.; Pereira, G. A. (2006). "Control by sugar of *Saccharomyces cerevisiae* flocculation for industrial ethanol production." FEMS Yeast Res **6**(2): 280-7.
- Denis, C. L.; Young, E. T. (1983). "Isolation and characterization of the positive regulatory gene ADR1 from *Saccharomyces cerevisiae*." Mol Cell Biol **3**(3): 360-70.
- DeRisi, J. L.; Iyer, V. R.; Brown, P. O. (1997). "Exploring the metabolic and genetic control of gene expression on a genomic scale." Science **278**(5338): 680-6.
- Devlin, T. M. (1997). Textbook of Biochemistry with clinical correlations, Wiley-Liss.
- Domingues, L.; Vicente, A. A.; Lima, N.; Teixeira, J. A. (2000). "Applications of yeast flocculation in biotechnological processes." Biotechnol Bioprocess Eng **5**: 288-305.
- Eddy, A. A.; Rudin, A. D. (1958). "Part of the yeast surface apparently involved in flocculation." J. Inst. Brew.(64): 19-21.
- Egilsson, V.; Evans, I. H.; Wilkie, D. (1979). "Toxic and mutagenic effects of carcinogens on the mitochondria of *Saccharomyces cerevisiae*." Mol. Gen. Genet.(174): 39-46.
- Epstein, C. B.; Waddle, J. A.; Hale, W. t.; Dave, V.; Thornton, J.; Macatee, T. L.; Garner, H. R.; Butow, R. A. (2001). "Genome-wide responses to mitochondrial dysfunction." Mol Biol Cell **12**(2): 297-308.
- Farrell, A. E.; Plevin, R. J.; Turner, B. T.; Jones, A. D.; O'Hare, M.; Kammen, D. M. (2006). "Ethanol can contribute to energy and environmental goals." Science **311**(5760): 506-8.
- Fox, T. D.; Folley, L. S.; Mulero, J. J.; McMullin, T. W.; Thorsness, P. E.; Hedin, L. O.; Costanzo, M. C. (1991). "Analysis and manipulation of yeast mitochondrial genes." Methods Enzymol **194**: 149-65.
- Gibson, B. R.; Lawrence, S. J.; Leclaire, J. P.; Powell, C. D.; Smart, K. A. (2007). "Yeast responses to stresses associated with industrial brewery handling." FEMS Microbiol Rev **31**(5): 535-69.
- Goffeau, A.; Barrell, B. G.; Bussey, H.; Davis, R. W.; Dujon, B.; Feldmann, H.; Galibert, F.; Hoheisel, J. D.; Jacq, C.; Johnston, M.; Louis, E. J.; Mewes, H. W.; Murakami,

- Y.; Philippsen, P.; Tettelin, H.; Oliver, S. G. (1996). "Life with 6000 genes." Science **274**(5287): 546, 563-7.
- Goldemberg, J. (2008). "The Brazilian biofuels industry." Biotechnol Biofuels **1**(1): 6.
- Goldemberg, J.; Nigro, F. E. B.; Coelho, S. T. (2008). Bioenergia no Estado de São Paulo: situação atual, perspectivas, barreiras e propostas. São Paulo, Imprensa Oficial do Estado de São Paulo.
- Govender, P.; Domingo, J. L.; Bester, M. C.; Pretorius, I. S.; Bauer, F. F. (2008). "Controlled expression of the dominant flocculation genes FLO1, FLO5, and FLO11 in *Saccharomyces cerevisiae*." Appl Environ Microbiol **74**(19): 6041-52.
- Griffiths, A. J. F.; Miller, J. H.; Suzuki, D. T.; Lewontin, R. C.; Gelbart, W. M. (1993). An Introduction to Genetics Analysis. New York, W. H. Freeman and Company.
- Guo, B.; Styles, C. A.; Feng, Q.; Fink, G. R. (2000). "A *Saccharomyces* gene family involved in invasive growth, cell-cell adhesion, and mating." Proc Natl Acad Sci U S A **97**(22): 12158-63.
- Guo, Z. P.; Zhang, L.; Ding, Z. Y.; Wang, Z. X.; Shi, G. Y. (2009). "Interruption of glycerol pathway in industrial alcoholic yeasts to improve the ethanol production." Appl Microbiol Biotechnol **82**(2): 287-92.
- Halme, A.; Bumgarner, S.; Styles, C.; Fink, G. R. (2004). "Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast." Cell **116**(3): 405-15.
- Hinrichs, J.; Stahl, U.; Esser, K. (1988). "Flocculation in *Saccharomyces cerevisiae* and mitochondrial DNA structure." Appl. Microbiol. Biotechnol **29**: 48-54.
- Hutter, A.; Oliver, S. G. (1998). "Ethanol production using nuclear petite yeast mutants." Appl. Microbiol. Biotechnol **49**(5): 511-516.
- Ishida-Fujii, K.; Goto, S.; Sugiyama, H.; Takagi, Y.; Saiki, T.; Takagi, M. (1998). "Breeding of flocculent industrial alcohol yeast strains by self-cloning of the flocculation gene FLO1 and repeated-batch fermentation by transformants." J Gen Appl Microbiol **44**(5): 347-353.
- Iung, A. R.; Coulon, J.; Kis, F.; Ekome, J. n. N.; Vallner, J.; Bonaly, R. (1999). "Mitochondrial function in cell wall glycoprotein synthesis in *Saccharomyces cerevisiae* NCYC 625 (Wild type) and [rho0] mutants." Applied and Environmental Microbiology **65**(12): 5398-5402.
- Jank, M. S. (2007). Da Alca para o álcool. O Estado de São Paulo. São Paulo: J3.

- Johnston, J. R.; Reader, H. P. (1983). Genetic control of flocculation. Yeast genetics: Fundamental and applied aspects. J. F. T. Spencer, D.M. and Smith, A.R.W. New York, Springer-Verlag: 205-224.
- Kobayashi, O.; Hayashi, N.; Kuroki, R.; Sone, H. (1998). "Region of FLO1 proteins responsible for sugar recognition." J Bacteriol **180**(24): 6503-10.
- Kobayashi, O.; Suda, H.; Ohtani, T.; Sone, H. (1996). "Molecular cloning and analysis of the dominant flocculation gene FLO8 from *Saccharomyces cerevisiae*." Mol Gen Genet **251**(6): 707-15.
- Lee, K. M.; DaSilva, N. A. (2005). "Evaluation of the *Saccharomyces cerevisiae* ADH2 promoter for protein synthesis." Yeast **22**(6): 431-40.
- Lima, U. A.; Aquarone E.; Borzani, W. (1975). Tecnologia das fermentações. **1**: 48-59.
- Martinelli, A. (1983). Tecnologia de vinhos e vinagres de frutas.
- Mastrandrea, M. D.; Stephen, H. S. (2005). Global Warming, World Book Online Reference Center. **2009**.  
<http://www.worldbookonline.com/wb/Article?id=ar226310>.
- Meleiro, L. A. C.; Maciel, R. (2000). "State and parameter estimation based on a nonlinear filter applied to an industrial process control of ethanol production." Braz J Chem Eng **17**(4-7): 991-1001.
- Miki, B. L.; Poon, N. H.; James, A. P.; Seligy, V. L. (1982). "Possible mechanism for flocculation interactions governed by gene FLO1 in *Saccharomyces cerevisiae*." J Bacteriol **150**(2): 878-89.
- Nelson, N. (1944). "A fotometric adaptaion of Somogyi method for the determination of glucose." J.Biol.Chem. **153**: 375 - 380.
- Nevoigt, E. (2008). "Progress in metabolic engineering of *Saccharomyces cerevisiae*." Microbiol Mol Biol Rev **72**(3): 379-412.
- Ooi, B. G.; Lankford, K. R. (2009). "Strategy for adapting wine yeasts for bioethanol production." Int J Mol Sci **10**(1): 387-4.
- Otterstedt, K.; Larsson, C.; Bill, R. M.; Stahlberg, A.; Boles, E.; Hohmann, S.; Gustafsson, L. (2004). "Switching the mode of metabolism in the yeast *Saccharomyces cerevisiae*." EMBO Rep **5**(5): 532-7.
- Panoutsopoulou, K.; Hutter, A.; Jones, P.; Gardner, D. C. J.; Oliver, S. G. (2001). "Improvement of Ethanol Production by an Industrial Yeast Strain via Multiple Gene Deletions." J Inst Brew **107**(1): 49-53.



- Pessoa-Jr, A.; Roberto, I. C.; Menossi, M.; dos Santos, R. R.; Filho, S. O.; Penna, T. C. (2005). "Perspectives on bioenergy and biotechnology in Brazil." Appl Biochem Biotechnol **121-124**: 59-70.
- Price, V. L.; Taylor, W. E.; Clevenger, W.; Worthington, M.; Young, E. T. (1990). "Expression of heterologous proteins in *Saccharomyces cerevisiae* using the ADH2 promoter." Methods Enzymol **185**: 308-18.
- Rando, O. J.; Verstrepen, K. J. (2007). "Timescales of genetic and epigenetic inheritance." Cell **128**(4): 655-68.
- Rajo, F. (2008). "Biofuels from microbes: a comprehensive view." Microbial Biotechnology **1**(3): 208-210.
- Salomão, A.; Onaga, M. (2006). Etanol o mundo quer. O Brasil tem. Revista Exame. **0870**.
- Sanchez, O. J.; Cardona, C. A. (2008). "Trends in biotechnological production of fuel ethanol from different feedstocks." Bioresour Technol **99**(13): 5270-95.
- Shen, H.; Iha, H.; Yaguchi, S.; Tsurugi, K. (2006). "The mechanism by which overexpression of Gts1p induces flocculation in a FLO8-inactive strain of the yeast *Saccharomyces cerevisiae*." FEMS Yeast Res **6**(6): 914-23.
- Slonimski, P. P.; Perrodin, G.; Croft, J. H. (1968). "Ethidium bromide induced mutation of yeast mitochondria: complete transformation of the cells into respiratory deficient non-chromosomal petites." Biochem Biophys Res Commun **30**: 232-239.
- Somogyi, M. (1945). "A new reagent for determination sugars." A New Sugar Reagent **May**: 61 - 68.
- Stan, S.; Despa, F. (2000). "On the doublet formation in the flocculation process of the yeast cells." Biosystems **57**(3): 139-45.
- Stratford, M. (1996a). "Induction of flocculation in brewing yeasts by change in pH value." FEMS Microbiol Lett **136**(1): 13-8.
- Stratford, M. (1996b). "Yeast flocculation: restructuring the theories in line with recent research." Cerevisia **21**(1): 38-45.
- Teunissen, A. W.; Steensma, H. Y. (1995). "Review: the dominant flocculation genes of *Saccharomyces cerevisiae* constitute a new subtelomeric gene family." Yeast **11**(11): 1001-13.
- Tortora, G. J.; Funke, B. R.; Case, C. L. (2002). Microbiology - An Introduction, **Benjamin Cummings**.

- Touhami, A.; Hoffmann, B.; Vasella, A.; Denis, F. A.; Dufrêne, Y. F. (2003). "Aggregation of yeast cells: direct measurement of discrete lectin-carbohydrate interactions." Microbiology **149**: 2873-2878.
- Van Mulders, S. E.; Christianen, E.; Saerens, S. M.; Daenen, L.; Verbelen, P. J.; Willaert, R.; Verstrepen, K. J.; Delvaux, F. R. (2009). "Phenotypic diversity of Flo protein family-mediated adhesion in *Saccharomyces cerevisiae*." FEMS Yeast Res **9**(2): 178-90.
- Varela, C.; Cardenas, J.; Melo, F.; Agosin, E. (2005). "Quantitative analysis of wine yeast gene expression profiles under winemaking conditions." Yeast **22**(5): 369-83.
- Verstrepen, K. J.; Derdelinckx, G.; Delvaux, F. R.; Winderickx, J.; Thevelein, J. M.; Bauer, F. F.; Pretorius, I. S. (2001). "Late fermentation expression of *FLO1* in *Saccharomyces cerevisiae*." Journal of the American Society of Brewing Chemists **59**(2): 69-76.
- Verstrepen, K. J.; Derdelinckx, G.; Verachtert, H.; Delvaux, F. R. (2003). "Yeast flocculation: what brewers should know." Appl Microbiol Biotechnol **61**(3): 197-205.
- Verstrepen, K. J.; Jansen, A.; Lewitter, F.; Fink, G. R. (2005). "Intragenic tandem repeats generate functional variability." Nat Genet **37**(9): 986-90.
- Verstrepen, K. J.; Klis, F. M. (2006). "Flocculation, adhesion and biofilm formation in yeasts." Mol Microbiol **60**(1): 5-15.
- Verstrepen, K. J.; Reynolds, T. B.; Fink, G. R. (2004). "Origins of variation in the fungal cell surface." Nat Rev Microbiol **2**(7): 533-40.
- Wach, A.; Brachat, A.; Pohlmann, R.; Philippsen, P. (1994). "New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*." Yeast **10**(13): 1793-808.
- Wackett, L. P. (2008). "Biomass to fuels via microbial transformations." Curr Opin Chem Biol **12**(2): 187-93.
- Wang, F. Z.; Shen, W.; Rao, Z. M.; Fang, H. Y.; Zhan, X. B.; Zhuge, J. (2007). "Construction of a flocculating yeast for fuel ethanol production." Biotechnol Lett.
- Watari, J.; Kudo, M.; Ogawa, M.; Nishikawa, N.; Kamimura, M. (1989). "Molecular cloning of a flocculation gene in *Saccharomyces cerevisiae*." Agric Biol Chem **53**: 901-903.
- Wheals, A. E.; Basso, L. C.; Alves, D. M.; Amorim, H. V. (1999). "Fuel ethanol after 25 years." Trends Biotechnol **17**(12): 482-7.

- Wilkie, D.; Evans, I. (1982). "Mitochondria and the yeast cell surface: implications for carcinogenesis." Trends Biochem Sci **7**(4): 147 - 151.
- Winston, F.; Dollard, C.; Ricupero-Hovasse, S. L. (1995). "Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C." Yeast **11**(1): 53-5.
- Yu, J. H.; Hamari, Z.; Han, K. H.; Seo, J. A.; Reyes-Dominguez, Y.; Scazzocchio, C. (2004). "Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi." Fungal Genet Biol **41**(11): 973-81.

**ANEXO I**

**INDUSTRIAL ASPECTS OF BIOETHANOL**

**SECTION 2**

**BIOTECHNOLOGY AND ENGINEERING SCIENCE INTEGRATION FOR  
CHEMICAL AND FUEL INDUSTRIES**

**CHAPTER 4**

**INDUSTRIAL POTENTIAL OF YEAST BIOTECHNOLOGY IN THE  
PRODUCTION OF BIOETHANOL IN BRAZIL: THE EXAMPLE OF  
CONDITIONAL FLOCCULATION**

*ANDERSON F. CUNHA, SILVIA K. MISSAWA AND GONÇALO A. G. PEREIRA*



INDUSTRIAL  
PERSPECTIVES  
FOR BIOETHANOL



**DIRETORIA EXECUTIVA**

**Mauricio Prates de Campos Filho**

Diretor Executivo

**Nelson Antonio Pereira Camacho**

Diretor para Assuntos Administrativos e Financeiros

**Saul Gonçalves d'Ávila**

Diretor para Assuntos Científicos e Tecnológicos

# INDUSTRIAL PERSPECTIVES FOR BIOETHANOL

Eduardo P. Carvalho Isaias C. Macedo José Luiz Olivério  
Pedro Wongtschowski Francisco I. Pellegrini Flávio C. Cavalcanti  
Anderson F. Cunha Silvia K. Missawa Gonçalo A. G. Pereira  
Silvio R. Andrietta Maria da Graça Stupiello Andrietta  
Adrie J. J. Straathof Çagri Efe Luuk A. M. van der Wielen  
Peter M. M. Nossin Telma T. Franco Carlos Eduardo Vaz Rossell  
H. J. Veringa P. Alderliesten

Edited by  
Telma Teixeira Franco



**Coleção UNIEMP Inovação**

INDUSTRIAL PERSPECTIVES FOR BIOETHANOL

ISBN 85-98951-06-4



**Av. Paulista, 2.200 - 16º andar**

**Tel. (11) 2178-0466**

**Fax. (11) 3283-3386**

**uniemp@uniemp.org.br**

**www.uniemp.org.br**

**01310-300 São Paulo - SP**

**Produzido no Brasil**

**2006**



## LIST OF CONTENTS

Preface  
Telma T. Franco  
9

Section 1  
Industrial Perspectives for Bioethanol:  
a Brazilian approach  
11

**Chapter 1**  
**Ethanol from Sugar Cane in Brazil:  
the Industrial Technology Perspectives**  
Eduardo P. Carvalho and Isaias C. Macedo  
13

**Chapter 2**  
**Technological evolution of the Brazilian  
sugar and alcohol sector:  
Dedini's contribution**  
José Luiz Olivério  
19

**Chapter 3**  
**Insertion of bioethanol into a new paradigm  
for the Brazilian chemical industry**  
Pedro Wongtschowski, Francisco I. Pellegrini  
and Flávio C. Cavalcanti  
33

Section 2  
BioTechnology and Engineering Science integration  
for Chemical and Fuel Industries  
57

Chapter 4  
**Industrial Potential of Yeast Biotechnology  
in the Production of Bioethanol in Brazil:  
the Example of Conditional Flocculation**  
Anderson F. Cunha, Silvia K. Missawa  
and Gonçalo A. G. Pereira  
59

Chapter 5  
**Optimization and Limits of Ethanol Production**  
Silvio R. Andrietta  
and Maria da Graça Stupiello Andrietta  
77

Chapter 6  
**Innovation and sustainability through  
industrial biotechnology**  
Adrie J. J. Straathof, Çagri Efe,  
Luuk A. M. van der Wielen,  
Peter M. M. Nossin and Telma T. Franco  
91

Chapter 7  
**Advances of the Brazilian production of chemicals  
and other products from biomass**  
Telma T. Franco  
107

Section 3  
Role of Bioethanol in the Energy Landscape  
121

**Chapter 8**

**Conversion of lignocellulose biomass (bagasse and straw)  
from the sugar-alcohol industry into bioethanol**

Carlos Eduardo Vaz Rossell

123

**Chapter 9**

**Advanced techniques for generation of energy  
from biomass and waste**

H. J. Veringa and P. Alderliesten

143

# Industrial Potential of Yeast Biotechnology in the Production of Bioethanol in Brazil: the Example of Conditional Flocculation

Anderson Ferreira da Cunha  
Silvia Kazue Missawa  
Gonçalo Amarante Guimarães Pereira

## Introduction

The existence of life on earth as we know it is due fundamentally to the reduction of CO<sub>2</sub> levels primarily existing in the atmosphere, mainly through the fixation of this compound into carbon chains. These long-term processes combine with geologic events, although they are basically biological activities.

Man as a species has developed in this rather milder atmosphere, although his social and economic progress, in the last hundred years, has reversed the direction of carbon fixation and is now jeopardizing biological maintenance of the atmosphere. Since the introduction of fossil fuels (gasoline, diesel, etc.) utilization, we have systematically released gigantic amounts of carbonic gas back to the atmosphere effectively reversing millions of years of carbon fixation. This means that our social progress is taking us back to our biological environmental starting point, which will certainly have major consequences.

In recent years a number of research studies, as well as scientific and social organizations, have pointed to this problem, although without much success. The largest producers of CO<sub>2</sub> even admit that this process is occurring, but up to now, they have only considered it as something that will occur in the remote future. Within this context, weather changes

that have occurred in the past few years, particularly floods and the record number of hurricanes, have brought to light an unmistakable demonstration that the consequences of fossil fuels are already here. Therefore, there is no more time to wait, alternative and renewable energy sources must be found.

The history of economic events show that during great crises, great opportunities arise and the ability to capitalize during these times is key to economic success. In the specific scenario of the fuel crisis a huge opportunity has been created for countries that have ample land, sun and water. This is undoubtedly the case of Brazil.

Although most of the developed world has reached a situation that is considered critical, with dependence on imported petroleum and alarming and overwhelming natural phenomena linked to climate changes, Brazil started take positive step to reduce this situation about 30 years ago, for economical reasons. With the oil shortage that occurred in 1973 and 1979, oil prices rose from US\$ 4.00 to US\$ 40.00 per barrel. Between 1973 and 1974, fuel costs jumped from US\$ 600 million to over US\$ 2 billion. This gigantic impact on the balance of payments showed the strategic vulnerability of Brazil, which at the time was importing nearly 80% of the oil consumed by the Country, and therefore was highly susceptible to an energy collapse.

In 1975, Brazil launched Proalcool, the Brazilian Program for Alcohol Fuel. Sugarcane alcohol soon provided an ideal alternative fuel for gasoline. In order to make this program feasible, the Brazilian government offered a financial plan that was supported by the World Bank, which made it possible to increase the areas planted to sugarcane, build new plants and develop and enlarge boiler shop industries (source: UNICA, the Sao Paulo State Sugarcane Agribusiness Union – [www.unica.com.br](http://www.unica.com.br)).

Once available, this so-called green fuel was harnessed in two forms. Initially, it was mixed with gasoline as anhydrous alcohol. This process has increased the amount of alcohol in gasoline over time, i.e. the gasoline distributed in Sao Paulo city changed from a 20% v/v mixture in 1977 to 22% in 1980, which has been adopted all over the country (Demetrius 1990). Also during this period the first cars that exclusively

utilized hydrated alcohol were introduced, unfortunately, these vehicles used poorly adapted gasoline motors and were not well received. Only during the 1980s, after the second oil shortage did automobile manufacturers launch new models that were especially manufactured to work with alcohol. Also during this time alcohol prices were kept favorable to gasoline, through governmental subsidies. The success of these automobiles exceeded all predictions. In 1984, alcohol-powered cars represented 94.4% of the assembly line production in Brazil, which accounted for 19 out of 20 cars.

However, this huge increase in demand was not maintained and combined with increases in the Brazilian production of crude oil and the reduction in the world price of this commodity, there was limited interest in maintaining the alcohol subsidies. Moreover, ethanol producers started an irrational process of increasing prices, which led to a shortage of the product, resulting in a lack of confidence in the consumer market. As a combined consequence, the production of alcohol-powered cars fell from 88.4% in 1988 to 61% in 1989, to 19.9% in 1990 and finally to 0.3% in 1996.

However, in spite of the program's limited adoption, the government viewed this as strategically important to the country and kept alcohol production at high levels. In 1991 a new federal law mandated the addition of 22% anhydrous alcohol to gasoline, which helped to promote the production of ethanol for fuel purposes. Associated with this strategy, environmental factors in cities like Sao Paulo, where air pollution had reached alarming levels, could also be improved by the addition of alcohol to gasoline.

More recently, the alcohol program has once again returned to the spotlight, at first because a consistent and apparently irreversible increase in the international crude oil prices and due to the development of the so-called FLEX engines that are capable of burning hydrated alcohol/gasoline mixture at any proportion. Cars with these new motors have given Brazilian consumers a choice in which fuel to use and thus a relevant power to resist increases in gasoline prices.

Thus, given the global demand for alternative fuels, Brazil now presents itself as the country capable of leading the world. Brazil has a huge production capacity and has developed new technologies for production,

distribution and utilization of ethanol, as well as demonstrating increased implementation of these technologies. For example, the sugar mill extract process has increased the amount of sugar that can be processed from sugarcane. The sugar extraction index, that was 89%, at the beginning of the Proalcool program, has reached 97% today. In parallel, the sugarcane residues (bagasse) burning technology to produce energy has been widely adopted, thus making approximately 95% of the sugar mills self-sufficient in terms of electricity (van Haandel 2005), and in some case providing excess electric that is redirected into the national electrical networks.

However, there are still technological advances that remain to be developed in order to maximize the inherently biological processes. In the case of sugarcane, over the last few years an intensive research effort has been developed to improve the plant. Genomics have been used to in an attempt to understand the biochemical processes connected to the plant's physiology, and these have lead to potential features of agronomic importance (Vettore et al. 2003; Vincentz et al. 2004). This advancement has led to the recent creation of companies focused on the use of biotechnology in the development of new cultivated varieties, as in the case of Canavialis ([www.canavialis.com.br](http://www.canavialis.com.br)). Curiously, this has not happened in the case of *Saccharomyces cerevisiae*. Even though this yeast is an essential part in fermentation, very little research has been done to improve its efficiency and specificity. This lack of research presents huge opportunities for biotechnological improves. This issue will be discussed in detail in the following items.

### The Biochemistry of Ethanol Production

The ethanol production process in Brazil occurs almost exclusively by fermentation of sugarcane juice by yeasts, mainly the *Saccharomyces cerevisiae* species, an organism that, given its technological importance, is widely used as an experimental model to understand eukaryotic cells (Goffeau et al. 1996).

Biochemically, alcohol fermentation is the degradation of hexoses into ethanol with the production of 2 ATP molecules in the process. In

the case of sugarcane juice, the fermentation process metabolizes the disaccharide sucrose that is formed by one molecule of fructose and one molecule of glucose. The first step is the hydrolysis of sucrose by invertase enzyme that is encoded by the SUC2 gene. Both glucose and fructose enter the glycolytic pathway and are converted into pyruvate via a sequence of reactions (Figure 1). Pyruvate is then decarboxylated to form acetaldehyde and this releases CO<sub>2</sub> and causes intense bubbling when the cells are grown in liquid medium. Later, the acetaldehyde molecule is reduced to ethanol, and this reaction is catalyzed by the enzyme alcohol dehydrogenase 1 (ADH1). The glucose depletion results in the end of the fermentation phase. During this period the yeast cells vigorously multiply, which is characterized by an exponential growth phase. From the depletion point on the yeasts seem to enter a stationary phase, but they actually continue to grow for a long period, and actually consume the ethanol previously produced through respiration. This change from fermentation metabolism into respiratory metabolism is called diauxic shift, which involves intense genetic reprogramming (De Risi et al. 1997; Lashkari et al. 1997), and is based on an increase in the transcription of certain genes and repression of others.

To better understand these processes it is important to briefly mention how these genes are organized and how they work. Basically, the genes that encode proteins are organized into two independent regions, namely, a promoter region and a coding region. The promoter region is responsible for the transcription regulation that defines when a gene should be switched on or off. When the gene is switched on, in a portion close to the coding region, a messenger RNA (mRNA) starts to form and copies the information from the coding region, which will be translated later into a protein by ribosomes. It is fundamental to note that these regions of the gene act independently and may be interchangeable through genetic engineering, for example, leading to the formation of new genes with unique expression patterns (Lewin 1990).

In particular, the growth in a medium containing glucose leads to the repression of a series of genes involved with respiration and the use of alternative sources of carbon. The SUC2 gene is repressed under these conditions, thus preventing the release of further amounts of glucose



and fructose stored in the sucrose molecule (Gancedo 1998). The ADH1 gene, on the other hand, is induced by this sugar.

The end of fermentation activates a set of genes required for respiration and leads to intense multiplication of mitochondria (Mahler et al. 1975; Ulery et al. 1994). In particular, the absence of glucose in the medium activates the transcription of ADH2 (alcohol dehydrogenase 2) gene (Ciriacy 1979). The enzyme produced by the expression of this gene oxidizes ethanol to acetaldehyde via the reverse of the ADH1 pathway; acetaldehyde is then converted into acetate, which participates in the later stages of the Krebs cycle that is the major aerobic respiration cycle (Devlin 1997) (Figure 1).

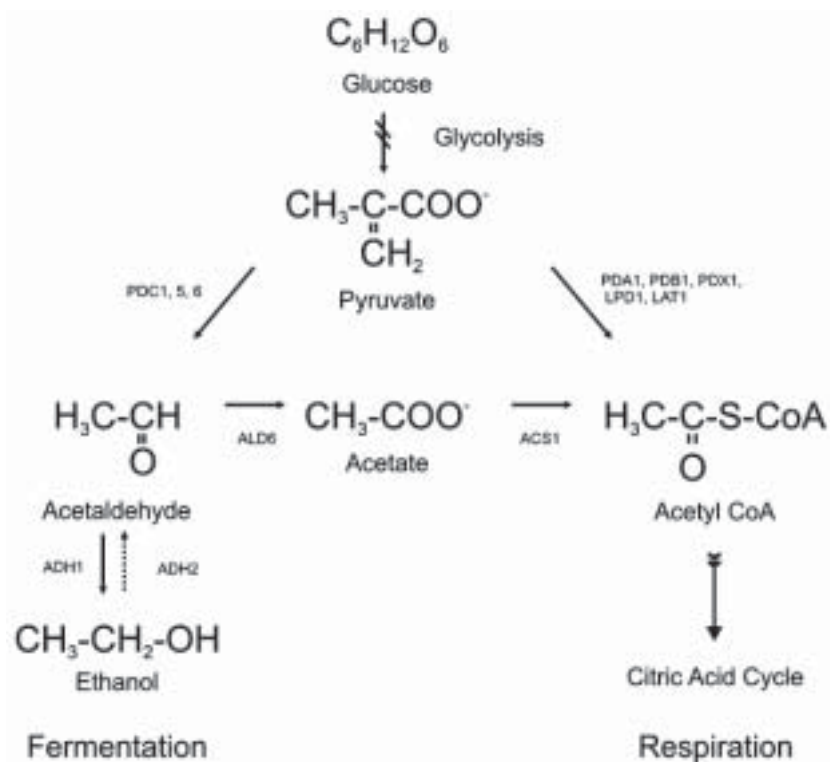


Figure 1. Schematic diagram of the fermentation and respiration reactions. The dotted arrows indicate the pathway that is activated after glucose depletion (ethanol oxidation). PDC1, 5, 6: pyruvate decarboxylase; PDA1, PDB1, PDX1, LPD1, LAT1: enzymes form the pyruvate dehydrogenase complex; ALD6: aldehyde dehydrogenase; ACS1: acetyl-CoA synthetase; ADH1: alcohol dehydrogenase 1; ADH2: alcohol dehydrogenase 2.

## Industrial Production of Ethanol

An interesting question is: why do yeasts produce ethanol? This would seem to be a waste if we consider it from the energetic perspective; respiration produces a much higher amount of energy compared to fermentation. However, from the strategic perspective, this seems to have been an effective evolutionary pathway. After all, only a few organisms are able to live in high concentrations of ethanol, a compound that is largely used as an antiseptic. While the yeast developed the ability to produce high concentrations of ethanol, the species also developed a system to resist this compound. Therefore, during the production of ethanol, the yeast can reduce significantly the competition from other organisms for the existing substrates and afterwards can consume the ethanol that was produced; while waiting for new stocks of sugar to be made available. If we look at the natural environments, this seems to be the logical strategy used by yeasts in fruits like grapes. As fruits fall from the vine, an environment is created that is similar to a fermentation vat, where ethanol is rapidly produced in large amounts. When new fruits fall from the vine this provides new stocks of sugar, and the entire cycle begins again. This parallel shows that the industrial fermentation of alcohol is just a scaled increase of a natural phenomenon. This process is not a human invention, but rather a human discovery.

Industrially, the production of alcohol follows basically two processes: fed-batch and continuous fermentation, with the first being the most used. In the fed-batch process, fermentation occurs in independent vats, beginning with low concentrations of yeast cells and high concentrations of substrate. The whole mixture of must (mixture of sugarcane juice and molasses), nutrients and yeast is added to the vat and after approximately 8 hours, once the fermentation has ended, the entire contents are removed and centrifuged. The "wine" – a name used in the sugar mills industry to indicate the fermented sugarcane juice without the yeast – is conveyed to the distillation towers, while the yeast milk, that is, the yeast broth obtained by centrifugation, is subjected to a decontamination treatment in the treatment vat consisting of a lowering of the pH with concentrated sulfuric acid. The yeast is then returned to

the vat and a new load of must is added, to start a new fermentation cycle (Wheals et al. 1999) (Figure 2).

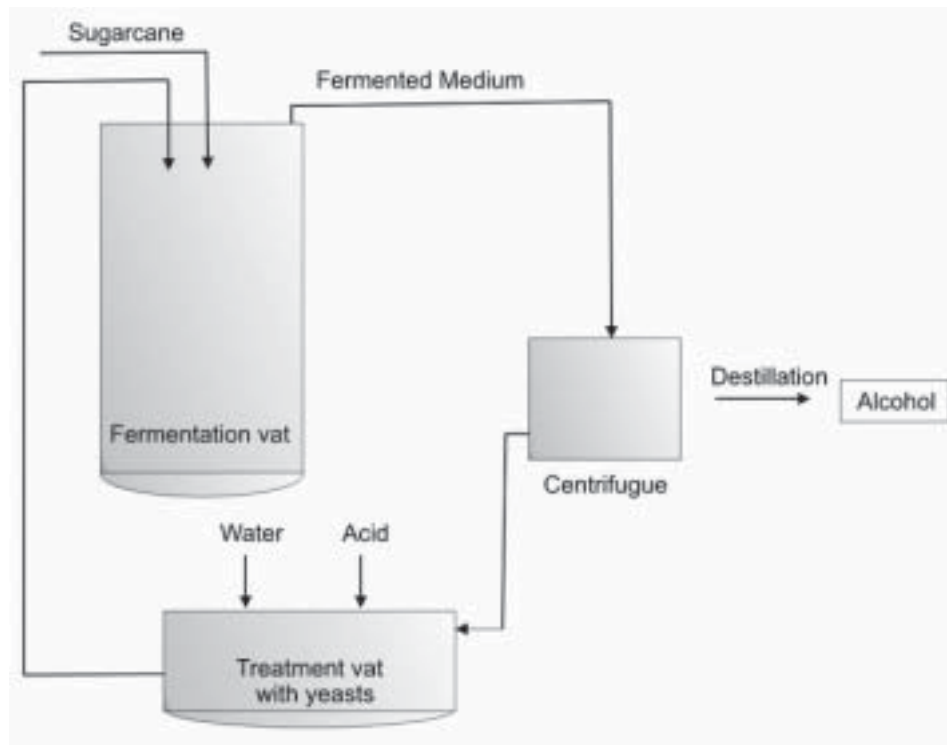


Figure 2. The fed-batch process of ethanol production. The whole mixture of must, nutrients and yeast is added to the vat and, once the fermentation has ended, the whole contents are removed and centrifuged. The wine is distilled and the yeast milk is returned to the vat, where a new load of must is added, starting a new fermentation cycle.

To start the fermentation, the original yeast used by sugar mills is usually multiplied under sterile conditions up to 50 liters and sent to a pilot stage until it reaches the volume required to start the production. This volume varies according to the size of the mill and can reach a value of up to 600,000 liters (Zarpelon and Andrietta 1992, Wheals et al. 1999).

An important observation that was only recently ascertained was that, although the yeast is inoculated at high levels, in some sugar mills it is rapidly replaced with wild yeast strains within a few ferment-

tation cycles and is then maintained until the end of the cropping season (da Silva Filho et al. 2005). In many cases this replacement is not even noted and there is no decrease in the productivity of the mill, possibly because the invading yeasts are better adapted to the conditions of the mill. Recently we have detected this fact in a large sugar mill in the Campinas region that for many years had been inoculating a strain to start the crop, and after a few weeks this strain was replaced with a wild strain that increased the process. This fact demonstrates clearly the low importance given to the fermentation agent, which is considered an abiotic factor. These factors may actually open up new opportunities for the development of new technological advances for improving the yeast strains used.

### The Yeast

This microorganism has been used in the production of bread and beverages for centuries. Since the nineteenth century it became an important experimental model, to the extent that it was the first eukaryotic organism to have its genome fully sequenced (Goffeau et al. 1996; Goffeau 1998), which tremendously simplified investigations to understand its metabolism. Moreover, *Saccharomyces cerevisiae* is an organism that can be easily genetically manipulated, which has led to a systematic deletion of all of its genes, either individually or in combination, in order to analyze their functions and determine how their products interact (Dujon 1998; Hauser et al. 1998). Full information is available for real time searches at the site [www.yeastgenome.org](http://www.yeastgenome.org).

The effective production of ethanol requires that yeasts included in the process have some essential characteristics such as: (1) ability to ferment rapidly and, effectively produce ethanol; (2) a high tolerance to ethanol; (3) a high osmotolerance (ability to ferment concentrated solutions of carbohydrates such as, for example, sugar molasses used in many sugar mills to produce ethanol); (4) genetic stability; (5) cellular viability and tolerance toward repeated fermentation cycles; (6) tolerance to temperature variations; (7) a high competitive capacity. There are several yeast species capable of producing ethanol on a large scale;

however, over 97% of processes involve the use of the species *Saccharomyces cerevisiae* (Stewart et al. 1987).

Therefore based on the knowledge we have of this organism, it is possible to manipulate it genetically to adequate it even further to the fermentation process. Several methods that have been used in the past have been to employed classical methods such as: cross hybridizations, selection and isolation of new strains.

The advent of molecular biology and the development of genetic engineering techniques made it possible to manipulate genes to increase the possibility of changing undesired characteristics or adding new desirable ones. Through transformation of an organism with modified genes, it is now possible to make the organism express certain characteristics that are not identified in natural populations. This means that it could be possible to build strains with specific functions, for industrial uses.

#### Yeast flocculation and its use in alcohol fermentation processes

A phenomenon that plays an important role in industrial production is yeast flocculation, that can be defined as the ability of certain strains to aggregate and form multicell masses that will separate by sedimentation or flotation (Stratford 1996; Stan and Despa 2000). This feature has been employed successfully in the beer industry to facilitate the removal of cells once the fermentation has ended (Stratford 1996). This precedent makes us think about it could possibly be used in a similar fashion in the ethanol production industry, since it would eliminate the need for centrifugation, an expensive and complicated step in the process.

In fact, this idea has led to the development of industrial procedures of alcohol production based on flocculating yeasts (Domingues et al. 2000, Domingues et al. 2001, Kondo et al. 2002). However, it was shown that continuous flocculation has a series of disadvantages and can lead to a reduction in the overall production of ethanol, as well as to in-process problems such as duct clogging (Stratford 1992, Verstrepen et al. 2003).

The physical and genetic mechanism of flocculation has been widely studied in recent years (Verstrepen et al. 2003). It involves primarily the interaction of surface proteins of a certain lecithin-like, sugar-binding

cell, with exposed sugar residues on surrounding cells. Thus, a cell producing a flocculation protein is capable of interacting with cells that do not produce the protein, and interact with the exposed sugars of those cells (Stratford and Carter 1993) (Figure 3).

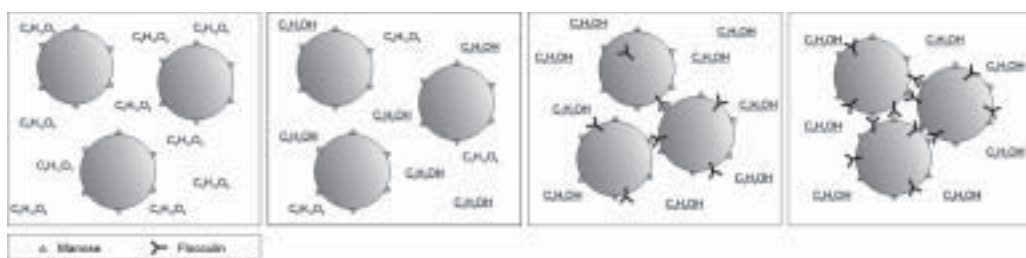


Figure 3. Schematic diagram of flocculation mechanism showing that all cells have sugar residues on their cell walls and that, after glucose ( $C_6H_{12}O_6$ ) depletion, flocculin – the flocculation-related protein – is formed.

Several genes have been genetically identified as being responsible for flocculation. Such genes express proteins called flocculins. The control mechanism for flocculins and proteins encoded them is different and has been intensely studied (Cormack 2004; Halme et al. 2004). The option used by industry has been to control flocculation by using engineering solutions. However, it seems that there is a huge potential for using biotechnological techniques to develop strains more adequate to the industrial process.

Based on this, we developed the following scenario: the ideal yeast should be perfectly soluble in the medium while this medium would include a fermentation substrate, i.e. sugar. The cells would be able to efficiently exploit this substrate because they would multiply and reach the maximum total contact surface. However, once the sugar was depleted, ideally we should have the opposite response, that is, the cells would aggregate and be deposited on the bottom of the vat. This would prevent ethanol consumption and the cells could be separate naturally, thus eliminating the need for centrifugation.

To determine if this idea was viable we used laboratory strains of *S. cerevisiae*, which are easily manipulated. At the same time we created a hybrid gene containing the promoter region of the ADH2 gene asso-

ciated with the coding region of the FLO5 flocculation gene. From this point on we had a desired effect. As mentioned before, the ADH2 gene promoter is repressed by glucose, and induced when this sugar is depleted from the medium. This mechanism selectively insures that the cells will continue to produce ethanol while preventing its oxidization to acetaldehyde when there is a flow in the opposite direction (Figure 1). At this point the flocculin encoded by the FLO5 gene is produced, and mobilized to the cell surface thus leading to a synchronized sedimentation (Figure 4). Starting with this simple system we were able to achieve a situation where we could sense the sugar levels in the medium, quench the fermentation at the appropriate moment and removing the yeast cells at that point. In a broader sense, we developed intelligent yeast – or, at least, smart yeast.

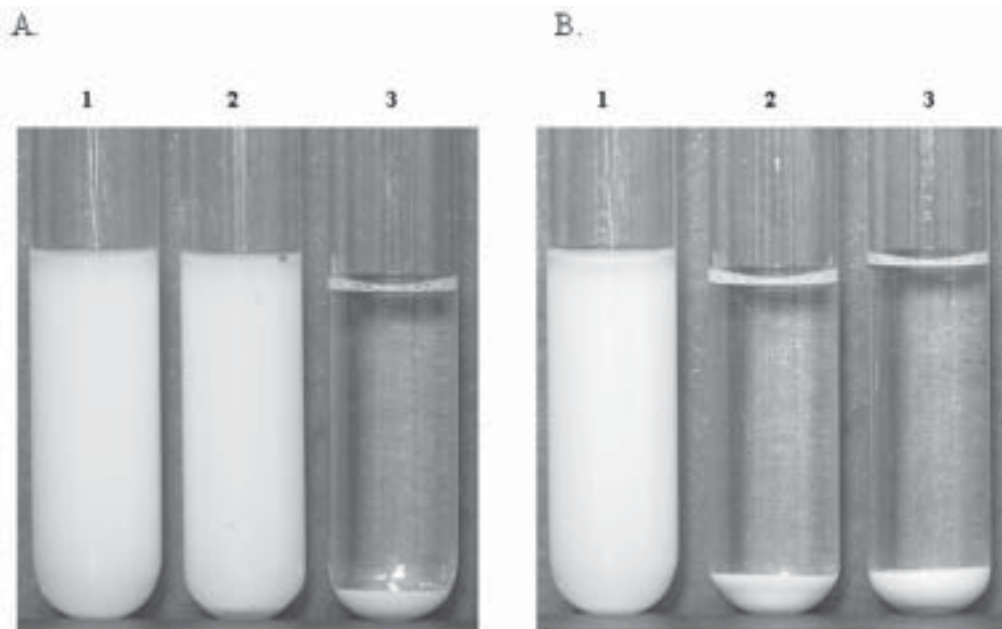


Figure 4. Expression of FLO5 gene under the control of the ADH2 gene promoter regulatory region.

1- Wild strain; 2- Wild strain with the flocculation gene controlled by glucose levels; 3- Wild strain with constitutively activated flocculation gene. A. Culture medium with glucose; B. Culture medium after glucose depletion. As shown in the figure, the strain whose flocculation gene is controlled by glucose levels is deposited after carbohydrate depletion.



This prototype worked perfectly in the laboratory, was patented (INPI # 0001122), and was awarded a Honorable Mention in the 2000 issue of the National Contest of Brazilian Inventions. It was the only laureate invention in the biotechnology area.

However, we know that a huge challenge is in front of us. As mentioned before, in the fermentation vats there is strong competition among microorganisms, and only the best-adapted strains reside in the process. Thus, our invention has no potential to produce a default strain that could be largely employed. On the contrary, we must be capable of identifying specific strains from the different sugar mills and modify them genetically so that they acquire the ability to flocculate according to the schematic diagram shown above. Recently we achieved this in the first industrial strains (Cunha et al. 2006) by integrating the flocculation cassette into the arginine permease gene present in all isolates of the species. This makes cells resistant to a drug called canavanine, allowing the selection of the transforming agents. However, the production of alcohol using these transforming agents is still unsatisfactory, and is the object of intense investigation.

Additionally, our laboratory team is committed to developing strains that are more resistant to ethanol and less capable of consume this compound. Theoretically, this can be achieved by deleting a few already-identified genes. This is an ongoing study and preliminary results are not yet ready to be disclosed at this time. With these first steps accomplished we are hopeful that it is possible to achieve higher concentrations of ethanol in the fermentation stage.

### Conclusion

Based on all current information, it appears that the demand for ethanol is no longer a national issue, but a global concern. This change in scale suggests that future ethanol programs will need to have a lot more stability than previous programs in order to meet future requirements. Thus, the search for an increase in ethanol production and in the efficacy of the process will dramatically increase and will be an opportunity for Brazil to develop high tech solutions on its own. Innovative improve-



ments have already been made to fermentation engineering and the automotive industry, particularly with the launching of FLEX fuel cars. However, major advances in the biological process of fermentation, or more specifically in the genetics of the yeast *S. cerevisiae* have yet to be accomplished. This report presents the development of conditional flocculating yeasts that have the potential of eliminating the need for centrifugation from the industrial fermentation process. When this system is available on the industrial scale, the possibility of having small alcohol/sugar mills spread all over the country using the fed-batch production process, will be feasible since the costly centrifugation step will be replaced with genetically modified “smart yeast”.

Anderson Ferreira da Cunha, Silvia Kazue Missawa  
and Gonçalo Amarante Guimarães Pereira

Institute of Biology, Department of Genetics and Evolution  
State University of Campinas - UNICAMP  
Campinas, SP – Brazil

## References

- Ciriacy M (1979). Isolation and characterization of further cis- and trans-acting regulatory elements involved in the synthesis of glucose-repressible alcohol dehydrogenase (ADHII) in *Saccharomyces cerevisiae*. *Mol Gen Genet* 176(3): 427-31.
- Cormack B (2004). Can you adhere me now? Good. *Cell* 116(3): 353-4.
- Cunha AF, Missawa SK, Gomes LH, Reis SF, Pereira GA (2006). Control by sugar of *Saccharomyces cerevisiae* flocculation for industrial ethanol production. *FEMS Yeast Res* 6(2): 280-7.
- da Silva-Filho EA, Brito dos Santos SK, Resende Ado M, de Morais JO, de Morais MA, Jr., Ardaillon Simoes D (2005). Yeast population dynamics of industrial fuel-ethanol fermentation process assessed by PCR-fingerprinting. *Antonie Van Leeuwenhoek* 88(1): 13-23.
- Demetrius FJ (1990). Brazil's national alcohol program: technology and development in authoritarian regime. New York.
- DeRisi JL, Iyer VR and Brown PO (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278(5338): 680-6.
- Devlin TM (1997). Textbook of Biochemistry with clinical correlations, Wiley-Liss.
- Domingues L, Lima N, Teixeira JA (2001). Alcohol production from cheese whey permeate using genetically modified flocculent yeast cells. *Biotechnol Bioeng* 72(5): 507-14.
- Domingues L, Vicente AA, Lima N, Teixeira JA (2000). Applications of yeast flocculation in biotechnological processes. *Biotechnol Bioprocess Eng* 5: 288-305.
- Dujon B (1998). European Functional Analysis Network (EUROFAN) and the functional analysis of the *Saccharomyces cerevisiae* genome. *Electrophoresis* 19(4): 617-24.
- Gancedo JM (1998). Yeast carbon catabolite repression. *Microbiol Mol Biol Rev* 62(2): 334-61.
- Goffeau A (1998). [The yeast genome]. *Pathol Biol (Paris)* 46(2): 96-7.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996). Life with 6000 genes. *Science* 274(5287): 546, 563-7.

Halme A, Bumgarner S, Styles C, Fink GR (2004). Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. *Cell* 116(3): 405-15.

Hauser NC, Vingron M, Scheideler M, Krems B, Hellmuth K, Entian KD, Hoheisel JD (1998). Transcriptional profiling on all open reading frames of *Saccharomyces cerevisiae*. *Yeast* 14(13): 1209-21.

Kondo A, Shigechi H, Abe M, Uyama K, Matsumoto T, Takahashi S, Ueda M, Tanaka A, Kishimoto M, Fukuda H (2002). High-level ethanol production from starch by a flocculent *Saccharomyces cerevisiae* strain displaying cell-surface glucoamylase. *Appl Microbiol Biotechnol* 58(3): 291-6.

Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO, Davis RW (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci USA* 94(24): 13057-62.

Lewin B (1990). Commitment and activation at pol II promoters: a tail of protein-protein interactions. *Cell* 61(7): 1161-4.

Mahler HR, Assimos K, Lin CC (1975). Derepression in *Saccharomyces cerevisiae* can be dissociated from cellular proliferation and deoxyribonucleic acid synthesis. *J Bacteriol* 123(2): 637-41.

Stan S, Despa F (2000). On the doublet formation in the flocculation process of the yeast cells. *Biosystems* 57(3): 139-45.

Stewart GG, Russel I, Klein RD, Hiebsch RR (1987). Biological Research on Industrial Yeasts.

Stratford M (1992). Yeast flocculation: a new perspective. *Adv Microb Physiol* 33: 2-71.

Stratford M (1996). Induction of flocculation in brewing yeasts by change in pH value. *FEMS Microbiology Letters* 136: 13-18.

Stratford M (1996). Yeast flocculation: restructuring the theories in line with recent research. *Cerevisia* 21(1): 38-45.

Stratford M, Carter AT (1993). Yeast flocculation: lectin synthesis and activation. *Yeast* 9(4): 371-8.

Ulery TL, Jang SH, Jaehning JA (1994). Glucose repression of yeast mitochondrial transcription: kinetics of derepression and role of nuclear genes. *Mol Cell Biol* 14(2): 1160-70.

van Haandel AC (2005). Integrated energy production and reduction of the environmental impact at alcohol distillery plants. *Water Sci Technol* 52(1-2): 49-57.

Verstrepen KJ, Derdelinckx G, Verachtert H, Delvaux FR (2003). Yeast flocculation: what brewers should know. *Appl Microbiol Biotechnol* 61(3): 197-205.

Vettore AL, da Silva FR, Kemper EL, Souza GM, da Silva AM, Ferro MI, Henrique-Silva F, Giglioti EA, Lemos MV, Coutinho LL, Nobrega MP, Carrer H, Franca SC, Bacci Junior M, Goldman MH, Gomes SL, Nunes LR, Camargo LE, Siqueira WJ, Van Sluys MA, Thiemann OH, Kuramae EE, Santelli RV, Marino CL, Targon ML, Ferro JA, Silveira HC, Marini DC, Lemos EG, Monteiro-Vitorello CB, Tambor JH, Carraro DM, Roberto PG, Martins VG, Goldman GH, de Oliveira RC, Truffi D, Colombo CA, Rossi M, de Araujo PG, Sculaccio SA, Angella A, Lima MM, de Rosa Junior VE, Siviero F, Coscrato VE, Machado MA, Grivet L, Di Mauro SM, Nobrega FG, Menck CF, Braga MD, Telles GP, Cara FA, Pedrosa G, Meidanis J, Arruda P (2003). Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. *Genome Res* 13(12): 2725-35.

Vincentz M, Cara FA, Okura VK, da Silva FR, Pedrosa GL, Hemerly AS, Capella AN, Marins M, Ferreira PC, Franca SC, Grivet L, Vettore AL, Kemper EL, Burnquist WL, Targon ML, Siqueira WJ, Kuramae EE, Marino CL, Camargo LE, Carrer H, Coutinho LL, Furlan LR, Lemos MV, Nunes LR, Gomes SL, Santelli RV, Goldman MH, Bacci M, Jr., Giglioti EA, Thiemann OH, Silva FH, Van Sluys MA, Nobrega FG, Arruda P, Menck CF (2004). Evaluation of monocot and eudicot divergence using the sugarcane transcriptome. *Plant Physiol* 134(3): 951-9.

Wheals AE, Basso LC, Alves DM, Amorim HV (1999). Fuel ethanol after 25 years. *Trends Biotechnol* 17(12): 482-7.

Zarpelon F, Andrietta SR (1992). Fermentação contínua para obtenção de álcool. *STAB (March-April)*: 23-28.



**ANEXO II**

**GENOME STRUCTURE OF A *SACCHAROMYCES CEREVISIAE* STRAIN  
WIDELY USED IN BIOETHANOL PRODUCTION.**

**JUAN LUCAS ARGUESO, MARCELO F. CARAZZOLLE,  
PIOTR A. MIECZKOWSKI, FABIANA M. DUARTE, OSMAR V. C. NETTO,  
SILVIA K. MISSAWA, FELIPE GALZERANI, GUSTAVO G. L. COSTA, RAMON  
O. VIDAL, MELLINE F. NORONHA, MARGARET DOMINSKA, MARIA G. S.  
ANDRIETTA, SÍLVIO R. ANDRIETTA, ANDERSON F. CUNHA, LUIZ H.  
GOMES, FLAVIO C. A. TAVARES, ANDRÉ R. ALCARDE,  
FRED S. DIETRICH, THOMAS D. PETES, AND GONÇALO A. G. PEREIRA**

**ARTIGO ACEITO PARA PUBLICAÇÃO NA GENOME RESEARCH**

Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production.

Juan Lucas Argueso<sup>1,9,10</sup>, Marcelo F. Carazzolle<sup>3,9</sup>, Piotr A. Mieczkowski<sup>6,9</sup>, Fabiana M. Duarte<sup>3</sup>, Osmar V. C. Netto<sup>3</sup>, Silvia K. Missawa<sup>3</sup>, Felipe Galzerani<sup>3</sup>, Gustavo G. L. Costa<sup>3</sup>, Ramon O. Vidal<sup>3</sup>, Melline F. Noronha<sup>3</sup>, Margaret Dominska<sup>1</sup>, Maria G. S. Andrietta<sup>4</sup>, Sílvia R. Andrietta<sup>4</sup>, Anderson F. Cunha<sup>5</sup>, Luiz H. Gomes<sup>7</sup>, Flavio C. A. Tavares<sup>7</sup>, André R. Alcarde<sup>8</sup>; Fred S. Dietrich<sup>1,2</sup>, John H. McCusker<sup>1</sup>, Thomas D. Petes<sup>1</sup>, and Gonçalo A. G. Pereira<sup>3,10</sup>

<sup>1</sup> Department of Molecular Genetics and Microbiology; and <sup>2</sup> Institute for Genome Sciences and Policy, Duke University Medical Center, Durham-NC, USA.

<sup>3</sup> Laboratório de Genômica e Expressão, Departamento de Genética e Evolução, Instituto de Biologia; and <sup>4</sup> Laboratório de Biotecnologia e Bioprocessos, Centro Pluridisciplinar de Pesquisas Químicas e Biológicas, Universidade Estadual de Campinas, Campinas-SP, Brazil.

<sup>5</sup> Departamento de Genética e Evolução, CCBS, Universidade Federal de São Carlos, São Carlos-SP, Brazil.

<sup>6</sup> Department of Genetics, School of Medicine, University of North Carolina, Chapel Hill-NC, USA.

<sup>7</sup> Departamento de Genética; and <sup>8</sup> Departamento de Agroindústria, Alimentos e Nutrição, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba-SP, Brazil.

<sup>9</sup> These authors contributed equally to this work.

<sup>10</sup> Corresponding authors.

Corresponding authors:

Juan Lucas Argueso, PhD

Department of Molecular Genetics and Microbiology, Duke University Medical Center,  
371 CARL Building, Box 3054, Durham-NC, 27710, USA.

Phone: 1-919-684-5814; Fax: 1-919-684-6033; email: lucas.argueso@duke.edu

Gonçalo A. G. Pereira, PhD

Laboratório de Genômica e Expressão, Departamento de Genética e Evolução, Instituto de Biologia,  
Universidade Estadual de Campinas, CP 6109, Campinas-SP, 13083-970, Brazil.

Phone: 55-19-3521-6237; Fax: 55-19-3521-6235; email: goncalo@unicamp.br

## ABSTRACT

In the search for renewable sources of energy, bioethanol stands out as a benchmark biofuel because its production is based on a proven technological platform. Bioethanol is produced mainly from the fermentation of carbohydrates derived from agricultural feedstocks by the yeast *Saccharomyces cerevisiae*. One of the most widely adopted strains is PE-2, a heterothallic diploid naturally adapted to the sugar cane fermentation process used in Brazil. Here we report the molecular genetic analysis of a PE-2 derived diploid (JAY270), and the complete genome sequence of a haploid derivative (JAY291). The JAY270 genome is highly heterozygous (~2 SNPs per kilobase), and has several structural polymorphisms between homologous chromosomes. These chromosomal rearrangements are confined to the peripheral regions of the chromosomes, and appear to reflect ectopic homologous recombination between repetitive DNA sequences. Despite the complex karyotype of JAY270, this diploid, when sporulated, had a high frequency of viable spores (93%). Crosses of haploids derived from JAY270 to a haploid of the unrelated laboratory strain S288c also resulted in diploids that had good spore viability (75-95%). Thus, the rearrangements that exist near the ends of chromosomes do not impair meiosis and spore viability, as they do not span regions that contain essential genes. This observation is consistent with a model in which the peripheral regions of chromosomes represent plastic domains of the genome that are free to recombine ectopically and experiment with alternative structures. We also explore features of the JAY270 and JAY291 genomes that help explain their high adaptation to industrial environments, exhibiting desirable phenotypes such as high cell mass production and fermentation kinetics, high temperature growth and oxidative stress tolerance. The genomic manipulation of such strains could enable the creation a new generation of industrial organisms, ideally suited for use as delivery vehicles for future bioenergy technologies.



## INTRODUCTION

As concerns mount over the alarming effects of climate change and of the continued volatility in petroleum markets, nations throughout the world are increasingly adopting policies to promote the use of renewable and domestic sources of energy (Robertson et al. 2008). Among the most viable alternatives, bioethanol stands out as a benchmark biofuel because its production is based on a proven technological platform. Bioethanol is produced through microbial fermentation of carbohydrates derived from agricultural feedstocks, mainly starch and sucrose. While the United States and Brazil are the dominant players in bioethanol, sharing approximately 70% of the global market, their production systems differ in many respects (Sanderson 2006; Goldemberg 2007). North American bioethanol is currently produced from enzymatically-hydrolyzed starch from grains. Farming this type of feedstock is energy intensive, resulting in a final energy balance that is only marginally positive. This process is currently regarded as a transitional technology, which in the coming years is expected to give way to advanced cellulosic biofuels (Gura 2009). The Brazilian system uses sugar cane as feedstock, a tropical grass crop that abundantly accumulates sucrose that can be converted to bioethanol directly by the yeast *Saccharomyces cerevisiae* without enzymatic pre-treatment. Among other advantages, sugar cane is semi-perennial and has symbiotic interactions with nitrogen-fixing microorganisms, allowing this system to produce an energy surplus estimated at about eight-fold (Goldemberg 2007; Robertson et al. 2008).

In the microbial fermentation process used in Brazil, the yeast inoculum is often recycled from one fermentation tank to the next (Wheals et al. 1999). In many distilleries, this recycling spans the entire eight-month sugar cane harvesting season (Basso et al. 2008), imposing both biotic and abiotic stresses on the fermenting yeast strain. While yeast strains genetically bred for high-efficiency in fermentation have been available since the 1980's (Tavares and Echeverrigaray 1987), they are often outcompeted in the industrial environment by more robust "wild" yeasts that contaminate the sugar cane stalks (Wheals et al. 1999; da Silva et al. 2005). In the 1990's, an alternative strategy was adopted: selection among the wild yeast contaminants for those that combined high fermentation efficiency with prolonged persistence in the system. In recent years, such strains have been widely adopted by the industry. One of the most successful examples is the PE-2 wild isolate (Basso et al. 2008), currently used by ~30% of Brazilian distilleries, generating about 10% of the world's bioethanol supply.

Thirteen years ago, the *S. cerevisiae* laboratory strain S288c became the first eukaryote to have its genome completely sequenced (Goffeau et al. 1996). Since then, other haploid strains from diverse backgrounds have been sequenced (RM11-1a, YJM789, M22, YPS163, and AWRI1631) (*S. cerevisiae* RM11-1a Sequencing Project - <http://www.broad.mit.edu/>) (Wei et al. 2007; Borneman et al. 2008; Doniger et al. 2008), and more recently a large scale effort to determine the genome sequences of many others has been completed (Liti et al. 2009).

Extensive analysis has been done to examine the nucleotide sequence diversity between these strains (Liti et al. 2009; Schacherer et al. 2009), whereas studies of structural variation have mostly focused on comparisons of *S. cerevisiae* to other related species (Fischer et al. 2000; Goffeau 2004; Scannell et al. 2007; Gordon et al. 2009). Microarray-based whole genome hybridization studies of wild, industrial, and laboratory *S. cerevisiae* strains (Winzeler et al. 2003; Carreto et al. 2008; Schacherer et al. 2009; Faddah et al. 2009) have uncovered a recurrent pattern of copy number variation (CNV) near the ends of chromosomes, suggesting a role for repetitive DNA sequences in structural genome diversification. Despite these valuable insights, two central questions regarding the role of chromosomal rearrangements in genome evolution in *S. cerevisiae* remain unanswered: First, it is still unclear how these rearrangements contribute to long-term fitness in natural environments; and second, it is not known if they are compatible with the formation of viable meiotic spores that would allow their sexual dissemination between natural *S. cerevisiae* populations.

In this study, we report a detailed molecular genetic characterization of a PE-2 derived diploid (JAY270), and the complete genome sequence of a derived haploid (JAY291). We have found extensive structural differences between the genome of JAY270 and those of other sequenced *S. cerevisiae* strains, including chromosomal translocations and insertions of large blocks of DNA absent in the reference strain's genome. These differences were confined to the peripheral regions of the chromosomes, while the central portions have remained structurally conserved. This pattern of structural variation was fully compatible with the formation of viable meiotic spores, and in some cases, the chromosome rearrangements resulted in the amplification of genes implicated in environmental stress tolerance. This observation supports a model in which *S. cerevisiae* chromosomes are organized in two structural domains (Pryde et al. 1997). Our data suggest a general chromosome organization in which the central core sectors, harboring the essential genes, are refractory to rearrangements, while the peripheral regions are highly plastic and are free to undergo ectopic recombination.

## RESULTS

### Genetic and phenotypic characterization

Before initiating a detailed characterization of PE-2, we used Pulse-Field Gel Electrophoresis (PFGE) to show that the commercially-available stocks of this strain contain a mixture of cells that have slightly diverged karyotypes (Supplemental Fig. 1). Consequently, we purified a single colony isolate with a representative karyotype. This isolate, JAY270, was indistinguishable from the original PE-2 culture in phenotypic tests (Fig. 1a, and data not shown) and was used in all subsequent experiments.

Upon induction of meiosis, JAY270 displayed rapid and efficient sporulation producing mostly asci with four ascospores. We dissected 104 tetrads and observed four viable spores per tetrad in 88 (93.3% spore viability). By PCR analysis, we found 2:2 segregation at the mating type locus (*MAT*; data not shown). These results show that JAY270 is a diploid, does not contain recessive lethal mutations, and is not heterozygous for

chromosome aberrations that span essential genes. The haploid segregants derived from JAY270 had stable *MATa* or *MAT $\alpha$*  genotypes, indicating that the strain is naturally heterothallic. We crossed haploid strains (both *MATa* and *MAT $\alpha$* ) derived from JAY270 to S288c-derived strains of the opposite mating type. The hybrid diploids produced in these crosses sporulated normally, had high spore viability (75-95%), and exhibited Mendelian segregation for all auxotrophic markers from the S288c-derived parent. These results argue that the chromosome structure of JAY270 is mostly collinear with that of S288c, at least in the regions required for viability.

We compared JAY270 and its derivatives to S288c in several phenotypic tests. In sugar cane extract fermentation assays that simulate the conditions found in the industry (Fig. 1a), JAY270 and all four spore clones from a JAY270 tetrad (segregants JAY289, JAY290, JAY291 and JAY292), produced roughly 50% more ethanol than the S288c strain. The observation that all four spore clones were equally proficient in ethanol production argues that the major alleles contributing to this complex phenotype are homozygous in the JAY270 diploid. We also compared growth phenotypes under a variety of stress conditions. Tolerance to ethanol, acid washing and short-term heat killing were not differently affected in these strains (Supplemental Fig. 2); however, JAY270 was much more tolerant to oxidative stress and high temperature growth than S288c (Fig 1b). When grown in the presence of menadione, an inducer of reactive oxygen species, JAY270 grew normally at concentrations that were inhibitory to S288c. JAY270 was also highly tolerant to temperature stress. While S288c had reduced growth rates at 37C, and did not form isolated colonies at 41C, JAY270 formed larger colonies at 37C than it did at 30C (Supplemental Fig. 3). At 41C, JAY270 displayed slower growth, but maintained full viability. These traits are likely important for long-term viability and competitiveness in the industrial setting, where oxidative stress (Landolfo et al. 2008) and heat are generated during high-pitch fermentation (fermentation involving large inocula).

Finally, we examined the patterns of genetic inheritance associated with two of the most important industrial traits of JAY270: cell mass accumulation and kinetics of ethanol production. To bypass the influence of factors specific to sugar cane extract (e.g. sucrose utilization), the fermentations were carried out in rich media with 10% glucose. We mated an S288c-isogenic strain to a haploid spore derived from JAY270 (JAY291) and generated hybrid diploids. Two reciprocal F<sub>1</sub> generation diploids (JAY361 and JAY365) were created by mating  $\rho^0$  mitochondrial *petites* derived from either parent (Fox et al. 1991). These two diploids were phenotypically indistinguishable from each other (data not shown), indicating that differences between the mitochondrial genomes of S288c and JAY291 do not have a major influence on the traits examined. We then evaluated the JAY291 and S288c haploid parents and the F<sub>1</sub> hybrid JAY361 in time-course fermentations. JAY291 displayed a higher growth rate and finished the fermentation with a cell density 34% higher than S288c; the JAY361 hybrid had an intermediate phenotype (Fig. 2a). Under the conditions used in this assay all three strains reached the same final concentration of ethanol at the end of the fermentation (Fig. 2b), but, JAY291 and

JAY361 were much more efficient, displaying higher kinetics of ethanol production and requiring ~30% less time than S288c to complete the fermentation. We have not tested diploidized versions of the JAY291 and S288c parents in these assays, therefore we cannot firmly ascertain dominance effects associated with these traits. However, if ploidy is not an influencing factor, our results would suggest that the cell mass accumulation trait is semi-dominant, whereas the kinetics of ethanol production trait is fully dominant.

We sporulated and dissected asci from JAY361 to isolate  $F_2$  haploids segregating genomic segments from the two parent strains. Individual clones in this progeny were then examined for final cell mass accumulation, and ethanol concentration after 8 hours of fermentation (the point of highest differential between the parents) as an indicator of ethanol production kinetics. The phenotypic distributions for these two traits in the progeny were quite distinct, and gave preliminary indications of the genetic control associated with each. The distribution of cell mass accumulation (Fig. 2c) appeared to be bimodal, with two distinct peaks centered around the phenotypes of the parents and a valley which corresponded roughly to the phenotype of the  $F_1$  hybrid. One interpretation of this pattern is that it was created by segregation of the two alleles of a single gene that is the major determinant of the cell mass accumulation phenotype. In addition to this major locus, the segregation of other minor contributing alleles was responsible for the bell-shaped distribution around each peak. The distribution of ethanol production kinetics (Fig. 2d) was more typical of a quantitative trait controlled by loci with roughly equal contributions to the phenotype. From this limited initial analysis, we can estimate that three or four genes are responsible for the segregation pattern observed for ethanol production kinetics.

We also asked whether there is a correlation between cell mass accumulation and ethanol production kinetics. This analysis (Fig. 2e) suggested that the two traits are largely independent of each other, and are presumably controlled by different genes. The calculated broad-sense heritabilities associated with both traits were very high ( $H^2 > 0.90$ ), indicating that most of the phenotypic variance observed between  $F_2$  individuals was due to differences in genetic composition.

### **Molecular karyotype analysis**

We compared the molecular karyotypes of JAY270 and S288c using a combination of PFGE, and microarray-based Comparative Genomic Hybridization (CGH-array) and Band-array (Argueso et al. 2008). PFGE analysis revealed clear chromosome (Chr) length polymorphisms between S288c and JAY270 (Fig. 3a). This analysis also showed polymorphisms between several pairs of homologs within the JAY270 diploid. Image tracing of the separated chromosomes in JAY270 revealed that some chromosomes were represented by two bands, each present in one copy in the diploid. For example, Chr6 is represented by short (Chr6S) and long (Chr6L) homologs, differing in size by about 60 kb. Interestingly, karyotypic analysis of the meiotic haploid spores revealed parental-sized chromosomes in JAY289 and JAY290, and chromosomes with the sizes expected for reciprocal recombinants between Chr6S and Chr6L in JAY291 and JAY292 (Fig. 4a; second band

from the bottom). In addition to Chr6, at least Chr3, Chr5, Chr9, Chr11, Chr14, and either Chr7 or Chr15 had length polymorphisms between homologs in JAY270.

We examined the structural polymorphisms between the Chr6 homologs by excising their respective chromosomal bands from PFGE, labeling the DNA and hybridizing them to microarrays covering the entire S288c reference genome (Band-array; Fig. 5). The JAY270 Chr6 homologs differed from each other and from Chr6 in S288c only near the chromosome ends, in regions distal to the first and last essential genes (*SEC53* and *RPN12*). This analysis showed that Chr6L contained a ~16 kb region on the left arm that was missing in Chr6S (including the *AGP3* gene), and that Chr6S contained a ~12 kb translocation from the right arm of Chr1 (including the *YAR064W* gene) that was missing in Chr6L. Southern analysis (Fig. 4b-c) confirmed this arrangement and revealed an additional translocation involving *AGP3* on one of two Chr10 homologs of JAY270. Notably, the cumulative Band-array hybridization signal for Chr6L (~322 kb in size, Fig. 2a), accounted for only ~250 kb of sequences, suggesting the presence of about 70 kb of DNA that did not hybridize to the microarrays.

We used CGH-arrays to look for sequences that were missing or duplicated in JAY270 relative to S288c (CNVs; Fig 3b). Most regions of the genome were present in equal dose between these strains. The regions shown in green near the telomeres of Chr1, Chr4, Chr6, Chr10, and Chr15 were underrepresented in the JAY270 genome. These could be either missing in JAY270, or represent diverged sequences that do not hybridize to the S288c-based microarrays. The underrepresented regions that were not telomeric corresponded to regions containing tandemly repeated genes in S288c (*HXT7*, *HXT6*, *HXT3* and *ENA5*, *ENA2*, *ENA1* in Chr4; and *ASP3-1* to *ASP3-4* in Chr12) that likely contracted in JAY270 or expanded in S288c as a consequence of unequal crossing-over. The regions shown in red represent sequences that were amplified in JAY270 relative to S288c. An amplification signal detected in the central section of Chr15 corresponded to the *HIS3* marker locus. This gene is present in JAY270, but was intentionally deleted in the derivative of S288c that was used in the CGH-array experiments. Thus, we can effectively detect a single amplified gene by these methods.

The amplification at the left end of Chr16 spans the *SAM3* and *SAM4* genes that are involved in metabolism of S-adenosylmethionine (AdoMet), a key biochemical cofactor that participates in a variety of metabolic pathways. The Sam4p has been proposed to participate in the recycling of AdoMet from the inactive, and possibly toxic, (*R,S*)-AdoMet isomeric state back to the biologically essential form (*S,S*)-AdoMet (Vinci and Clarke 2007). Based on CGH-array and Southern blot data (Fig. 3b, and data not shown), we estimate that four copies of these genes are present in JAY270 compared to two in the laboratory diploid strain.

Another interesting gene amplification resulted in two peaks at the left ends of Chr6 and Chr14. In S288c these regions contain, respectively, the duplicated *SNO3* and *SNZ3*, and *SNO2* and *SNZ2* genes which are involved in vitamin B6 metabolism, and have a role in oxidative stress tolerance (Padilla et al. 1998; Ehrenshaft et al. 1999; Rodriguez-Navarro et al. 2002). To determine the extent and chromosomal distribution

of the *SNO/SNZ* gene amplifications in JAY270, we did a Southern analysis of pulse-field gels for this region (Fig. 4d). We detected nine copies of the *SNO/SNZ* genes in JAY270, compared to four in the S288c diploid. Consistent with this observation, we found that expression levels of these genes were up-regulated about four-fold in JAY270 relative to S288c (Supplemental Table 1). These results also showed that *SNO/SNZ* genes were redistributed among the chromosomes in JAY270 with no copies on either Chr6 homolog and new copies on Chr9, Chr10, Chr13 or Chr16, Chr7 or Chr15, and Chr4 (Fig. 4d). Strikingly, three of the extra copies were hemizygous (only one copy of a gene is present in a pair of homologous chromosomes) in the JAY270 diploid as could be inferred from their Mendelian segregation among the haploid spores JAY289-JAY292.

### **Genome sequencing of JAY291, a haploid spore derived of JAY270**

In addition to the molecular genetic analysis described above, we employed massively parallel DNA sequencing to characterize the genome of JAY270. To avoid complications due to heterozygosity in the diploid, we sequenced the genome of a haploid derivative, JAY291. Two independent DNA sequencing platforms were used, Roche/454 and Illumina, resulting in 162-fold coverage (Table 1). Despite this high level of redundancy, because both of these methods generated relatively short sequence reads (<250 bp), we were still unable to assemble the genome into one single contiguous sequence per chromosome. Instead, our *de novo* assembly of the JAY291 genome resulted 452 genomic fragments (contigs; Materials and Methods). As expected, when we compared these contigs to the S288c reference genome, it became clear that most interruptions occurred at regions of dispersed repetitive DNA sequences, mostly Ty retrotransposon insertions. As described below, we compared the sequences of JAY291 to those of other sequenced *S. cerevisiae* strains: S288c, RM11-1a, and YJM789 (Wei et al. 2007). These sequence comparisons were done to characterize chromosome structure differences between these strains, as well as to determine the level of nucleotide sequence divergence.

### **Chromosome structure polymorphisms**

From the PFGE profiles of JAY270 and its haploid descendants (Figs. 3a and 4a) it was clear that this diploid has chromosomes that vary in size, not only in comparison to S288c, but also between the two homologs. To investigate the nature of these rearrangements, we focused our analysis on Chr6. In JAY270, the two homologs were approximately 261 kb (Chr6S) and 322 kb (Chr6L) in size. The Chr6 of JAY291 is a recombinant between Chr6S and Chr6L, and is about 290 kb in size. We assembled the complete sequence of Chr6 in JAY291 by searching the 452 whole genome contigs for known Chr6 sequences from the S288c, RM11-1a and YJM789 strains; this search identified 14 candidate contigs. We then used PCR to amplify across the small gaps between the contigs and to position telomeric repeats, confirming their relative order and orientation (data not shown). The size of the resulting assembled sequences (including one X and one Y' element as each telomere) was 290 kb, matching the observed PFGE size. The chromosome structure and nucleotide sequence

in the central portion of Chr6 were very similar to those of S288c, whereas in the regions near the telomeres the sequences were more diverged or completely different (Fig. 6a). At the left end of the Chr6 alignment, we found two segments (9.1 and 19.3 kb) that are absent in the S288c genome. We also aligned the JAY291 chromosome to the Chr6 sequence from *Saccharomyces paradoxus* (Kellis et al. 2003), the closest known relative of *S. cerevisiae*, and observed continual synteny through most of Chr6, including the entire left arm. This alignment suggests that the 9.1 and 19.3 kb insertions were present in the last common ancestor of *S. cerevisiae* and *S. paradoxus*, and that they were probably lost from the progenitor of S288c. The only structural difference in the central section of Chr6 was a Ty2 retrotransposon element insertion which is present in S288c (*YFLWTy2-1*), but absent in JAY291. Chr6 had no full-length retrotransposons in JAY270, but Ty elements were found in all other chromosomes (Supplemental Fig. 4).

Although we concentrated our sequencing efforts on the JAY291 haploid, from our analysis of the sequences of S288c, RM11-1a, and YJM789 and from our examination of the microarray data and Southern analysis (Figs. 3-5), we were able to predict the structures at the left ends of the two Chr6 homologs in the JAY270 diploid. We then designed specific PCR primers to confirm our predictions (Supplemental Fig. 5). Repetitive DNA sequences were found at the breakpoints (shown as boxed letters in Fig. 6b) of all chromosomal rearrangements in this region, suggesting that they formed through ectopic homologous recombination.

We determined that JAY270 Chr6S has a portion of Chr1 that includes the *YAR064W* gene attached to the left end of Chr6 (Fig. 4c). This translocation is also present in RM11-1a and YJM789, and the breakpoint is between *YAR062W* in Chr1 (boxed "A") and *YFL051C* in Chr6, both are pseudogene members of the *FLO* gene family. The *YFL051C* pseudogene is also the site of the 19.3 kb region which is present in JAY270 Chr6L and deleted in Chr6 from S288c. While *YFL051C* is short and non-functional in S288c, in JAY270 Chr6L this ORF extends into the 19.3 kb region to form a full length *FLO* gene (boxed "B"). Near the left end of this insertion region, there was a 4.3 kb translocated section of S288c Chr10, with breakpoints at *HXT8* (boxed "C"; a hexose transporter gene family member) and *YJL216C* (boxed "D"; a gene with similarity to maltase genes). Finally, the leftmost rearrangement in JAY270 Chr6L contains the 9.1 kb sequence not represented in S288c. The breakpoint for this rearrangement was at the *THI5* gene (boxed "E"; a member of a gene family involved in the thiamine biosynthesis). In S288c, the *SNO3* and *SNZ3* genes are located immediately distal to *THI5*, whereas this region of Chr6L of JAY270 contains a copy of the *MPR2* gene involved in tolerance to oxidative stress and ethanol toxicity (Du and Takagi 2007). *MPR2* is also present in the RM11-1a and  $\Sigma$ 1278b strains, but strikingly, the genes in *MPR* family are found at different genomic locations in these strains: in RM11-1a, *MPR2* is at the left end of Chr10; whereas in  $\Sigma$ 1278b, *MPR2* is at the right end of Chr10 and *MPR1* is at the left end of Chr14 (Takagi et al. 2000).

This detailed analysis also allowed us to examine the rearrangements between the two Chr10 homologs in JAY270 (arbitrarily designated A and B). The left end structure of Chr6L was essentially identical to that

found in Chr10 of RM11-1a and in JAY270 Chr10A (inherited by the JAY289 and JAY292 spores). A similar Chr6/Chr10 translocation with a breakpoint at *YJL216C* (boxed “D”) has been described recently in the M22 vineyard strain (Doniger et al. 2008). The Chr10 structure in YJM789 is similar to RM11-1a, M22, JAY270 Chr10A, with the exception of the extreme left end at positions distal to *THI5* (boxed “E”) where YJM789 is similar to S288c Chr6. Interestingly, the YJM789 Chr10 configuration is precisely what would be predicted from a crossover occurring between S288c Chr6 and RM11-1a Chr10 anywhere between *THI5* and *YFL051C*. An analogous meiotic crossover event within translocated chromosomal segments has been recently observed in the S288c/Y101 laboratory hybrid (Faddah et al. 2009). The second JAY270 Chr10 homolog (Chr10B; inherited by JAY290 and JAY291) had at its left end a ~15 kb translocated segment not found in the S288c genome. The breakpoint in this chromosome is near an unannotated Ty1 delta LTR element (boxed “F”) that is located on S288c Chr10 between SGD coordinates 29000 and 29500 and is also present in the JAY291 Chr10 sequence. Finally, it is important to note that all rearrangements described above occurred at sites distal to the first essential genes in the Chr6 (*SEC53*) and Chr10 (*PRP21*).

### **Nucleotide polymorphisms**

We compared the sequences of JAY291 to S288c, estimating the number of SNPs at 5.4 per kilobase (about 65000 for the entire genome). To compare JAY291 to other *S. cerevisiae* strains, we analyzed a ~49 kb region from Chr14 that has been sequenced for several unrelated strains (Steinmetz et al. 2002) and used it to construct an unrooted phylogenetic tree (Fig. 7a). This analysis indicated that JAY291 is highly diverged relative to S288c, YJM789, and RM11-1a and, therefore, will be useful in characterizing the rich sequence diversity present in *S. cerevisiae*.

We also sought to estimate the degree of heterozygosity in the JAY270 diploid genome. We performed a limited sequence analysis of JAY292, a haploid spore derived from the same tetrad as JAY291. Because the genomic coverage in this case was relatively low (15-fold with Illumina paired-end reads), we did not attempt to assemble the genome *de novo*. Instead, we assembled the JAY292 sequence using the JAY291 contigs as reference. After filtering out repetitive elements, regions of high divergence, and regions of low quality and/or coverage, we estimated a density of about 1.3 SNPs/kb in the single-copy genomic regions. Because JAY291 and JAY292 are sibling haploid strains, about half of their genomes should be identical by descent and we found that about half of the genomic regions had a very low level of polymorphisms (data not shown). We estimate, therefore, that the JAY270 diploid has about 2.6 SNPs/kb between allelic regions in homologous chromosomes.

We also examined heterozygosity restricting the analysis to the Chr6 sequences of JAY291 and JAY292. Because these two haploids carry the two reciprocal products of a meiotic crossover between the two parental Chr6 homologs (Fig. 4a), it should be possible to directly find all Chr6 SNPs in the diploid by comparing the JAY291 and JAY292 sibling sequences (assuming that they formed as a result of exactly one crossover).



We found 465 SNPs in the 208 kb region common to both the Chr6L and Chr6S homologs (Fig. 5). The Chr6 SNP density (~2.2 SNPs/kb) was consistent with our estimate for the whole genome.

### Overview of specific gene polymorphisms important for bioethanol production

We compared the sequences of several JAY291 genes to those of other strains to establish possible links between the genotype and the observed phenotypes. High-Temperature Growth (HTG) has been investigated in *S. cerevisiae* as a model for complex quantitative inheritance (Steinmetz et al. 2002; Sinha et al. 2008), and three major contributing loci have been identified: *NCS2*, *MKT1*, and *END3*. We found that JAY291 had all three of the alleles that are related to HTG<sup>+</sup> (Fig. 7b). However, HTG, as is often the case for quantitative traits, is known to be influenced by genetic background (Sinha et al. 2006). We therefore tested the role of these polymorphisms in the JAY291 background by constructing isogenic strains carrying HTG<sup>-</sup> allele replacement versions of *NCS2* and *MKT1* from S288c, and *END3* from YJM789. We found that the allele replacement strains, individually, were not compromised for HTG, at least within the resolution of our co-culture competition assays (~5 fold; data not shown). This result was analogous to the background-specific phenotypic response described by Sinha et al. (2006), and suggested that other uncharacterized alleles present in the JAY291 genome are contributing to its distinct HTG<sup>+</sup> phenotype.

We also examined the sequence of the *HO* gene that encodes the endonuclease that stimulates mating type switching in homothallic strains (Herskowitz 1988). The *ho* allele in JAY291 has three missense mutations, including the H475L substitution that is present in S288c and has been shown to significantly reduce HO activity (Meiron et al. 1995; Ekino et al. 1999). In addition, an in-frame deletion between 8 bp direct repeats within the gene removed 36 amino acid residues from the DNA binding domain (residues 524-559). These mutations likely eliminate the endonuclease activity of HO, explaining the heterothallic life cycle of JAY270.

Flocculation is a mechanism through which yeast cells aggregate to form clumps, or flocs. While flocculation is a desirable trait in brewing, it is problematic in sugar cane batch fermentation because it significantly slows fermentation kinetics, and can cause excessive foaming and clogging of pipes. One of the attractive properties of JAY270 is that it rarely flocculates and produces very little foam in fermentation tanks (Basso et al. 2008). The JAY291 haploid is also non-flocculant. We examined the sequences of genes encoding cell surface adhesins involved in flocculation (*FLO1*, *FLO5*, *FLO11*, and others) and found that they are present in JAY291, although, due to their internal repeat structures (Verstrepen et al. 2004), these sequences were sometimes split between two contigs in our assembly. In addition, we looked at the *FLO8* gene, a positive regulator of flocculation. In S288c, *FLO8* contains a nonsense mutation (*flo8-1*) that renders the cells non-flocculant (Liu et al. 1996). The *FLO8* gene in JAY291 appears to be functional, indicating that the non-flocculant phenotype of JAY291 involves a block in a different step of the pathway. Interestingly, among 120 F<sub>2</sub> progeny from the S288c x JAY291 cross described above (Fig. 2) we observed 19 individuals that

flocculated in fermentation media and 101 that did not; roughly a 1:7 phenotypic ratio expected for a trait controlled by three unlinked genes ( $\chi^2 = 1.22$ ;  $p = 0.27$ ). We genotyped the flocculant progeny by PCR and restriction digest and observed that all 19 flocculant progeny inherited the *FLO8* allele from JAY291, therefore confirming the requirement of a functional Flo8p for flocculation.

In the S288c genome, there is a Ty1 retrotransposon insertion in the *HAP1* gene (Gaisne et al. 1999). *HAP1* encodes a heme-mediated transcription factor that controls the expression of genes involved in fermentation (Mense and Zhang 2006), and in the biosynthesis of ergosterol which is important for ethanol tolerance (Inoue et al. 2000). As in other industrial strains (Tamura et al. 2004), this interrupting Ty1 element is absent in JAY291. Consistent with this result, we found that several of the Hap1p regulated genes were expressed at a higher level in JAY270 than in S288c (Supplemental Table 1).

Another gene important for industrial yeast strains is *MIP1*, a nuclear-encoded mitochondrial DNA polymerase (Foury 1989). A polymorphism present in Mip1p (Mip1p-661A) has been shown recently to be responsible for the high frequency of spontaneous cytoplasmic respiratory mutations ( $\rho^-$  *petites*) observed in S288c (Baruffini et al. 2007). The *MIP1* sequence in JAY291 contains five non-synonymous substitutions relative to S288c, including the Mip1p-661T variant, which is associated with a low frequency of *petites*. This desirable trait was confirmed in semi-quantitative *petite* mutation assays in which JAY270 produced *petite* colonies at a frequency about five fold lower than S288c (data not shown).

Efficient sucrose utilization is essential in sugar cane fermentation. *S. cerevisiae* strains have long been known to vary in the number of copies of the *SUC* genes coding for invertase, the enzyme that breaks down sucrose in glucose and fructose which are then fermented to produce ethanol (Carlson and Botstein 1983). S288c only has one copy, *SUC2*, the only non-telomeric member of this gene family. Since JAY270 thrives in the sugar cane extract, we expected to find several *SUC* copies in its genome. Instead, the only invertase gene found in the genome sequence of JAY291 was *SUC2*. Another gene we did not find in JAY291 was *RTM1* (Ness and Aigle 1995), involved in resistance to toxins found in molasses, a byproduct of sugar production sometimes used in bioethanol fermentation. Incidentally, *RTM1* is always found clustered with the telomeric copies of *SUC* genes (Ness and Aigle 1995).

A comprehensive list of all genes found in both JAY291 and S288c is presented in Supplemental Table 1, including the local density of SNPs inside the coding sequences and in the upstream regulatory regions, and the relative level of gene expression between these strains measured with cDNA microarrays. In addition to the shared genes discussed above, it is possible that the genes present in JAY291 but absent in S288c also contribute to the properties of JAY270. At least sixteen such genes have been identified in the genome of JAY291 (Supplemental Table 1). An attractive example of such genes are two copies of putative S-adenosylmethionine-dependent methyltransferases found in the JAY291 genome (contig404-gene1 and contig386-gene1; their predicted protein sequences share 92% similarity). These new genes participate in the

same processes as the amplified *SAM3* and *SAM4* genes discussed above. Interestingly, a protein identical to the one coded by contig404-gene1 has been recently identified and found to be specific to the West African lineage of *S. cerevisiae* strains (hypothetical protein 5; Liti et al. 2009).

## DISCUSSION

### Diverse genome structure of JAY270 and its implications for genome evolution

Our molecular analysis of the JAY270 diploid revealed that its genome is highly heterogeneous, both structurally and at the nucleotide level. Thus far, the genomic analysis of *S. cerevisiae* had been mostly limited to haploids, often derived from homothallic diploid strains. In this type of life cycle, immediately after meiosis, the haploid spores undergo mating type switching, followed by self mating to restore a diploid state (Herskowitz 1988). This cycle generates diploids that are homozygous for the entire genome, except at the *MAT* locus. In contrast, heterothallic spores, such as those derived from JAY270, are unable to switch mating type because of a mutation in the *HO* gene. The spores restore the diploid state by mating to spores of the opposite mating type, either siblings from the same ascus or by outcrossing (Knop 2006). The heterozygous JAY270 genome richly illustrates how this process is effective in preserving and amplifying genetic diversity. It is important to note, however, that both the homothallic and heterothallic life cycles require the production of viable haploid spores in meiosis, and thus both types of strains are expected to continuously remove from the gene pool recessive lethal mutations and chromosomal rearrangements associated with the loss of essential genes.

We have found that the genomes of JAY270 and S288c are collinear through most regions and hybrids generated by crosses between these genetic backgrounds have good spore viability. Structural variation, though abundant, was limited to the end sections of chromosomes. A similar pattern has been previously observed in microarray-based surveys of genome variation among different *S. cerevisiae* strains (Winzeler et al. 2003; Carreto et al. 2008; Schacherer et al. 2009), which found that most polymorphisms, including SNPs and CNVs, were more common within 25 kb of the telomeres. In these studies, however, the full extent of the structural variation in these regions could not be revealed because the microarrays used were based on the S288c genome, and regions of high sequence divergence or sequences that were absent in the reference strain could not be examined. This gap has been recently filled by a whole genome sequencing-based survey of dozens of *S. cerevisiae* and *S. paradoxus* strains (Liti et al. 2009). This study identified DNA sequences, including 38 genes, that are present in some of these strains but are absent in the reference S288c genome. Importantly, these new genes are mostly subtelomeric.

The peripheral regions of chromosomes in *S. cerevisiae* typically contain genes that participate in alternative carbon source and vitamin metabolisms, ion and amino acid transport, flocculation and other processes that are not essential for viability. In the S288c genome, ~8% of all genes are located at sites distal to essential genes (Supplemental Table 1). Some of these genes, like *SNO/SNZ* and *SAM*, are functionally

clustered. Previous studies have shown that there are very few functionally clustered genes in *S. cerevisiae*, and that those which are clustered tend to be associated with specific functions that differ between strains (Hall and Dietrich 2007). Analogous distributions of non-essential genes are also found in other species. For example, in *Plasmodium falciparum*, the peripheral regions encode variable surface antigen genes (contingency genes) that help this parasite evade the host's immune defenses (Freitas-Junior et al. 2000; Barry et al. 2003). Many of the genes found in such regions are members of multi-copy gene families that allow structural variation to occur through ectopic homologous recombination events similar to those detected in JAY270. In addition to their repetitive nature, the telomere-proximal regions of chromosomes are often spatially co-localized at the nuclear envelope, forming clusters of chromatin that are believed to facilitate ectopic recombination interactions (Palladino et al. 1993; Gasser et al. 2004). Pryde *et al.* (1997) integrated these observations to formulate a model in which the yeast genome is structured in chromosomal sectors that are either rigid (central core) or plastic (peripheries). A schematic model of this concept as it relates to a set of *S. cerevisiae* haplotypes is shown in Fig. 8.

We recently demonstrated a mechanism in diploid strains that explains how repetitive elements participate in ectopic homologous recombination through their differential interactions with DNA double strand breaks (DSBs; Argueso et al. 2008). While lesions in single-copy regions of the genome are efficiently repaired to restore the original structures, DSBs that occur within repetitive sequences often lead to chromosomal rearrangements by engaging in homologous recombination with ectopic repeats. This experiment imposed no restrictions regarding the fitness effect or the haploid viability associated with the rearrangements, thus, we identified re-structuring events spanning the entire *S. cerevisiae* genome. Because most rearrangements were non-reciprocal and involved regions containing essential genes, it is likely that those events would be deleterious in meiosis, generating inviable spores and therefore eliminating them from the population. In contrast, the naturally-occurring rearrangements observed in JAY270 are not detrimental, since the translocations do not span genes essential for viability. At least six chromosomes in JAY270 were polymorphic in length and yet the diploid had excellent spore viability. Further, crosses between spores derived from JAY270 and S288c generated hybrid diploids in which structural differences were present in virtually every homolog pair. Despite their heterogeneous karyotype, these hybrids produced mostly viable spores (75-95%). We propose that all DSBs that occur in repetitive elements of the genome are able to promote ectopic recombination and generate chromosomal rearrangements. Those that span essential genes result in poor spore viability and are, therefore, eventually eliminated from the population, whereas those that involve the peripheral chromosome regions do not affect spore viability and may lead to increased fitness.

A distinguishing feature of the JAY270 genome architecture that may contribute to its fitness in competitive environments is the hemizygous distribution of useful stress tolerance genes. It is possible that this configuration offers alternative mechanisms for diploid cells to adjust the copy number of these genes through

mitotic crossover, break-induced replication, or meiotic segregation followed by re-mating of sibling haploid spores (Paques and Haber 1999; Knop 2006; Llorente et al. 2008). For example, these mechanisms could, in just a few generations, produce diploids containing 0 to 4 copies of *MPR* genes and 6 to 12 copies of *SNO/SNZ* genes. Notably, once the homozygous state is achieved, it may be reversed over time through outcrossing or ectopic homologous recombination between neighboring repeats. This scenario may explain how the highly plastic genome structure at the peripheral regions of chromosomes could enable *S. cerevisiae* to rapidly respond to a changing environment.

Finally, it is interesting to contrast the pattern of spatially restricted chromosomal rearrangements in *S. cerevisiae* to the much more radical chromosomal rearrangements recently described in the haploid asexual yeast pathogen *Candida glabrata* (Polakova et al. 2009). These very complex genome reconfiguration events included chromosome fusions, large translocations, aneuploidy and even possible circularizations, that are completely incompatible with meiosis. These authors proposed that *C. glabrata* has taken genome restructuring to the extreme, sacrificing its sexual life cycle to fully explore the adaptation potential provided by a highly unstable genome structure. We believe that *S. cerevisiae* has found the middle ground by confining the genome rearrangements to the ends of chromosomes.

### **Implications of highly adapted industrial strains to future bioenergy technologies.**

In addition to the insights into genome evolution discussed above, the analysis of the JAY270 genome offers a number of opportunities for the development of a new generation of industrial yeast strains. The finding that JAY270 is genetically compatible with laboratory strains indicates that the full arsenal of research tools developed to study yeast as a model system can be readily applied to the modification of this naturally adapted strain. We envision the use of JAY270 as an ideal delivery vehicle for future renewable energy technologies. One of the most active fields of investigation in the effort to develop these technologies has been the search for strategies to efficiently break down cellulosic feedstocks and to utilize its complex derived sugars in bioethanol fermentation (Gura 2009). These advances would dramatically improve the overall energy balance in bioethanol production. While the early signs of success in this effort are very encouraging, it is important to remember that even when the solutions for this problem are found, there will be a need to efficiently deliver these technologies to the industrial setting. Decades of experience in the development of industrial strains have convincingly demonstrated that any large-scale attempt to introduce genetically modified yeast strains in the bioethanol industry will be futile, unless they are based on naturally adapted strains. Therefore, strains such as JAY270 are likely to play a key role in facilitating the transition from laboratory technological breakthroughs to industrial scale field applications.

## **METHODS**

**Growth media and yeast strains.** Yeast culture, mating, sporulation, and tetrad dissection were conducted using standard procedures (Rose et al. 1990). Menadione was obtained from Sigma-Aldrich. The PE-2 stock used to purify JAY270 was from our own laboratory collection, which was derived from commercially available dry active yeast sold to bioethanol distilleries (LNF-Latino Americana). The S288c-isogenic strain used as a control in the fermentation assays in Fig. 1a was the BY4741 haploid (Giaever et al. 2002). For the initial genetic characterization of JAY270, the S288c-isogenic strains used were from the FY series (Winston et al. 1995). FY833 (*Mata*, *ura3-52*, *leu2Δ1*, *trp1Δ63*, *his3Δ200*, *lys2Δ202*) and FY834 (*Mata $\alpha$* , *ura3-52*, *leu2Δ1*, *trp1Δ63*, *his3Δ200*, *lys2Δ202*) were crossed to the JAY270-derived haploid spores JAY289 (*Mata*), JAY290 (*Mata $\alpha$* ), JAY291 (*Mata*), and JAY292 (*Mata $\alpha$* ) to generate hybrid diploids. FY23 (*Mata*, *ura3-52*, *leu2Δ1*, *trp1Δ63*) and FY86 (*Mata $\alpha$* , *ura3-52*, *leu2Δ1*, *his3Δ200*) were crossed to each other to generate the JAY309 diploid used as a control in comparative stress tolerance assays (Fig. 1b, and Supplemental Fig. 2 and 3). Finally, the S288c-isogenic strains used in the quantitative genetic analysis described in Fig. 2 were S1 and S97 (Steinmetz et al. 2002). Genomic DNA was also prepared from RM11-1a and YJM789 strains.

**Fermentation assays.** The fermentation assays shown in Fig. 1a were carried out by inoculating 4 g of fresh yeast cells in 36 ml of autoclaved sugar cane extract (pH4.0, 18.2% Total Reducing Sugars - mostly sucrose). Fermentation cultures were incubated at 30C without agitation for exactly 15 hours. The cells were recovered by centrifugation and the supernatant was distilled for the measurement of ethanol by densimetry. Cells were recycled for five consecutive fermentation cycles, and four independent replicates were analyzed for each genotype. In each replicate, only the data from cycles 2 to 5 were considered, since the data from the first cycle (adaptation round) is usually highly variable. The final ethanol concentration measurements from cycles 2 to 5 from all four replicates were combined to generate the average and standard error shown.

The fermentation assays used in the quantitative genetic analysis (Fig. 2) were carried out by inoculating 0.08 g of fresh yeast cells in 8 ml of media containing 0.5% yeast extract, 1% peptone, and 10% glucose. Cells were pre-grown in the same media for two cycles to provide an adaptation round. The cell mass concentration was normalized to 1% prior to the third cycle which was used to record cell growth and ethanol production kinetics. Cultures were incubated at 30C in a slow rotating drum to avoid decantation. Samples were taken at regular intervals and centrifuged in pre-weighted tubes to determine the weight of the cell pellets. The supernatant was frozen and later used to determine the ethanol concentration using the BioChain (Hayward, CA) Saccharide Removal and Ethanol Assay kits in sequence according to the manufacturer's recommendations.

**Microarray and physical analysis of DNA.** CGH-array and Band-array experiments, and PFGE procedures were conducted as previously described (Argueso et al. 2008). Image tracing of PFGE bands was obtained

using BioRad Quantity One software. All PCR reactions were conducted using BioRad iProof High-Fidelity DNA polymerase. For the differential gene expression experiments shown in Supplemental Table 1, total RNA was extracted from exponentially growing cells of JAY309 and JAY270. Three independent biological replicates were grown in YP 2% glucose liquid media, with 30C incubation and 200 RPM agitation. First strand cDNA was synthesized and labeled with dUTP-Cy3 or dUTP-Cy5, respectively, using the Invitrogen SuperScript Direct labeling kit. Labeled cDNAs were then competitively hybridized to microarrays containing all S288c ORFs. The Log<sub>2</sub> (Cy5/Cy3) signal for all three replicates and average for each microarray probe is shown. The CGH/Band-array, and the differential gene expression microarray data are available under GEO accession numbers GSE14601 and GSE17578, respectively.

**Parallel sequencing and genome assembly.** Total genomic DNA from JAY291 was isolated using Qiagen Maxi kit. We performed sequencing in the Illumina Genome Analyzer II and Roche Genome Sequencer FLX (454) platforms (Table 1), using standard manufacturer recommended sample preparation procedures. The Illumina and 454 reads were assembled into longer contigs using *de novo* assembler Edena (Hernandez et al. 2008) and Newbler (Roche), respectively. The hybrid assembly was performed through the combination of 3581 Illumina contigs and 1259 454 contigs resulting in 452 hybrid contigs (pipeline developed in our group; Zorro: A hybrid assembly pipeline for next-gen sequencing data; G. G. L. Costa, R. O. V., M. F. C., J. M. C. Mondego, M. J. Guiltinam, S. C. Schuster, J. E. Carlson, P. A. M., J. L. A., L. W. Meinhardt and G. A. G. P.; *manuscript in preparation*). Briefly, the zorro pipeline consists of (1) masking repeat regions in the contigs, (2) overlap detection, (3) unmasking repeat regions and (4) assembly of hybrid contigs. In phase (1), the repeat regions were determined based on counting occurrences of k-mers in the Roche 454 reads, the assembler then masks k-mers in the contigs that occur at high frequency. The absence of repeats produces correct overlap detection in phase (2). The contigs are then unmasked in phase (3), and the correct hybrid assembly can be obtained in (4), by merging all overlapping contigs into hybrid contigs. The overlaps detection and consensus generation were performed using the minimus package (Sommer et al. 2007). The hybrid contigs were ordered and oriented with Bambus program (Pop et al. 2004) using paired-end information and manual verification producing 427 scaffolds. These sequences were aligned to the reference genome (S288c) using Mummer package (Kurtz et al. 2004). The genome comparisons were visualized using Artemis Comparison Tool - ACT (Carver et al. 2005). *ab initio* and extrinsic gene predictions were obtained using Augustus (Stanke and Morgenstern 2005) and Exonerate (Slater and Birney 2005), respectively. For the Augustus analysis, we used the training sets available for S288c, and for Exonerate we used the reference curated ORFs (SGD; <http://www.yeastgenome.org>) that were aligned into our hybrid assembly.

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession ACFL00000000. The version described in this paper is the first version, ACFL01000000.

## **ACKNOWLEDGEMENTS**

We are grateful to Chris Vinci, Steven Clarke, Dina Raveh, Joseph Heitman, and Eric Alani for useful discussions and comments on the manuscript. We thank CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenho de São Paulo) for data processing support. Research in the Pereira, Tavares, Alcarde and Andrietta laboratories is supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), and CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior). Research in the Pereira laboratory is also supported by BRASKEM and ETH Bioenergia. Work in the Petes laboratory is funded by NIH grants GM52319 and GM24110.



## FIGURE LEGENDS

### Figure 1. Comparative phenotypic analysis of JAY270 and S288c.

**a.** Sugar cane extract fermentation assays (~18% sucrose). The ethanol concentration shown is the average reached in four fermentation assay repetitions, each comprised of four consecutive 15 hours of fermentation at 30C with cell recycling (see Materials and Methods for details). **b.** Reactive oxygen (menadione) and temperature effects on colony growth. Ten-fold serial dilutions of saturated cultures spotted (5  $\mu$ l) in rich medium (YPD) containing various concentrations of menadione were incubated at the indicated temperatures, and scanned after the indicated incubation period. The S288c-isogenic control strain used in **(a)** was the BY4741 haploid; and in **(b)** was the JAY309 diploid.

### Figure 2. Genetic analysis of cell mass accumulation and ethanol production kinetics.

**a** and **b.** Kinetics of cell mass accumulation and ethanol production, respectively, during fermentation of rich media with 10% glucose. The results show the average values and standard error for three biological replicates from each strain. The S288c-isogenic strains used in these assays were the *MAT $\alpha$*  haploids S1 ( $\rho^+$ ) and S97 ( $\rho^0$ ). The JAY361 diploid was obtained by mating JAY291 and S97; this diploid inherited 100% of its mitochondrial genome from JAY291. **c.** and **d.** show the distributions of cell mass accumulation after 18 hours and ethanol production after 8 hours (respectively) among haploid  $F_2$  spores derived from JAY361.  $F_2$  individuals were grouped in bins according to their phenotypes, and the bars represent the number of individuals in each phenotypic bin. **e.** Scatter plot of cell density (x axis) vs. ethanol concentration (y axis) for all  $F_2$  individuals tested in both assays (gray dots).  $P_1$ ,  $F_1$  and  $P_2$  indicate the relative phenotypes of JAY291, JAY361 and S1, respectively.  $n$  is the number of  $F_2$  individuals in each dataset,  $H^2$  is the broad-sense heritability calculated from phenotypic and environmental variances, and  $r^2$  is coefficient of correlation between the two traits analyzed. Only data from non-flocculant  $F_2$  individuals was used in this analysis.

### Figure 3. Molecular karyotype and gene copy number variation in JAY270.

**a.** PFGE and densitometric analysis of individual chromosomal bands in S288c and JAY270. The size of the peaks reflects the intensity of the ethidium bromide staining for each chromosomal band as determined by image analysis using BioRad QuantityOne software. The predicted chromosome sizes (in kb) are shown next to the corresponding chromosomal peaks were determined by comparison to the BioRad I molecular weight ladder (not shown). Note the presence of different sized homologs for Chr6 and Chr11 that appear at lower relative intensities. Contrast the abundance of each Chr6 homolog to the intensity of Chr3 for which both homologs are about the same size, and compare the abundance of each Chr11 homolog to the intensity of Chr10.

**b.** CGH-array relative gene dosage plots. Each horizontal line corresponds to a specific S288c chromosome; the signal of each array probe was smoothed in CGH-miner software in a seven-probe sliding window to reduce

noise (Wang et al. 2005). Gray areas indicate regions of similar genomic dosage, while positive/red and negative/green peaks indicate genomic regions over- or under-represented, respectively. The amplification signal on the left end of Chr1 included the *SEO1* gene encoding a putative amino-acid permease; the amplification peak on the right end of Chr3 did not include any known genes; all other amplifications are discussed in the text. None of the deletion/under-representation peaks spanned regions containing genes known to be essential in S288c. The S288c-isogenic control strain used in (a) and (b) was the JAY309 diploid.

**Figure 4. Molecular karyotype of meiotic products and segregation of chromosomal rearrangements.**

a. Ethidium bromide staining of a PFGE including the four meiotic spore clones from a JAY270 tetrad. The S288c-isogenic control strain used was the JAY309 diploid. b, c, and d show Southern hybridizations of the PFGE in (a), using as probes, respectively, the *AGP3*, *YAR064W*, and *SNO/SNZ* sequences that were PCR-amplified from S288c genomic DNA. The *SNO/SNZ* (2-3) probe detects only the duplicated *SNO2*, *SNZ2*, *SNO3*, and *SNZ3* genes. The diverged single copy genes *SNO1* and *SNZ1* on Chr13 are not detected. The numbers to the left indicate the S288c chromosomes to which these probes hybridized. The *YAR064W* gene is duplicated in S288c and JAY270 at the right end of Chr8.

**Figure 5. Band-array analysis of Chr6L and Chr6S homologs in JAY270.**

The curves indicate the normalized hybridization signal of specific chromosomal DNA samples (y axis) to probes arranged according to their chromosomal coordinates (x axis) in S288c Chr6 (a) and Chr1 (b). The dashed vertical lines indicate the breakpoints where JAY270 Chr6S and Chr6L differ from each other and from the S288c chromosomes. Only the data for Chr6 and Chr1 probes are shown. No hybridization signal was detected for other S288c chromosomes, with the exception of the right end of Chr8 that is essentially identical to Chr1. *SEC53* and *RPN12* indicate the most distal essential genes in S288c Chr6, and *SNO3*, *SNZ3*, *AGP3* and *YAR064W* indicate the position of Southern blot probes used in Fig. 4. The signal valley in the central position of Chr6 corresponds to the position of the *YFLWTy2-1* retrotransposable element (arrow) which is not present in the JAY270 Chr6 homologs. The black circles indicate the centromere positions.

**Figure 6. Genome rearrangements near the ends of chromosomes.**

a. Full-length Chr6 sequences aligned with Artemis Comparative Tool software (Carver et al. 2005). Red lines connect regions of sequence similarity higher than 85%, gaps in white indicate lower or absent similarity. Green indicates S288c Chr6 sequences and thick areas indicate regions conserved in JAY291, whereas thin areas indicate non-conserved regions. The small segment in Blue is a translocated fragment from S288c Chr10. Black indicates *S. cerevisiae* sequences not found in the S288c genome. The gray thick line corresponds to *S. paradoxus* Chr6 assembled from contigs 346, 345, 344, 434 and 433, from left to right, in this order (Kellis et al.

2003). *SEC53* and *RPN12* indicate the positions of the most distal essential genes in Chr6. Black circles indicate the position of the centromere (*CEN6*) and the arrow represents the *YFLWTy2-1* element. Y' and X sub-telomeric sequences are not represented in this alignment. **b.** Multi-strain chromosome alignment around the left end of Chr6. The horizontal lines represent the left ends of the designated chromosomes, with the exception of S288c Chr1 (top line) for which the inverted right end is shown. The source strain for each sequence is indicated to the left in bold, other strains with similar chromosome structures are also indicated. The rectangles represent ORFs, and their positions above and below the central line designate the Watson and Crick orientations, respectively. Chromosomal regions are color coded according to their correspondence to S288c Chr1 (orange), Chr6 (green) and Chr10 (blue), while regions in black correspond to sequences not found in the S288c genome. The dotted line in S288c Chr6 represents a discontinuity in the alignment at the site of a 19.3 kb insertion in JAY291. The rectangles hatched in red between chromosomes indicate a high level of sequence similarity (>85%; analogous to the red lines in panel a). All sequence similarities are indicated in red except for the left end of YJM789 Chr10 which contains the *SNO3* and *SNZ3* genes and is nearly identical to the collinear region in S288c Chr6. Boxed letters (A-F) indicate the specific chromosomal rearrangements discussed in the text, and *SNO3*, *SNZ3*, *AGP3* and *YAR064W* indicate the positions of Southern blot probes used in Fig. 4. The source sequences for this figure were the complete Chr6 sequences from S288c and JAY291, contigs 1.67 (Chr10) and 1.109 (Chr6) from RM11-1a, contigs 100 (Chr10) and 7 (Chr6) from YJM789, and contig 386 (Chr10) from JAY291. The 10 kb scale bar indicates the size scale for panel **b**.

**Figure 7. Phylogenetic placement and High-Temperature Growth (HTG) alleles in JAY291.**

**a.** Unrooted phylogenetic tree of 15 *S. cerevisiae* strains based on a 49 kb region from Chr14 containing the three HTG QTLs (Sinha et al. 2008). **b.** HTG QTL allele distribution in various *S. cerevisiae* strains. The three critical amino acid residues implicated in High Temperature Growth are shaded in gray, and HTG (+, -, nd - not determined) phenotype is shown to the right.

**Figure 8. Structural diversity in *S. cerevisiae*: rigid and plastic domains of the genome.**

The model depicts a set of homologs of a hypothetical chromosome in several unrelated *S. cerevisiae* strains. The top line (continuous gray) depicts the structural configuration of this chromosome in the first sequenced strain (*i.e.* S288c; reference), whereas the chromosomes shown below represent the rearrangements found in other strains. The diverged structural configurations (colors) in the peripheral regions harbor genes that are not required for viability, but that may contribute to fitness in specific environments. The entire set shares structural conservation in the central core region (delimited by the most distal essential genes - arrows), therefore meiotic crossovers between the unrelated haplotypes can generate new combinations, while remaining compatible with haploid viability.

Table 1. Summary of JAY291 genome sequencing and assembly parameters.

Haplotype	JAY291
Contigs	452
Contig size (average)	25.5 Kb
N50 <sup>a</sup>	65 Kb
Assembly size	11.6 Mb
Coverage depth <sup>b</sup> :	
454 single end	12x
Illumina single end	55x
Illumina paired-ends	95x
Gene models <sup>c</sup> :	
Extrinsic	5864
<i>Ab initio</i>	5488
Total gene models	5880

a. 50% of the genome assembly is represented by contigs of N50 size and longer.

b. Roche/454 reads were about 250 bp, Illumina single end reads were 35 bp, and paired-end reads correspond to two 35 bp reads physically linked by ~125 bp of undetermined sequence.

c. Extrinsic gene models were predicted based on the established S288c ORFs; *ab initio* gene models were predicted without regard to S288c data. 5471 genes were predicted by both models; 392 and 16 genes were found exclusively in the extrinsic and *ab initio* models, respectively.

## REFERENCES

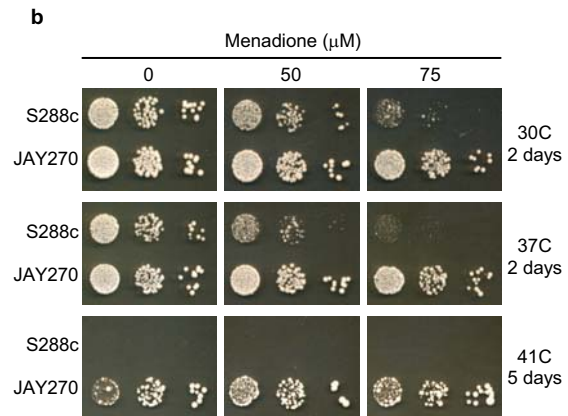
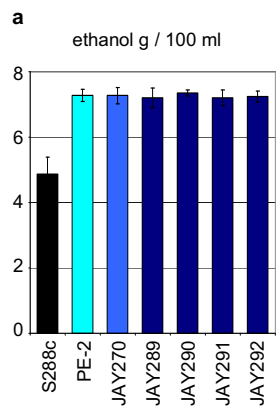
- Argueso, J.L., Westmoreland, J., Mieczkowski, P.A., Gawel, M., Petes, T.D., and Resnick, M.A. 2008. Double-strand breaks associated with repetitive DNA can reshape the genome. *Proc.Natl.Acad.Sci.U.S.A* **105**:11845-11850
- Barry, J.D., Ginger, M.L., Burton, P., and McCulloch, R. 2003. Why are parasite contingency genes often associated with telomeres? *Int.J.Parasitol.* **33**:29-45
- Baruffini, E., Lodi, T., Dallabona, C., and Foury, F. 2007. A single nucleotide polymorphism in the DNA polymerase gamma gene of *Saccharomyces cerevisiae* laboratory strains is responsible for increased mitochondrial DNA mutability. *Genetics* **177**:1227-1231
- Basso, L.C., de Amorim, H.V., de Oliveira, A.J., and Lopes, M.L. 2008. Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res.*
- Borneman, A.R., Forgan, A.H., Pretorius, I.S., and Chambers, P.J. 2008. Comparative genome analysis of a *Saccharomyces cerevisiae* wine strain. *FEMS Yeast Res.* **8**:1185-1195
- Carlson, M. and Botstein, D. 1983. Organization of the *SUC* gene family in *Saccharomyces*. *Mol.Cell Biol.* **3**:351-359
- Carreto, L., Eiriz, M.F., Gomes, A.C., Pereira, P.M., Schuller, D., and Santos, M.A. 2008. Comparative genomics of wild type yeast strains unveils important genome diversity. *BMC Genomics* **9**:524
- Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., and Parkhill, J. 2005. ACT: the Artemis Comparison Tool. *Bioinformatics.* **21**:3422-3423
- da Silva, E.A., dos Santos, S.K.B., Resende, A.D., de Morais, J.O.F., de Morais, M.A., and Simoes, D.A. 2005. Yeast population dynamics of industrial fuel-ethanol fermentation process assessed by PCR-fingerprinting. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **88**:13-23
- Doniger, S.W., Kim, H.S., Swain, D., Corcuera, D., Williams, M., Yang, S.P., and Fay, J.C. 2008. A catalog of neutral and deleterious polymorphism in yeast. *PLoS.Genet.* **4**:e1000183
- Du, X. and Takagi, H. 2007. N-Acetyltransferase Mpr1 confers ethanol tolerance on *Saccharomyces cerevisiae* by reducing reactive oxygen species. *Appl.Microbiol.Biotechnol.* **75**:1343-1351
- Ehrenshaft, M., Bilski, P., Li, M.Y., Chignell, C.F., and Daub, M.E. 1999. A highly conserved sequence is a novel gene involved in de novo vitamin B6 biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **96**:9374-9378
- Ekino, K., Kwon, I., Goto, M., Yoshino, S., and Furukawa, K. 1999. Functional analysis of *HO* gene in delayed homothallism in *Saccharomyces cerevisiae* wy2. *Yeast* **15**:451-458
- Faddah, D.A., Ganko, E.W., McCoach, C., Pickrell, J.K., Hanlon, S.E., Mann, F.G., Mieczkowska, J.O., Jones, C.D., Lieb, J.D., and Vision, T.J. 2009. Systematic identification of balanced transposition polymorphisms in *Saccharomyces cerevisiae*. *PLoS.Genet.* **5**:e1000502
- Fischer, G., James, S.A., Roberts, I.N., Oliver, S.G., and Louis, E.J. 2000. Chromosomal evolution in *Saccharomyces*. *Nature* **405**:451-454
- Foury, F. 1989. Cloning and sequencing of the nuclear gene *MIP1* encoding the catalytic subunit of the yeast mitochondrial DNA polymerase. *J.Biol.Chem.* **264**:20552-20560
- Fox, T.D., Folley, L.S., Mulero, J.J., McMullin, T.W., Thorsness, P.E., Hedin, L.O., and Costanzo, M.C. 1991. Analysis and manipulation of yeast mitochondrial genes. *Methods in Enzymology* **194**:149-165

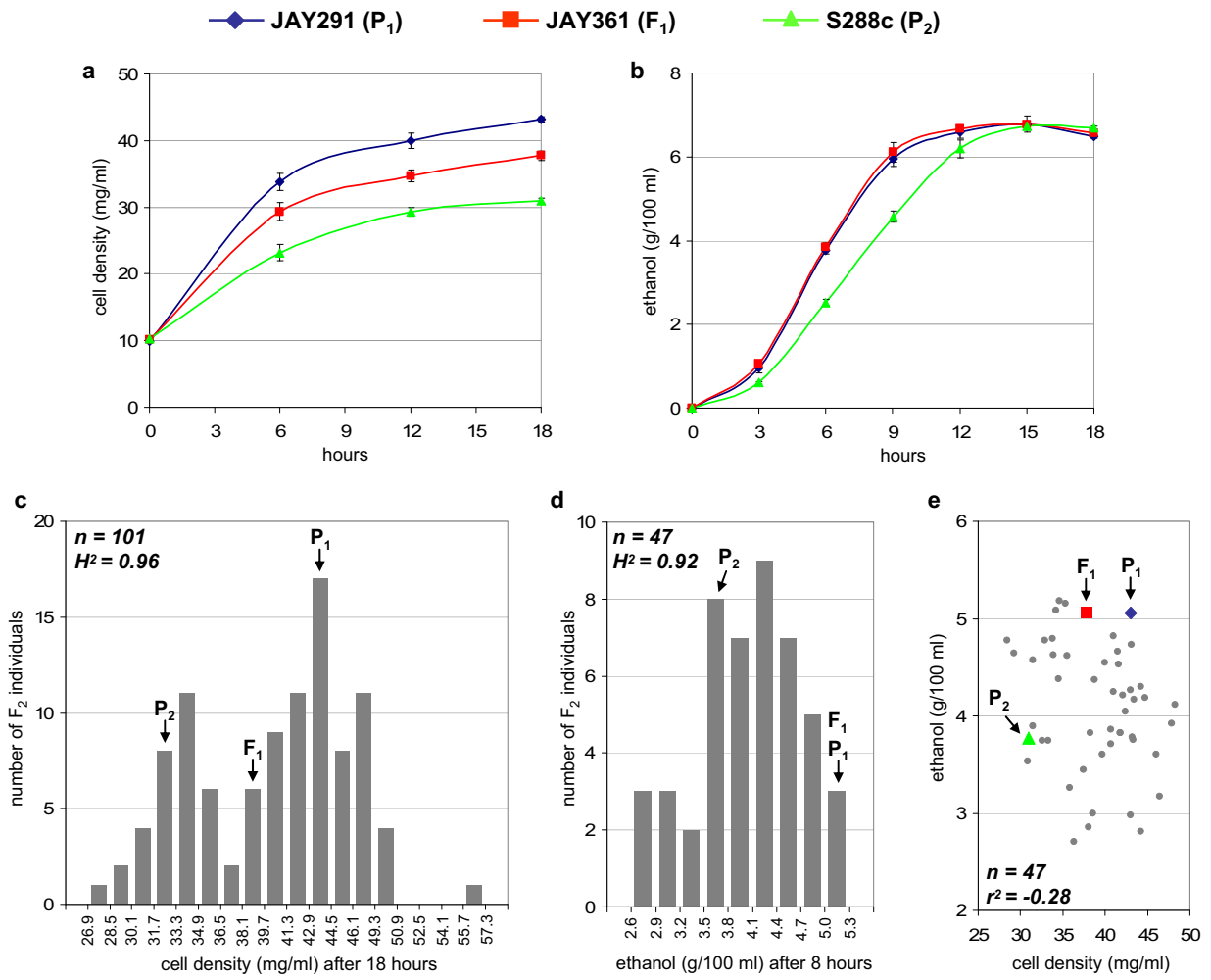
- Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. 2000. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* **407**:1018-1022
- Gaisne, M., Becam, A.M., Verdiere, J., and Herbert, C.J. 1999. A 'natural' mutation in *Saccharomyces cerevisiae* strains derived from S288c affects the complex regulatory gene *HAP1* (*CYP1*). *Curr.Genet.* **36**:195-200
- Gasser, S.M., Hediger, F., Taddei, A., Neumann, F.R., and Gartenberg, M.R. 2004. The function of telomere clustering in yeast: the circe effect. *Cold Spring Harb.Symp.Quant.Biol.* **69**:327-337
- Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**:387-391
- Goffeau, A. 2004. Evolutionary genomics: seeing double. *Nature* **430**:25-26
- Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., et al. 1996. Life with 6000 genes. *Science* **274**:546, 563-546, 567
- Goldemberg, J. 2007. Ethanol for a sustainable energy future. *Science* **315**:808-810
- Gordon, J.L., Byrne, K.P., and Wolfe, K.H. 2009. Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. *PLoS.Genet.* **5**:e1000485
- Gura, T. 2009. Driving Biofuels from Field to Fuel Tank. *Cell* **138**:9-12
- Hall, C. and Dietrich, F.S. 2007. The reacquisition of biotin prototrophy in *Saccharomyces cerevisiae* involved horizontal gene transfer, gene duplication and gene clustering. *Genetics* **177**:2293-2307
- Hernandez, D., Francois, P., Farinelli, L., Osteras, M., and Schrenzel, J. 2008. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res.* **18**:802-809
- Herskowitz, I. 1988. Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiol.Rev.* **52**:536-553
- Inoue, T., Iefuji, H., Fujii, T., Soga, H., and Satoh, K. 2000. Cloning and characterization of a gene complementing the mutation of an ethanol-sensitive mutant of sake yeast. *Bioscience Biotechnology and Biochemistry* **64**:229-236
- Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E.S. 2003. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**:241-254
- Knop, M. 2006. Evolution of the hemiascomycete yeasts: on life styles and the importance of inbreeding. *Bioessays* **28**:696-708
- Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* **5**:R12
- Landolfo, S., Politi, H., Angeozzi, D., and Mannazzu, I. 2008. ROS accumulation and oxidative damage to cell structures in *Saccharomyces cerevisiae* wine strains during fermentation of high-sugar-containing medium. *Biochimica et Biophysica Acta-General Subjects* **1780**:892-898
- Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A., Davey, R.P., Roberts, I.N., Burt, A., Koufopanou, V., et al. 2009. Population genomics of domestic and wild yeasts. *Nature* **458**:337-341
- Liu, H.P., Styles, C.A., and Fink, G.R. 1996. *Saccharomyces cerevisiae* S288C has a mutation in *FL08* a gene required for filamentous growth. *Genetics* **144**:967-978

- Llorente, B., Smith, C.E., and Symington, L.S. 2008. Break-induced replication: what is it and what is it for? *Cell Cycle* **7**:859-864
- Meiron, H., Nahon, E., and Raveh, D. 1995. Identification of the heterothallic mutation in HO-endonuclease of *S. cerevisiae* using HO/ho chimeric genes. *Curr.Genet.* **28**:367-373
- Mense, S.M. and Zhang, L. 2006. Heme: a versatile signaling molecule controlling the activities of diverse regulators ranging from transcription factors to MAP kinases. *Cell Res.* **16**:681-692
- Ness, F. and Aigle, M. 1995. *RTM1*: a member of a new family of telomeric repeated genes in yeast. *Genetics* **140**:945-956
- Padilla, P.A., Fuge, E.K., Crawford, M.E., Errett, A., and Werner-Washburne, M. 1998. The highly conserved, coregulated *SNO* and *SNZ* gene families in *Saccharomyces cerevisiae* respond to nutrient limitation. *Journal of Bacteriology* **180**:5718-5726
- Palladino, F., Laroche, T., Gilson, E., Axelrod, A., Pillus, L., and Gasser, S.M. 1993. SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. *Cell* **75**:543-555
- Paques, F. and Haber, J.E. 1999. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol.Mol.Biol.Rev.* **63**:349-404
- Polakova, S., Blume, C., Zarate, J.A., Mentel, M., Jorck-Ramberg, D., Stenderup, J., and Piskur, J. 2009. Formation of new chromosomes as a virulence mechanism in yeast *Candida glabrata*. *Proc.Natl.Acad.Sci.U.S.A* **106**:2688-2693
- Pop, M., Kosack, D.S., and Salzberg, S.L. 2004. Hierarchical scaffolding with Bambus. *Genome Res.* **14**:149-159
- Pryde, F.E., Gorham, H.C., and Louis, E.J. 1997. Chromosome ends: all the same under their caps. *Curr.Opin.Genet.Dev.* **7**:822-828
- Robertson, G.P., Dale, V.H., Doering, O.C., Hamburg, S.P., Melillo, J.M., Wander, M.M., Parton, W.J., Adler, P.R., Barney, J.N., Cruse, R.M., et al. 2008. Agriculture. Sustainable biofuels redux. *Science* **322**:49-50
- Rodriguez-Navarro, S., Llorente, B., Rodriguez-Manzanique, M.T., Ramne, A., Uber, G., Marchesan, D., Dujon, B., Herrero, E., Sunnerhagen, P., and Perez-Ortin, J. 2002. Functional analysis of yeast gene families involved in metabolism of vitamins B-1 and B-6. *Yeast* **19**:1261-1276
- Rose, M.D., Winston, F., and Hieter, P. 1990. *Methods in yeast genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanderson, K. 2006. US biofuels: a field in ferment. *Nature* **444**:673-676
- Scannell, D.R., Butler, G., and Wolfe, K.H. 2007. Yeast genome evolution--the origin of the species. *Yeast* **24**:929-942
- Schacherer, J., Shapiro, J.A., Ruderfer, D.M., and Kruglyak, L. 2009. Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* **458**:342-345
- Sinha, H., David, L., Pascon, R.C., Clauder-Munster, S., Krishnakumar, S., Nguyen, M., Shi, G., Dean, J., Davis, R.W., Oefner, P.J., et al. 2008. Sequential elimination of major-effect contributors identifies additional quantitative trait Loci conditioning high-temperature growth in yeast. *Genetics* **180**:1661-1670
- Sinha, H., Nicholson, B.P., Steinmetz, L.M., and McCusker, J.H. 2006. Complex genetic interactions in a quantitative trait locus. *PLoS.Genet.* **2**:e13
- Slater, G.S. and Birney, E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics.* **6**:31

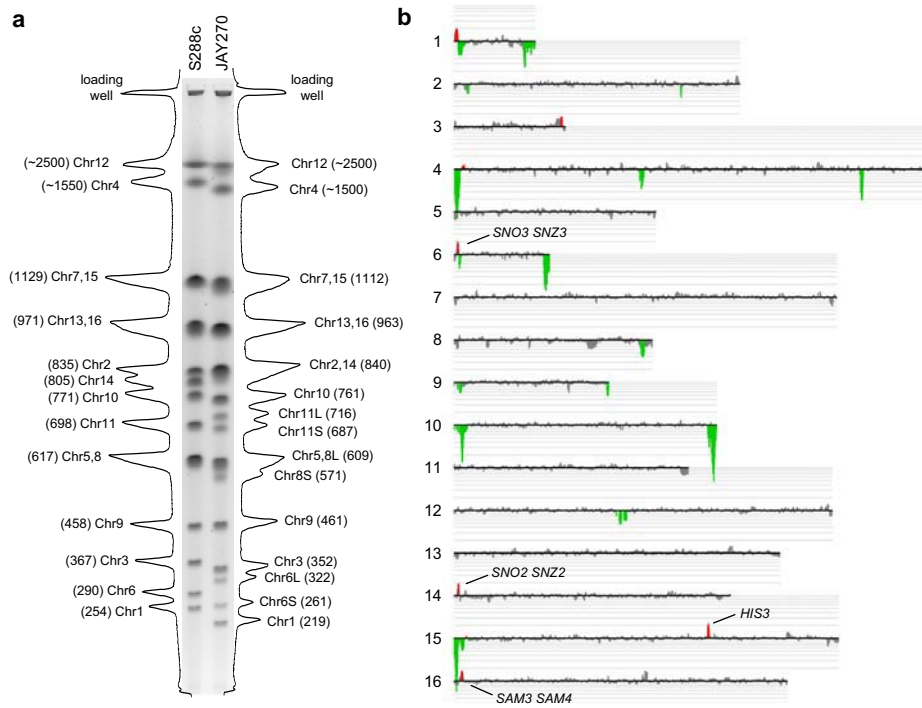
- Sommer, D.D., Delcher, A.L., Salzberg, S.L., and Pop, M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics*. **8**:64
- Stanke, M. and Morgenstern, B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* **33**:W465-W467
- Steinmetz, L.M., Sinha, H., Richards, D.R., Spiegelman, J.I., Oefner, P.J., McCusker, J.H., and Davis, R.W. 2002. Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**:326-330
- Takagi, H., Shichiri, M., Takemura, M., Mohri, M., and Nakamori, S. 2000. *Saccharomyces cerevisiae* sigma 1278b has novel genes of the N-acetyltransferase gene superfamily required for L-proline analogue resistance. *J.Bacteriol.* **182**:4249-4256
- Tamura, K., Gu, Y., Wang, Q., Yamada, T., Ito, K., and Shimoi, H. 2004. A *hap1* mutation in a laboratory strain of *Saccharomyces cerevisiae* results in decreased expression of ergosterol-related genes and cellular ergosterol content compared to sake yeast. *J.Biosci.Bioeng.* **98**:159-166
- Tavares, F.C.A. and Echeverrigaray, S. 1987. Yeast Breeding for fuel production. In Biochemistry and Biology of Industrial Yeasts (eds. G. Stewart, I. Russell, R. Klein, R. Hiebish), pp. 59-80. CRC, Florida.
- Verstrepen, K.J., Reynolds, T.B., and Fink, G.R. 2004. Origins of variation in the fungal cell surface. *Nat.Rev.Microbiol.* **2**:533-540
- Vinci, C.R. and Clarke, S.G. 2007. Recognition of age-damaged (R,S)-adenosyl-L-methionine by two methyltransferases in the yeast *Saccharomyces cerevisiae*. *J.Biol.Chem.* **282**:8604-8612
- Wang, P., Kim, Y., Pollack, J., Narasimhan, B., and Tibshirani, R. 2005. A method for calling gains and losses in array CGH data. *Biostatistics.* **6**:45-58
- Wei, W., McCusker, J.H., Hyman, R.W., Jones, T., Ning, Y., Cao, Z., Gu, Z., Bruno, D., Miranda, M., Nguyen, M., et al. 2007. Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *Proc.Natl.Acad.Sci.U.S.A* **104**:12825-12830
- Wheals, A.E., Basso, L.C., Alves, D.M., and Amorim, H.V. 1999. Fuel ethanol after 25 years. *Trends Biotechnol.* **17**:482-487
- Winston, F., Dollard, C., and RicuperoHovasse, S.L. 1995. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. *Yeast* **11**:53-55
- Winzeler, E.A., Castillo-Davis, C.I., Oshiro, G., Liang, D., Richards, D.R., Zhou, Y., and Hartl, D.L. 2003. Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays. *Genetics* **163**:79-89



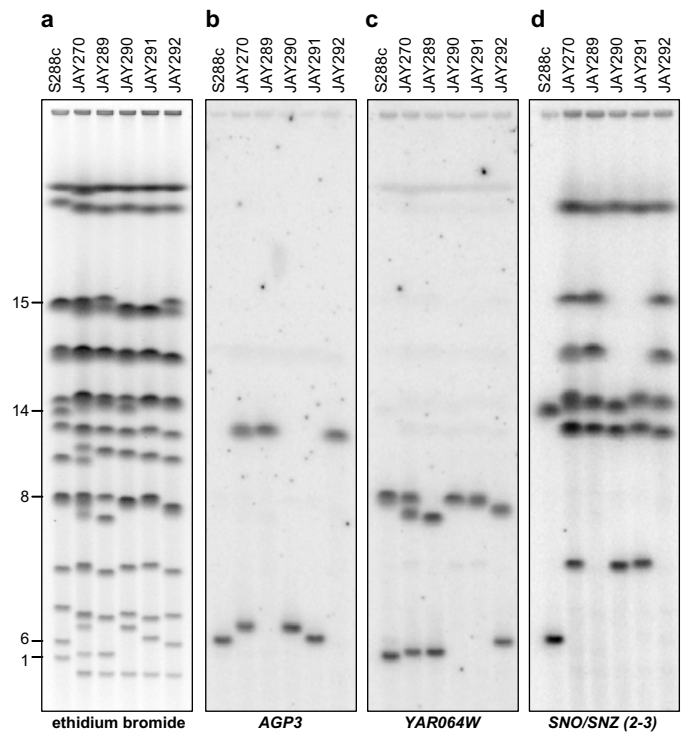


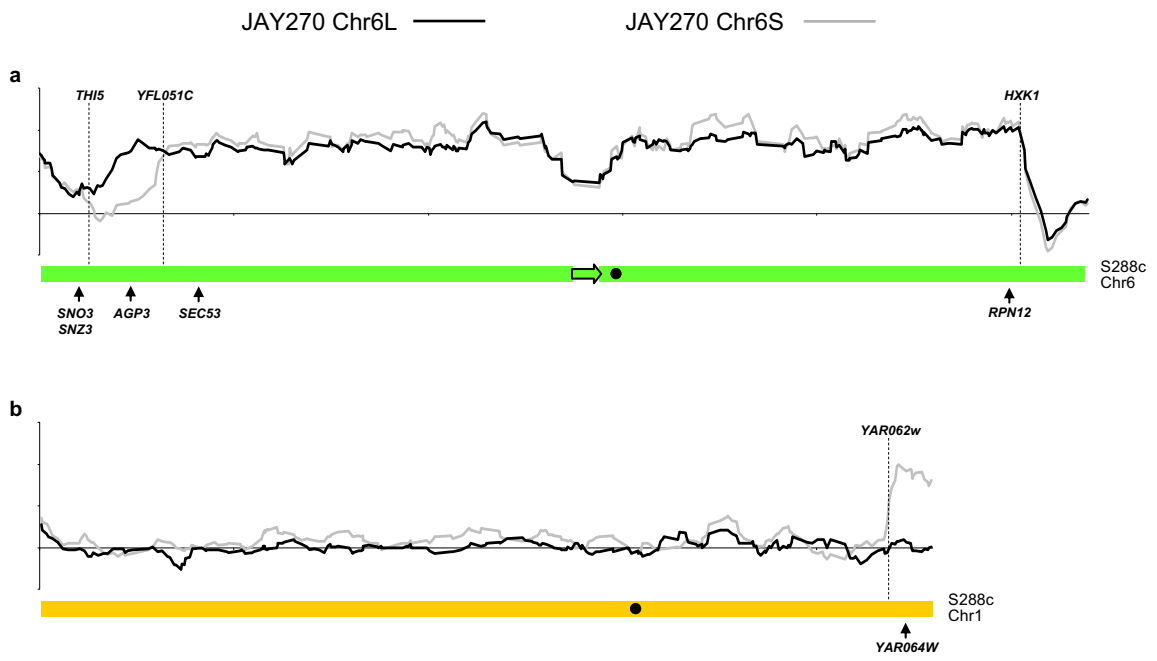


Argueso *et al* Figure 2

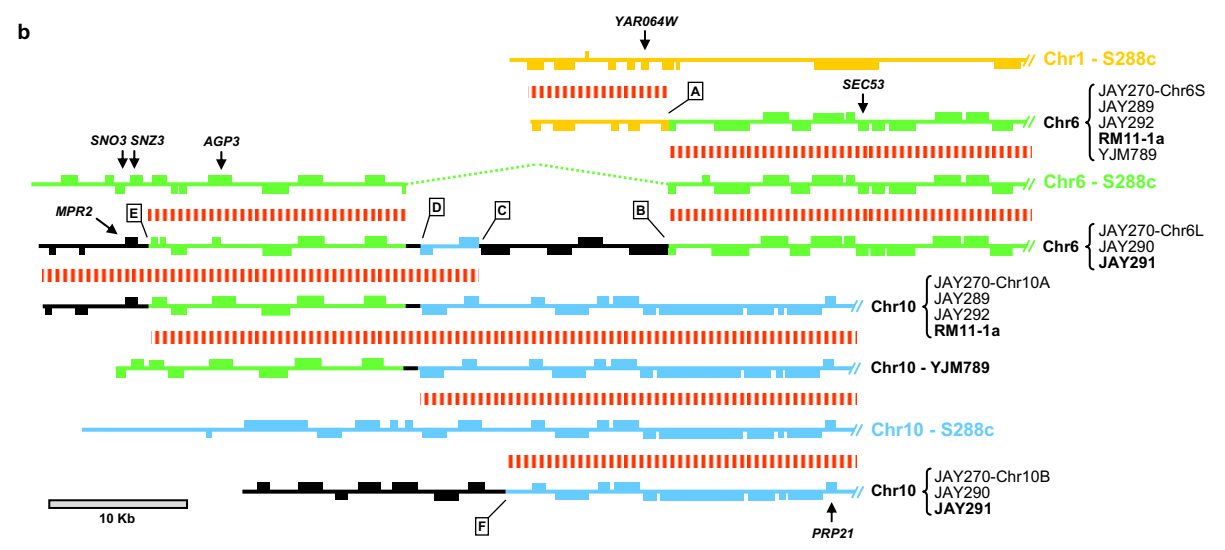
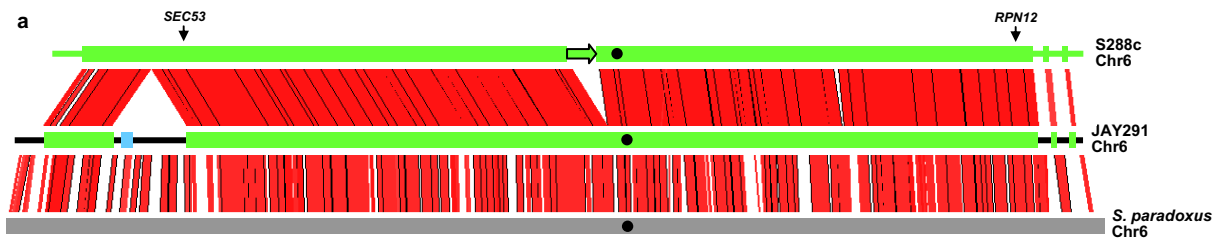


Argueso *et al* Figure 3

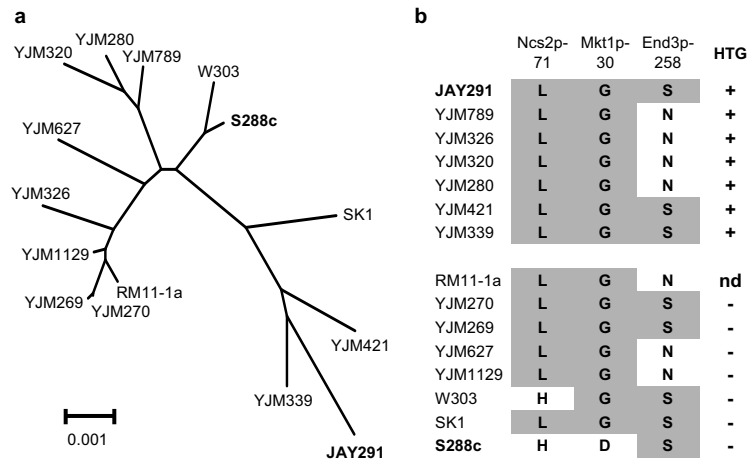


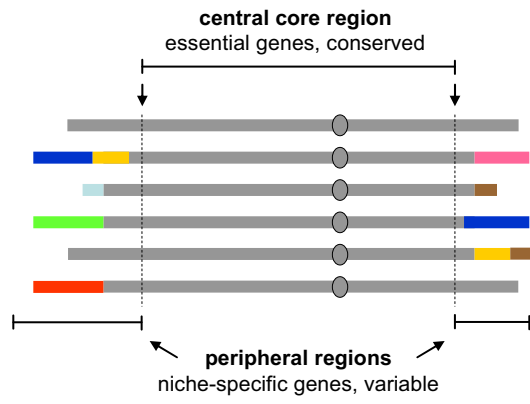


Argueso *et al* Figure 5

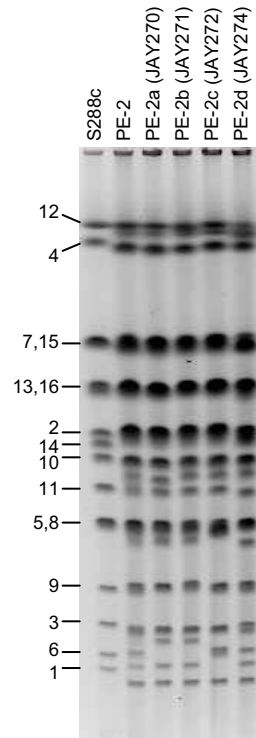


Argueso *et al* Figure 6



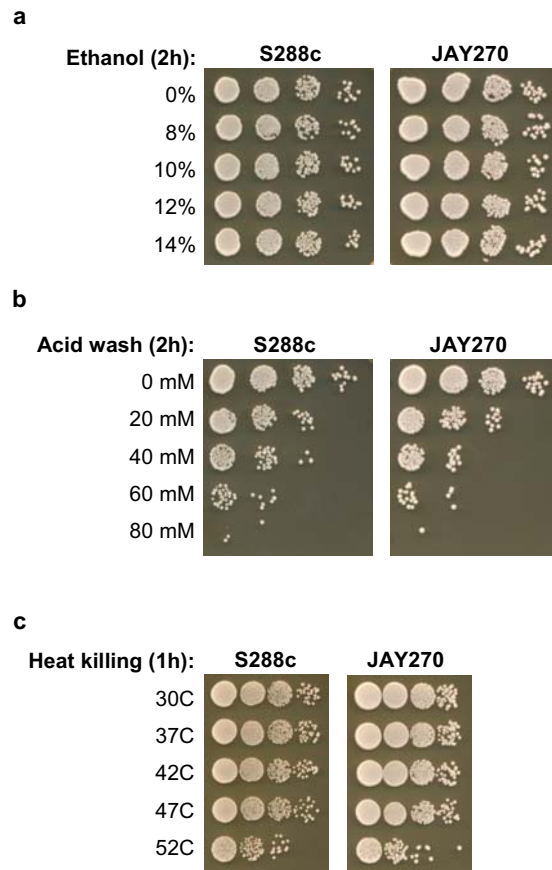






**Supplemental Figure 1.** Molecular karyotype analysis of PE-2 and its derived pure strains.

**a.** PFGE analysis of the original PE-2 mixed cultured and of four independent clonal single colony isolates derived from PE-2. Note the diffuse chromosomal banding pattern in PE-2 which was caused by the presence of subpopulations of cells with slightly divergent karyotypes. In contrast, the bands observed in the purified isolates (JAY270 to JAY274) are discrete and sharper as expected for clonal cell lineages, each carrying specific chromosomal variants.

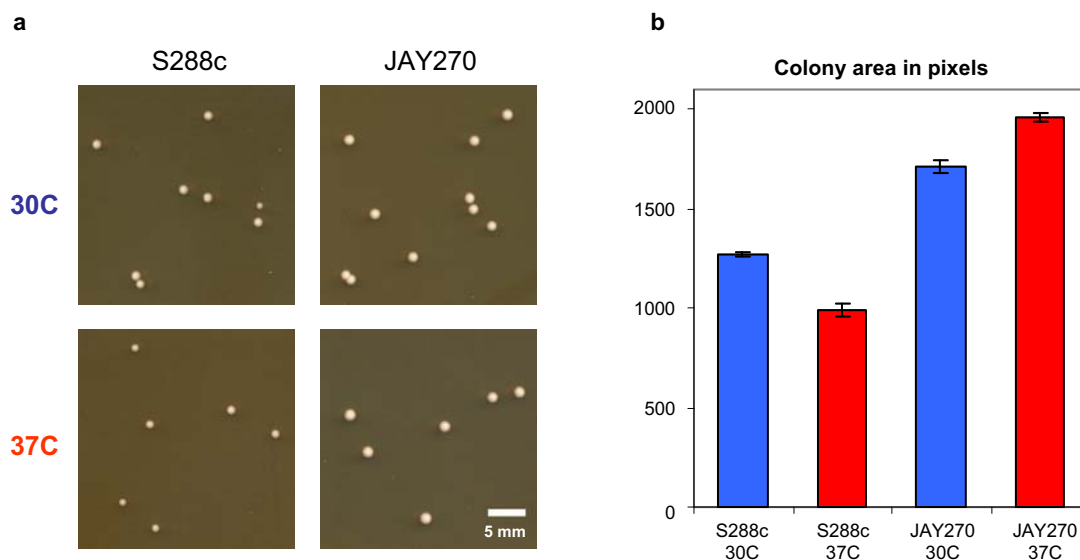


**Supplemental Figure 2.** Effect of incubation on various stressors on cell viability.

10<sup>-1</sup> dilutions of saturated overnight cultures of an S288c-derived diploid (JAY309) and JAY270 were incubated in the presence of various concentrations of ethanol (**a**), sulfuric acid (**b**), or at different temperatures (**c**). After the indicated time periods, these aliquots were serially diluted in sterile water in 10<sup>-1</sup> fold increments, spotted to rich medium and allowed to grow at 30C for 48 hours to assess cell viability.

No significant differences in viability were observed between the two strains for ethanol tolerance up to 14% concentration (**a**). Under standard sugarcane extract fermentation the final ethanol concentration rarely exceeds 12%.

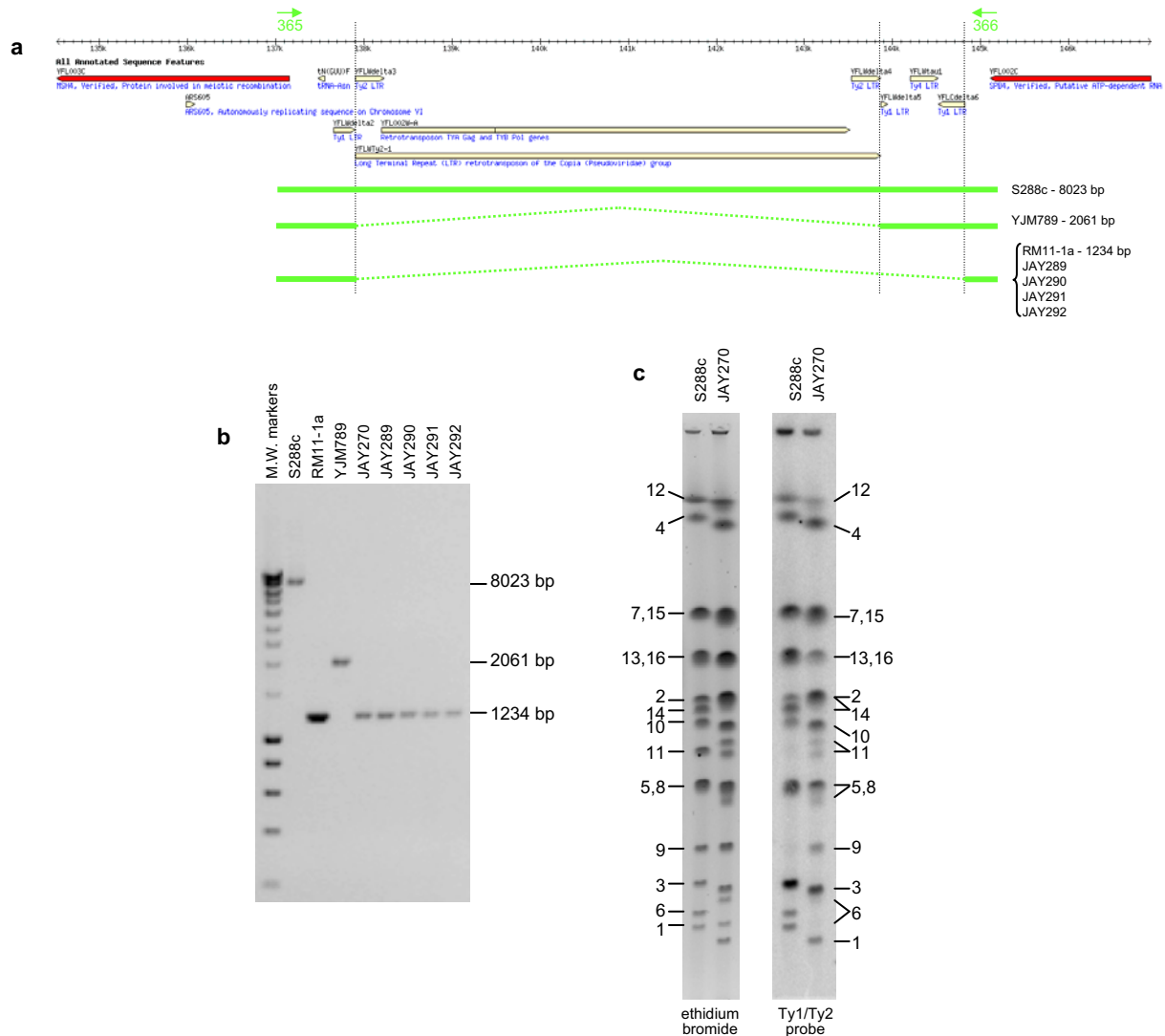
Also, no significant differences were observed in tolerance to short-term acid wash (**b**), or to short-term heat killing (**c**) at high temperatures beyond those compatible with cell growth (42C, 47C and 52C).



**Supplemental Figure 3.** Effect of incubation temperature on colony size.

**a.** Cropped sample Petri dish scans showing difference in colony size between S288c and JAY270 grown in rich medium (YEPD) for 48 hours at 30C and 37C. Scale bar is 5x1 mm.

**b.** Quantification of colony sizes. Cells were plated to a density of about 20 colonies per plate. After 48 hours of incubation, Petri dishes were scanned at 800 dpi and images were analyzed in Adobe Photoshop CS2 to measure the colonies. The average colony area in pixels with standard error are shown. In the images above, 1 mm<sup>2</sup> equals 986 pixels. In the specific experiment shown, the diameter of 13 well-spaced colonies was determined. Colonies were then sorted by size and two smallest and the two largest were ignored for the average area calculation to minimize the influence of possible *petite* colonies in the plate. The same trend was observed in another two independent experimental repetitions.

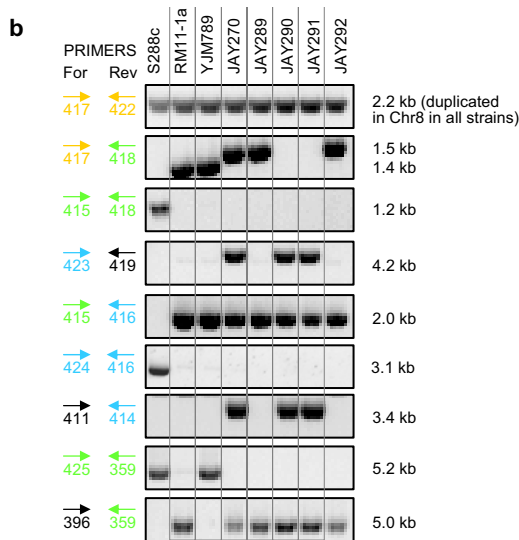
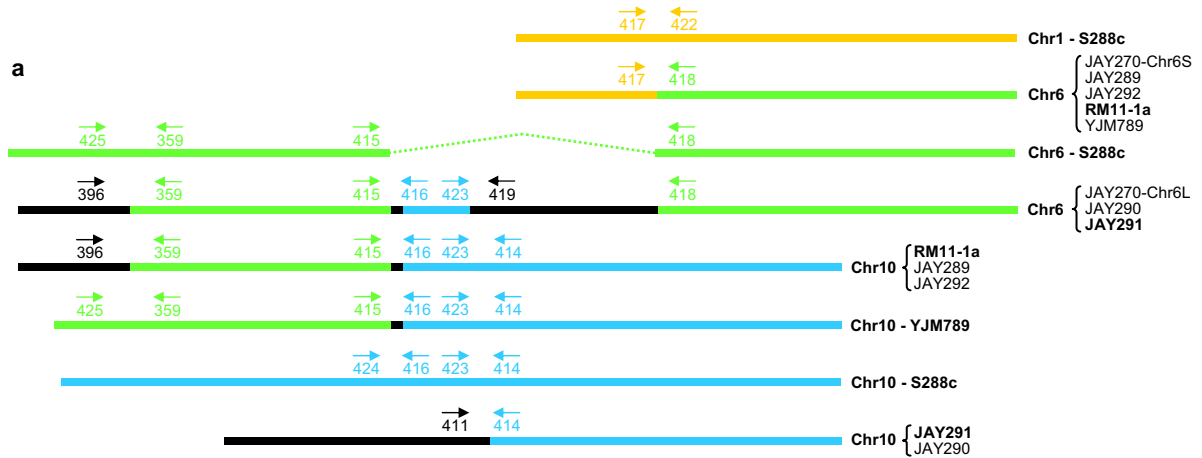


**Supplemental Figure 4.** PCR mapping of the *YFLWty2-1* region and Ty distribution in JAY270.

a. Screen capture of GBrowser software showing the *YFLWty2-1* region in Chr6 of S288c, including the *MSH4* and *SPB4* flanking genes. The forward PCR primer JAO365 is positioned at the 5' end of *MSH4*, and the reverse primer JAO366 is positioned at the 3' end of *SPB4*. The schematics below the map show the expected PCR products from this region in various *S. cerevisiae* strains. S288c contains multiple Ty element insertions upstream of the tN(GUU)f tRNA gene, including the full length *YFLWty2-1* element. YJM789 has the same LTRs as S288c, but lacks the Ty2 element. Finally, RM11-1a and the JAY270 ascospores contain only the YFLWdelta2 LTR, which likely marks the founding Ty insertion at this site.

b. Ethidium bromide stained agarose gel showing the JAO365 / JAO366 PCR products from (a).

c. PFGE and Southern analysis of Ty1 and Ty2 retrotransposons in JAY270. S288c chromosomes are indicated to the left and JAY270 chromosomes to the right. The probe used hybridizes to both Ty1 and Ty2 elements.



**c**

Primer	Nucleotide sequence (5'→3')
JAO359	TCGCTGTCTCTTCAATGTCGACAG
JAO365	AAACTAGATAGATTAGATTCACCTC
JAO366	GTTTCCAGCAAAGCTATCCAAGGC
JAO396	AGTTTATATTGTAGTGGCATGACTC
JAO411	CTTACTGATAGCGATGCCCGAGAC
JAO414	ACAGCTGTCAAAGGCACTAGCAGG
JAO415	TGTCAGTGTGCGCTTATCACG
JAO416	CAGTTCATTGATCTTGATAATGAC
JAO417	CATTGCTGCGATAGAATGGAG
JAO418	TGCAATTGATGCCTACCGGAG
JAO419	CCCACGACGTGCTTACAACGCACT
JAO422	AACTCAAGGACGAGGAGTATTCGG
JAO423	GTCATTATCAAGATCAATGAACTG
JAO424	AGAGTTCAGTTGTTGCGGACCGAG
JAO425	ACATCACCAAGATTAAGGCGGAG

**Supplemental Figure 5.** PCR mapping of chromosomal rearrangements at the left arm of Chr6 and Chr10.

**a.** Map of PCR primer locations in the region examined. This map matches the genome rearrangement breakpoints in Fig. 6b, but it is not drawn to scale. The chromosomal regions and primers are color coded according to their correspondence to S288c's Chr1 (orange), Chr6 (green) and Chr10 (blue), while regions in black correspond to sequences not found in the S288c genome. Arrows indicate primer forward or reverse orientations.

**b.** Cropped ethidium bromide stained agarose gels with PCR reactions using as template genomic DNA from the indicated strains. The primer pair used in each reaction is shown to the left, and the size of the expected target product is indicated to the right. Only the region of each gel containing the target PCR product is shown. The expected target band was observed in all cases. Occasional unspecific PCR products were also present in a few of the reactions, but were cropped out to facilitate visualization of the results.

**c.** PCR primers used in mapping of chromosomal rearrangements at the left arm of Chr6 and Chr10, and of the *YFLWY2-1* insertion (Supplemental Fig. 4).

## DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação/tese de Doutorado intitulada "Modificação de linhagens industriais de *Saccharomyces cerevisiae* para o aumento da produtividade de álcool e floculação condicional":

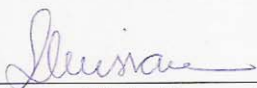
( ) não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

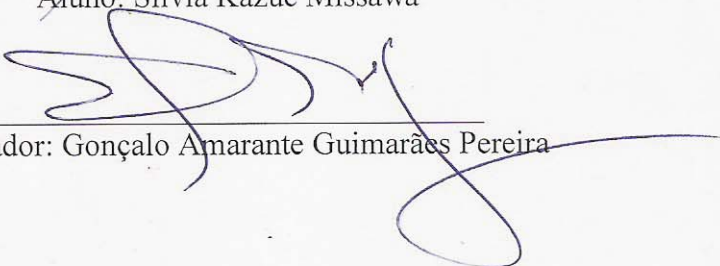
(X) está inserido no **Projeto CIBio/IB/UNICAMP** (Protocolo nº 2004/02), intitulado "Modificação de linhagens industriais de *Saccharomyces cerevisiae* para o aumento da produtividade de álcool e floculação condicional";

( ) tem autorização da **Comissão de Ética em Experimentação Animal/IB/UNICAMP** (Protocolo nº \_\_\_\_\_);

( ) tem autorização do **Comitê de Ética para Pesquisa com Seres Humanos/FCM/UNICAMP** (Protocolo nº \_\_\_\_\_);

( ) tem autorização de comissão de bioética ou biossegurança externa à UNICAMP.  
Especificar: \_\_\_\_\_


  
\_\_\_\_\_  
Aluno: Silvia Kazue Missawa

  
\_\_\_\_\_  
Orientador: Gonçalo Amarante Guimarães Pereira

Para uso da Comissão ou Comitê pertinente:

(X) Deferido ( ) Indeferido

Nome: \_\_\_\_\_

Função:   
Prof. Dra. HELENA COUTINHO F. DE OLIVEIRA  
Presidente  
Comissão Interna de Biossegurança  
CIBio/IB - UNICAMP