

Nádia Maria Vieira Sampaio

“Caracterização fenotípica e fisiológica de cepas de *Saccharomyces cerevisiae* mutadas para genes potencialmente envolvidos em fermentação sob condição industrial”

**CAMPINAS
2013**



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA

NÁDIA MARIA VIEIRA SAMPAIO

**“CARACTERIZAÇÃO FENOTÍPICA E FISIOLÓGICA DE
CEPAS DE *SACCHAROMYCES CEREVISIAE* MUTADAS PARA
GENES POTENCIALMENTE ENVOLVIDOS EM
FERMENTAÇÃO SOB CONDIÇÃO INDUSTRIAL”**

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)
NÁDIA MARIA VIEIRA SAMPAIO
e aprovada pela Comissão Julgadora.

Dissertação apresentada ao Instituto de
Biologia para obtenção do Título de
Mestre em Genética e Biologia
Molecular, na área de Genética de
Microrganismos.

Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira
Co-orientador: Prof. Dr. Juan Lucas Argueso Gomes de Almeida

CAMPINAS,
2013

Sampaio, Nadia Maria Vieira, 1989-
Sa47c Caracterização fenotípica e fisiológica de cepas de *Saccharomyces cerevisiae* mutadas para genes potencialmente envolvidos em fermentação sob condição industrial / Nádia Maria Vieira Sampaio. – Campinas, SP : [s.n.], 2013.

Orientador: Gonçalo Amarante Guimarães Pereira.
Coorientador: Juan Lucas Argueso Gomes de Almeida.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Biocombustíveis. 2. Etanol. 3. *Saccharomyces cerevisiae*. 4. Linhagem industrial. 5. Genômica funcional. I. Pereira, Gonçalo Amarante Guimarães, 1964-. II. Argueso, Juan Lucas, 1972-. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

Título em inglês: Phenotypic and physiological characterization of mutant strains of *Saccharomyces cerevisiae* for genes potentially involved in industrial fermentation

Palavras-chave em inglês:

Biofuels

Ethanol

Saccharomyces cerevisiae

Industrial strain

Functional genomics

Área de concentração: Genética de Microorganismos

Titulação: Mestra em Genética e Biologia Molecular

Banca examinadora:

Gonçalo Amarante Guimarães Pereira [Orientador]

Anderson Ferreira Cunha

Sandra Regina Ceccato Antonini

Data de defesa: 02-07-2013

Programa de Pós-Graduação: Genética e Biologia Molecular

Campinas, 02 de julho de 2013

BANCA EXAMINADORA

Prof. Dr. Gonçalo Amarante Guimarães Pereira
(orientador)



Assinatura

Prof. Dr. Anderson Ferreira Cunha



Assinatura

Profa. Dra. Sandra Regina Ceccato Antonini



Assinatura

Prof. Dr. Paulo Arruda

Assinatura

Prof. Dr. Flávio Henrique da Silva

Assinatura

**Dedico este trabalho a meus pais, Sérgio e Nilza,
por todo amor e apoio na realização desta conquista.**

AGRADECIMENTOS

Inicialmente, gostaria de agradecer meu orientador Prof. Dr. Gonçalo Amarante Guimarães Pereira, pela confiança em mim depositada desde o início, pela orientação científica e pessoal e pela oportunidade única que me concedeu, permitindo a realização deste trabalho. Obrigada por me contagiar com seu entusiasmo pela Ciência e, assim me motivar a explorar este campo a cada dia. O seu papel foi essencial para meu desenvolvimento acadêmico.

Com especial carinho, agradeço também a meu co-orientador, Prof. Dr. Lucas Argueso, por toda a atenção e valiosos ensinamentos dados durante estes dois anos. Você é para mim um exemplo de cientista, professor e orientador.

A Marcelo Carnier Dornelas, por todos os ensinamentos transmitidos durante minha iniciação científica e suporte ao longo de minha trajetória na Unicamp.

Aos membros de minha qualificação, pré-banca e banca. Prof. Dr. Luiz Humberto Gomes, Dr. Ana Carolina Deckman, Dr. Jorge Maurício Costa Mondego, Prof. Dr. Fábio Papes, Profa. Dra. Sandra Antonini e Prof. Anderson Ferreira Cunha, agradeço sinceramente por todas as sugestões e contribuições para este trabalho.

A todos os membros do Laboratório de Genômica e Expressão (LGE) pela amizade e apoio na realização deste trabalho. Especialmente, aos meus colegas Angélica M. Gomes, Beatriz Temer, Bruno Vaz de Oliveira, Daniela Thomazella, Gleidson S. Teixeira, Leandro V. dos Santos, Luciana S. Mofatto, Marcela Salazar, Marcelo F. Carazzolle, Odalys G. Cabrera, Pedro A. G. Tizei, Ricardo A. Tiburcio e Silvia K. Missawa, por toda ajuda na condução e delineamento dos experimentos e por todas as sugestões dadas.

A Osmar Vaz de Carvalho Netto, por todo o suporte experimental, pela paciência e discussões acerca de meu projeto. Sua ajuda foi fundamental para a conclusão deste trabalho.

A Luige Armando Llerena Calderon, pela amizade e apoio constante ao desenvolvimento deste trabalho. Obrigada por compartilhar comigo seu conhecimento sobre fermentação, por suas valiosas sugestões e por tornar meus dias de trabalho mais divertidos.

À Eliane Laranja Dias e Silvia Kazue Missawa por todo o apoio, paciência e por tornarem minha vida muito mais fácil ao longo deste trabalho.

Ao Instituto de Biologia da UNICAMP e à *Colorado State University* pela estrutura disponibilizada.

Ao CNPQ e FAPESP (processos 2010/14131-1 e 2012/09447-5) pelo apoio financeiro.

Aos meus queridos amigos, Laura M. Branco, Gabriel F. Zaniboni, Caio G. Pereira, Guilherme Barbosa, Suzana U. Benitez, Raísa M. Lorencini, Thaís H. Ferreira, Carolina C. Monte Bello, Gabriela Chiarotto e Gabriela Bonadirman, por serem a minha família em Campinas.

A Victor Augusti Negri, pela amizade e suporte na realização deste projeto. Suas sugestões, discussões e ajuda na condução dos experimentos foram fundamentais para conclusão deste trabalho. Obrigada por dividir comigo sua visão encantadora sobre Ciência e, com isso, me motivar a seguir em frente nesta carreira que escolhemos. Obrigada por todos os conselhos, risadas e por nunca, NUNCA ter me deixado sentir sozinha, mesmo que isso significasse a trabalhar a noite toda. Vou sentir muita saudade de ter você por perto.

A Marcelo Colika Bassalo, por ser um de meus maiores apoios durante todo o mestrado. Obrigada por dividir comigo todo seu conhecimento científico, por suas sugestões e pelos ensinamentos teóricos e práticos acerca de meu trabalho. Principalmente, obrigada pela amizade, pelas risadas, por me acalmar com suas opiniões (ou com sucos de maracujá) e por me fazer compreender o valor de uma amizade, em todos os seus sentidos. Sua ajuda foi fundamental para concretização desta dissertação e para manutenção da minha sanidade.

Aos meus pais, Sérgio Roberto V. Sampaio e Nilza Maria C. G. Sampaio, por serem os maiores exemplos de dedicação, caráter e amor que eu poderia ter. Obrigada pelo apoio incondicional a todas as minhas escolhas. A André Augusto V. Sampaio, meu irmão mais velho, meu herói. Sua importância para o desenvolvimento do meu mestrado com certeza é maior do que você imagina, obrigada. À Ana Maria V. Sampaio, minha irmã e melhor amiga para sempre. Sua vinda para Campinas foi o melhor presente que eu poderia ter. Obrigada por todo seu apoio durante esses anos e pelo incentivo a este trabalho.

Ao meu noivo, amigo e eterno companheiro, Luiz Rafael Ferrari Pesce. Sua energia e carinho foram meus maiores pilares durante esses anos. Obrigada por ser o maior incentivador da minha carreira, por compreender todas as minhas escolhas e por completar a minha vida. O seu apoio e segurança foram essenciais para a conclusão de mais esta etapa da minha carreira, e continuarão sendo por toda a minha vida.

A todos os meus familiares, pelo incentivo. Especialmente, à minha sobrinha Cecília, a mais nova princesa da família, por quem já tenho um amor sem limites.

Finalmente, agradeço a Deus, por me conceder sabedoria e saúde para realização deste trabalho.

***“We keep moving forward, opening new doors, and doing new things,
because we're curious and curiosity keeps leading us down new paths”.***

Walter Elias Disney

RESUMO

Nos últimos anos, a busca por fontes renováveis de energia tem sido intensificada em todo o mundo e, atualmente, o processo fermentativo de produção de bioetanol destaca-se como uma alternativa energética economicamente viável e menos poluente do que os combustíveis fósseis. Diferentes etapas do processo de produção desse biocombustível vêm sendo aprimoradas e, recentemente, o melhoramento genético de linhagens da levedura *Saccharomyces cerevisiae* é apontado como um ponto chave para tornar a produção do bioetanol ainda mais rentável. No Brasil, a linhagem JAY270/PE-2 é amplamente utilizada pelo setor industrial devido a sua elevada produtividade e robustez, consideravelmente superiores em relação à linhagem padrão de laboratório S288c. Em busca da compreensão das características genéticas associadas ao fenótipo da linhagem JAY270/PE-2, nosso grupo de pesquisa realizou o sequenciamento de um isolado haploide dessa linhagem, denominado JAY291, o qual revelou um total de 21 genes preditos não identificados no genoma da linhagem S288c. As significativas diferenças fenotípicas observadas entre essas linhagens motivou a investigação do envolvimento destes genes nas características de interesse industrial da linhagem JAY270/PE-2. Para isso, foi realizada uma investigação funcional de 4 desses genes, por meio da avaliação fenotípica de linhagens *knockout*, testadas quanto às suas principais características de interesse industrial. A partir dos testes realizados, não foi possível correlacionar estes genes a uma função específica. Apesar disso, a anotação dessas sequências, a conservação desses genes em outras linhagens industriais e, adicionalmente, a observação de que alguns deles apresentam-se diferencialmente expressos ao longo de uma fermentação industrial forneceram fortes indícios de que os mesmos apresentam um papel importante para a fisiologia da linhagem em condições industriais. Dessa forma, os dados gerados por este trabalho deverão servir como base para futuros estudos focados na elucidação do papel desses genes, a qual é de fundamental importância, tanto para a geração de conhecimento básico acerca da biologia dessa espécie, como para a identificação de potenciais alvos para o melhoramento genético da linhagem JAY270/PE-2.

ABSTRACT

In the past years, the search for renewable energy sources has increased worldwide and, recently, the bioethanol fermentation process rises as an economically viable and less pollutant alternative supply. Several steps of the production process have been enhanced and the genetic improvement of *Saccharomyces cerevisiae* strains plays a key role on that effort. In Brazil, the yeast strain JAY270/PE-2 is widely used in the industrial sector due to its high productivity and robustness, being remarkably superior to the reference laboratory strain S288c. In order to shed light to the genetic characteristics underlying these traits, our group sequenced the genome of the derivative haploid JAY291, which unveiled a total of 21 predicted genes absent in the S288c reference genome. The substantial differences observed between these strains motivated us to investigate the relationship between these predicted genes and the industrially relevant phenotype of JAY270/PE-2 strain. Thus, we accomplished the functional investigation of 4 putative genes through phenotypic evaluation of knockout strains. It was not possible to correlate those genes to a specific function. Although, considering their functional annotation, their conservation between industrial strains and also their differential expression during an industrial fermentation strengthens their potential to be related to these industrially valuable traits. Therefore, this study might work as a backbone for the future functional characterization of these genes. This knowledge is of cardinal importance for a comprehensive understanding of the biology of this specie and could possibly provide new targets for the genetic manipulation of the industrial strain JAY270/PE-2.

DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado/tese de Doutorado intitulada : "Caracterização fenotípica e fisiológica de cepas de *Saccharomyces cerevisiae* mutadas para genes potencialmente envolvidos em fermentação sob condição industrial" :

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

Tem autorização da(s) seguinte(s) Comissão(ões):

(X) CIBio – Comissão Interna de Biossegurança , projeto No. 2011/03, Instituição: Instituto de Biologia, Universidade Estadual de Campinas.

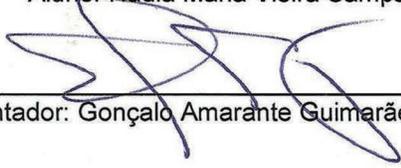
() CEUA – Comissão de Ética no Uso de Animais , projeto No. _____, Instituição:

() CEP - Comissão de Ética em Pesquisa, protocolo No. _____, Instituição:

** Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*

Nádia Maria V. Sampaio

Aluno: Nádia Maria Vieira Sampaio


Orientador: Gonçalo Amarante Guimarães Pereira

Para uso da Comissão ou Comitê pertinente:

(X) Deferido () Indeferido

Carimbo e assinatura


Prof. Dr. MARCELO LANCELOTTI
Presidente da Comissão Interna de Biossegurança
Instituto de Biologia - UNICAMP

Para uso da Comissão ou Comitê pertinente:

() Deferido () Indeferido

Carimbo e assinatura

Sumário

1	INTRODUÇÃO	1
1.1	O etanol como fonte de energia renovável	1
1.2	A via fermentativa na produção de etanol	1
1.3	Linhagens industriais brasileiras e o processo fermentativo industrial.....	2
1.4	Caracterização genética da linhagem industrial brasileira PE-2	4
1.5	Genoma e fisiologia de “linhagens industriais” de leveduras <i>S. cerevisiae</i>	5
1.6	Genes específicos da linhagem JAY291	7
2	OBJETIVOS E IMPORTÂNCIA DO TRABALHO	8
3	MATERIAIS E MÉTODOS	10
3.1	Análise da expressão gênica dos genes novos por <i>RNA-Seq</i>	10
3.2	Avaliação da expressão gênica por <i>Real-Time PCR</i>	11
3.2.1	Coleta de amostras	11
3.2.2	Extração de RNA.....	11
3.2.3	Gel de agarose desnaturante para verificar a integridade do RNA	12
3.2.4	Síntese de cDNA	12
3.2.5	<i>Real-time PCR</i>	13
3.3	Estimativa do número de cópias dos 21 genes novos da linhagem JAY291 ...	15
3.4	Sequenciamento do locus do gene 1947 nas linhagens JAY291 e JAY292....	16
3.5	Esporulação de diplóides e dissecação de tétrades	17
3.6	Obtenção de linhagens mutantes.....	18
3.6.1	Deleções do gene 1947 e do <i>cluster3</i>	18
3.6.2	Inserção do gene 1947 nas linhagens RM11-1a e S288c	19
3.7	Transformação de leveduras.....	20
3.8	Extração de DNA total de leveduras e confirmação dos transformantes	21
3.9	Curvas de crescimento	21
3.10	Teste de tolerância a etanol.....	21
3.11	Teste de tolerância a estresse osmótico.....	22
3.12	Teste de tolerância a temperaturas elevadas	22
3.13	Ensaio de competição.....	22
3.14	Ensaio fermentativos.....	24
3.15	Análise dos metabólitos produzidos durante a fermentação.....	24
3.16	Cálculo dos parâmetros de desempenho fermentativo.....	25

4	RESULTADOS E DISCUSSÃO	26
4.1	Caracterização genômica dos genes novos de JAY291	26
4.1.1	Organização dos genes putativos em <i>clusters</i>	26
4.1.2	Identidade dos genes putativos	27
4.1.3	Expressão gênica dos genes novos de JAY291 em condições industriais	28
4.1.4	<i>Screening</i> do efeito da deleção dos genes novos em linhagens diplóides	35
4.2	Caracterização funcional de quatro genes novos de JAY291	39
4.2.1	Caracterização geral do <i>cluster3</i>	39
4.2.2	Caracterização geral do gene 1947.....	42
4.2.3	Obtenção de linhagens mutantes	49
4.2.4	Ensaio fenotípico com as linhagens mutantes construídas.....	50
4.2.5	Avaliação do crescimento das linhagens <i>knockout</i>	51
4.2.6	Ensaio de tolerância a estresses	53
4.2.7	Ensaio de competição.....	55
4.2.8	Ensaio fermentativo	59
5	CONCLUSÃO	60
6	REFERÊNCIAS BIBLIOGRÁFICAS.....	62

1 INTRODUÇÃO

1.1 O etanol como fonte de energia renovável

A produção de combustíveis renováveis tem se tornado essencial para a manutenção das necessidades energéticas mundiais e alcance de um futuro sustentável, visto a inevitável depleção das reservas mundiais de combustíveis fósseis e o acelerado crescimento da população mundial - estimado para aumentar em aproximadamente três bilhões de pessoas nos próximos 40 anos (Sivakumar et al. 2010; Weber et al. 2010).

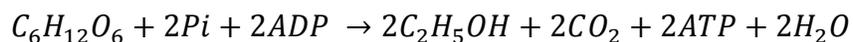
Atualmente, cerca de 90% da necessidade energética mundial provêm de fontes fósseis, como petróleo, carvão e gás natural, sendo o petróleo responsável por 40% desse total. No entanto, flutuações no fornecimento e nos preços deste combustível, devido a inúmeras razões políticas e econômicas, têm motivado a busca por alternativas energéticas (Sivakumar et al. 2010). Nesse contexto, o processo fermentativo de produção de etanol apresenta-se como o mais viável economicamente nas condições brasileiras, além de ser uma fonte energética renovável e menos poluente em relação aos combustíveis fósseis (Goldemberg 2008).

No Brasil, o incentivo à produção deste biocombustível vem sendo intensificado desde o estabelecimento do “Programa Nacional do Álcool”, em 1975, após uma crise no fornecimento do petróleo internacional (Wheals et al. 1999). Há um longo tempo, este programa está bem estabelecido e atualmente o etanol vem substituindo cerca de 40% da gasolina utilizada no país, tornando o Brasil um dos líderes mundiais na produção deste biocombustível (Goldemberg 2008).

1.2 A via fermentativa na produção de etanol

O etanol é o produto final reduzido da via metabólica da fermentação alcoólica, presente nas leveduras da espécie *Saccharomyces cerevisiae*. Nesse processo, a

glicose é oxidada enzimaticamente em diversos passos até a formação de piruvato, que é então convertido a etanol e gás carbônico (Jacques, Lyons, and Kelsall 2003). A reação simplificada é mostrada abaixo.



A partir desta equação é possível deduzir que o rendimento teórico em massa para o etanol é de 51% (massa/massa) e 49% (massa/massa) para o dióxido de carbono. Atualmente, a eficiência deste processo, nas usinas brasileiras de bioetanol, pode alcançar 93% do rendimento teórico máximo do processo fermentativo (Jacques, Lyons, and Kelsall 2003; Amorim et al. 2011). Esta elevada eficiência foi possível por meio do aprimoramento de diferentes etapas do processo produtivo industrial, como o desenvolvimento de novas variedades de cana-de-açúcar, o melhoramento dos métodos de mecanização e processamento nas usinas e, mais recentemente, a seleção de linhagens de levedura altamente adaptadas ao ambiente industrial (Basso et al. 2008), comumente denominadas “linhagens industriais”.

1.3 Linhagens industriais brasileiras e o processo fermentativo industrial

Tradicionalmente, as usinas brasileiras empregam linhagens de *S. cerevisiae* de panificação como inóculo inicial da fermentação alcoólica. No entanto, em diferentes usinas foi observado que, ao final do processo fermentativo, as leveduras comerciais inoculadas são substituídas nas dornas de fermentação por linhagens selvagens, mais robustas e adaptadas ao ambiente altamente estressante das dornas (Da Silva-Filho et al. 2005; Basso et al. 2008).

A partir destas observações, iniciou-se um processo de seleção das linhagens com melhor desempenho e altamente robustas, capazes de permanecer durante todo o processo de fermentação. Dentre as selecionadas, a linhagem de *S. cerevisiae*

denominada Pedra-2 (PE-2) está entre as de melhor desempenho e passou a ser utilizada como inóculo em cerca de 30% das usinas brasileiras (Basso et al. 2008; Argueso et al. 2009). Esta linhagem apresenta muitas características de interesse industrial, como: alta tolerância a etanol, elevada produtividade, osmotolerância (habilidade de fermentar soluções concentradas de carboidratos, como por exemplo, o xarope de cana-de-açúcar), viabilidade celular, tolerância a repetidos ciclos fermentativos, baixa produção de espuma, baixa floculação e tolerância à temperatura (Basso et al. 2008).

Todas essas características conferem a linhagem PE-2 uma vantagem competitiva para sobrevivência no ambiente altamente estressante encontrado nas dornas de fermentação. Nas usinas brasileiras, o caldo de cana-de-açúcar e o melaço são direcionados a grandes dornas (0,5 a 3 milhões de litros), onde a fermentação ocorre por um período de 6 a 10 horas. Uma alta densidade celular (10-15% massa/volume) é utilizada no processo e após seu término o mosto fermentado é centrifugado para separação das células. Até 90% das células são reutilizadas em novos ciclos fermentativos, por cerca de 200 a 300 dias, durante toda a safra (Da Silva-Filho et al. 2005). O vinho, fermentado de cana-de-açúcar sem leveduras, é conduzido para as torres de destilação para a separação do etanol e as células concentradas de leveduras são submetidas à um tratamento com ácido sulfúrico (pH 2,0 – 2,5) que visa a redução da contaminação bacteriana. Ao final de um ciclo fermentativo praticamente todo açúcar é utilizado e uma concentração de etanol de cerca de 7-11% é alcançada (Amorim et al. 2011). Apesar de a prática de reciclagem celular gerar vantagens, como um maior rendimento devido à menor conversão de açúcar a biomassa, essa prática também é responsável por uma exposição constante das células ao ambiente extremamente hostil das dornas de fermentação.

Dessa forma, uma melhor compreensão dos mecanismos envolvidos com a tolerância às condições de estresse inerentes ao processo fermentativo e a identificação dos fatores relacionados à eficiência, robustez e persistência da linhagem industrial PE-2, são essenciais para o melhoramento genético da linhagem e, conseqüentemente, da produção de etanol.

1.4 Caracterização genética da linhagem industrial brasileira PE-2

Apesar da grande importância da linhagem PE-2 no setor industrial brasileiro, suas características genéticas e fisiológicas eram completamente desconhecidas até poucos anos atrás. A falta de informações acerca da biologia dessa linhagem motivou o nosso grupo de pesquisa a realizar um trabalho pioneiro de caracterização genética de um isolado diplóide da linhagem PE-2, denominado JAY270. Além disso, foi realizado o sequenciamento do genoma de um isolado haploide da linhagem, denominado JAY291, o qual revelou um genoma altamente polimórfico, tanto estruturalmente quanto em nível da sequência de nucleotídeos (Argueso et al. 2009).

Análises realizadas por gel de eletroforese de campo pulsado (PFGE) revelaram claros polimorfismos de tamanho entre os cromossomos de JAY270 e da linhagem padrão de laboratório S288c. A maior variação entre o genoma das duas linhagens foi observada principalmente nas regiões teloméricas, enquanto as regiões centrais se mantêm mais conservadas. Esses resultados sugerem que os setores centrais dos cromossomos da linhagem JAY270, que contêm genes essenciais, são refratários a rearranjos cromossômicos, enquanto as regiões periféricas apresentam uma maior plasticidade. Os genes contidos nas regiões teloméricas, em geral, não estão envolvidos em processos essenciais para viabilidade celular e, em alguns casos, podem estar presentes em hemizigose, ou seja, apresentar apenas uma cópia em um par de cromossomos homólogos. Essa característica faz com que polimorfismos de tamanho também sejam observados entre vários pares de cromossomos homólogos de JAY270 (Argueso et al. 2009).

Adicionalmente, este trabalho confirmou uma significativa diferença fenotípica entre a linhagem PE-2 e a linhagem de laboratório S288c. Ensaio de fermentação revelaram que a linhagem JAY270 e seus esporos produzem uma quantidade significativamente maior de etanol durante a fermentação, em relação à linhagem de laboratório. O fenótipo destas linhagens também foi comparado em diferentes condições de estresse, tal como estresse oxidativo e de temperatura, e constatou-se que a linhagem PE-2 apresenta maior tolerância a estas condições adversas e inerentes ao processo industrial (Argueso et al. 2009).

1.5 Genoma e fisiologia de “linhagens industriais” de leveduras *S. cerevisiae*

As diferenças fenotípicas existentes entre as linhagens PE-2 e S288c já foram observadas em outras linhagens empregadas em diferentes setores industriais e, nos últimos anos, esse conhecimento vem motivando o interesse no sequenciamento do genoma dessas linhagens (Borneman et al. 2011). Alguns desses trabalhos têm gerado importantes contribuições na elucidação de fatores genéticos associados às características de interesse dessas linhagens.

Análises do genoma de diferentes linhagens industriais sugerem que a capacidade de adaptação a condições de estresse está relacionada com a elevada plasticidade de seus genomas. Essas linhagens comumente apresentam rearranjos cromossômicos, principalmente em regiões periféricas, que podem resultar em eventos como inserções, deleções, translocações e duplicações (Argueso et al. 2009; Novo et al. 2009; Babrzadeh et al. 2012). Em alguns casos, essas modificações estruturais podem conferir vantagens adaptativas a estas linhagens.

Como exemplo pode-se citar o aumento no número de cópias dos genes *SNO* e *SNZ*, relacionados à biossíntese da vitamina tiamina e de piridoxina (um precursor da tiamina), no genoma de cinco linhagens industriais utilizadas na produção de etanol, incluindo a linhagem PE-2 (Stambuk et al. 2009). Foi demonstrado que a amplificação no número de cópias desses genes está diretamente relacionada ao maior crescimento dessas linhagens em condições de elevada concentração de açúcar, o que sugere uma alta demanda dessa vitamina sob condições industriais (Stambuk et al. 2009). Adicionalmente, outros trabalhos evidenciaram um aumento na expressão e atividade de enzimas relacionadas à produção de tiamina sob condições de estresse osmótico e oxidativo, indicando que esses genes podem ser importantes para a sobrevivência celular sob condições encontradas durante o processo fermentativo (Kowalska and Kozik 2008; Kowalska et al. 2012). O conjunto desses dados sugere que a amplificação dos genes *SNO* e *SNZ* nessas linhagens industriais confere vantagens adaptativas às condições estressantes encontradas nas dornas de fermentação.

Uma translocação recíproca entre os cromossomos VIII e XV, presente em diferentes linhagens empregadas na produção de vinho, também é descrita como um exemplo claro da vantagem conferida por rearranjos cromossômicos para a adaptação dessas linhagens ao ambiente industrial. Essa translocação é responsável por um aumento na expressão do gene *SSU1*, que codifica uma proteína responsável pela extrusão de sulfito. A maior tolerância a este composto confere uma vantagem adaptativa a linhagens que apresentam essa translocação, devido a ampla utilização de sulfito na produção de vinho como um conservante (Pérez-Ortín et al. 2002).

Os rearranjos cromossômicos também podem estar relacionados a inserções ou deleções de genes. Como consequência de análises do genoma de diversas linhagens industriais tem se tornado evidente que uma significativa quantidade de *ORFs* (*Open Reading Frame*) apresentadas por estas linhagens estão ausentes em S288c (Tabela 1), linhagem de referência para o genoma de *S. cerevisiae* (Argueso et al. 2009; Novo et al. 2009; Borneman et al. 2011; Akao et al. 2011). Estes genes podem ter origem em eventos de duplicação ou podem ter sido adquiridos por transferência horizontal. Outra possibilidade é que a persistente propagação da linhagem S288c em condições de laboratório pode ter levado à perda de genes que não apresentam um papel específico para sobrevivência nessas condições de cultivo, resultando em uma representação reduzida do *pan*-genoma da espécie (Borneman et al. 2011).

Tabela 1. Exemplos de linhagens industriais de *S. cerevisiae* que contém ORFs ausentes no genoma da linhagem de referência S288c*.

Linhagem	Nº de ORFs ausentes em S288c	Referência
AWRI796	74	(Borneman et al. 2011)
EC1118	77	(Novo et al. 2009)
FostersB	36	(Borneman et al. 2011)
FostersO	48	(Borneman et al. 2011)
JAY291	16	(Argueso et al. 2009)
Kyokai No.7	48	(Akao et al. 2011)
QA23	110	(Borneman et al. 2011)
VIN13	45	(Borneman et al. 2011)
VL3	54	(Borneman et al. 2011)

* Adaptada de (Engel and Cherry 2013)

Análises de similaridade destas *ORFs* revelam que muitas estão potencialmente associadas com as características de interesse das linhagens industriais. O sequenciamento do genoma de um diplóide comercial da linhagem EC1118 revelou a presença de diversos genes ausentes na linhagem de referência S288c. Dentre eles, 34 estão distribuídos em três grandes regiões cromossômicas, perfazendo um total de 120kb. A análise desses genes, baseada em homologia de sequências com outros micro-organismos, sugere que a função da maioria está associada a pontos fundamentais do processo de produção do vinho, como metabolismo de carbono e nitrogênio, transporte celular e tolerância a estresse (Novo et al. 2009).

Além disso, foi demonstrado que algumas dessas *ORFs* ausentes no genoma da linhagem de referência S288c estão conservadas em diferentes linhagens industriais. Um exemplo é a presença de um *cluster* de cinco genes de sequências altamente conservadas presente em seis linhagens utilizadas na produção de vinho, no isolado natural de campos de vinho RM11-1a e na linhagem de bioetanol JAY291. Estes genes são preditos para codificar dois fatores de transcrição, uma floculina de superfície celular, uma permease para ácido nicotínico e uma 5-oxoprolinase (Borneman et al. 2011). A função dessas *ORFs* para a fisiologia dessas linhagens não está clara, no entanto, foi verificado que este *cluster* apresenta de uma a três cópias entre as linhagens. A amplificação do número de cópias desse *cluster* em algumas dessas linhagens é mais um indício de que o mesmo potencialmente confere características vantajosas às linhagens. A importância do conhecimento das propriedades do genoma dessas linhagens relacionadas à fisiologia das mesmas vem sendo reconhecida como fundamental para o melhoramento do processo fermentativo.

1.6 Genes específicos da linhagem JAY291

Neste contexto, o trabalho de caracterização da linhagem PE-2 também buscou correlacionar características do genoma desta linhagem com as diferenças fenotípicas de interesse industrial existentes entre a mesma e a linhagem de laboratório.

A análise comparativa das sequências dos dois genomas revelou que parte das diferenças estruturais observadas entre estas linhagens ocorrem devido a presença de *ORFs* identificadas no genoma da linhagem JAY291, que estão ausentes em S288c e vice-versa (Argueso et al. 2009). Este trabalho revelou um total de 16 *ORFs* específicas da linhagem JAY291, que entre outros fatores genéticos, podem estar associadas a significativa diferença fenotípica observada entre a linhagem PE-2 e a linhagem S288c durante o processo fermentativo.

Dando continuidade a este trabalho, nosso grupo realizou um estudo da expressão gênica da linhagem sob condições industriais (Tese de doutorado – Osmar Vaz de Carvalho Netto “Fisiologia da fermentação alcoólica: análise da expressão gênica de uma linhagem industrial de *Saccharomyces cerevisiae* durante o processo fermentativo”) utilizando a técnica de RNA-Seq (Wang, Gerstein, and Snyder 2009). Esse trabalho proporcionou um melhor entendimento da fisiologia dessa linhagem e dos mecanismos utilizados contra os estresses impostos pelo processo fermentativo industrial e vem servindo como base para ensaios de manipulação e melhoramento genético da linhagem PE-2 (dados não publicados). Adicionalmente, os dados gerados revelaram que os genes específicos da linhagem JAY291 apresentam uma expressão diferencial ao longo da fermentação industrial, o que suporta a hipótese de que essas *ORFs* podem apresentar relação com as características de interesse industrial da linhagem PE-2. A partir desse conjunto de informações, o presente projeto focou na investigação dessas *ORFs* exclusivas da linhagem PE-2, avaliando se as mesmas poderiam estar correlacionadas com as excelentes características fenotípicas desta linhagem em condições industriais.

2 OBJETIVOS E IMPORTÂNCIA DO TRABALHO

A linhagem industrial PE-2 é amplamente utilizada nas usinas brasileiras de bioetanol por aliar uma elevada eficiência fermentativa a uma excelente capacidade de adaptação ao ambiente altamente hostil encontrado nas dornas de fermentação. No

entanto, as bases genéticas associadas a estas características de interesse são pouco conhecidas.

Nesse cenário, nosso grupo de pesquisa vem dando passos importantes na caracterização genética e molecular da linhagem PE-2, com o objetivo final de desenvolver estratégias de manipulação, domesticação e melhoramento da linhagem. Entre as descobertas geradas pela elucidação do genoma da linhagem JAY291, um derivado haplóide de PE-2, está a identificação de grupos de genes inexistentes no genoma da linhagem de laboratório S288c. Dada a significativa superioridade fermentativa da linhagem PE-2 em comparação com a linhagem laboratorial, o principal objetivo deste trabalho foi investigar se estes genes podem estar relacionados às características de interesse da linhagem industrial PE-2, constituindo assim, potenciais alvos para manipulação genética e melhoramento da linhagem. Para alcançar estes objetivos, este trabalho foi dividido em duas etapas gerais. Os objetivos específicos delineados para cada etapa estão descritos a seguir.

Primeira etapa: caracterização genômica geral dos genes específicos da linhagem JAY291.

- Anotação, análise da localização, número de cópias, expressão em condições industriais e conservação desses genes entre diferentes linhagens de *S. cerevisiae*.

Segunda etapa: investigação funcional de quatro genes específicos da linhagem industrial JAY291.

- Confirmação da expressão dos genes em estudo em fermentações de escala laboratorial, por meio de *Real-Time* PCR;
- Construção de uma coleção de linhagens *knockout*, derivadas de JAY291 e JAY270, para estes genes;

- Inserção de um dos genes em estudo na linhagem de laboratório S288c e na linhagem isolada de campos de vinho RM11-1a;
- Avaliação das linhagens mutantes construídas em testes fenotípicos e fisiológicos.

Antes de iniciar a discussão dos resultados deste trabalho, é importante ressaltar que o uso da denominação dos genes em estudo como “genes novos” ou “específicos” da linhagem JAY291, tem função apenas didática. Como será possível observar na “sessão 4 - Resultados e Discussão”, a maioria dos genes em estudo apresenta similaridade com sequências de genes de outros micro-organismos. Ainda, foi possível atribuir uma função predita para a maioria desses genes e apenas seis deles fazem parte de um grupo que não apresenta similaridade com proteínas de função conhecida. Pode-se concluir, portanto, que nenhum dos genes em estudo, aqui referidos como “genes novos” constitui um *no hit* verdadeiro. Desse modo, a denominação de “genes novos” ou “específicos de JAY291” foi utilizada para representar os genes presentes no genoma da linhagem JAY291, mas ausentes no genoma da linhagem de laboratório S288c.

3 MATERIAIS E MÉTODOS

3.1 Análise da expressão gênica dos genes novos por *RNA-Seq*

Dados obtidos de bibliotecas de *RNA-Seq* geradas pelo trabalho de doutorado de Osmar Vaz de Carvalho Netto, intitulado “Fisiologia da fermentação alcoólica: análise da expressão gênica de uma linhagem industrial de *Saccharomyces cerevisiae* durante o processo fermentativo”, foram utilizados para a análise da expressão dos genes novos da linhagem JAY291 ao longo de uma fermentação industrial. As amostras avaliadas no trabalho citado foram coletadas diretamente de dornas da Usina

Nova América (Maracáí, SP), ao longo de 15 horas de uma fermentação em condições típicas, ou seja, células não floculadas, temperatura da dorna de 34°C e mosto com 17º BRIX (porcentagem de sólidos solúveis totais). Os dados de expressão gênica dos 21 genes novos de JAY291 foram representados pela medida de RPKM (*Reads PerKilobase of the transcript per Million mapped reads*). Esta medida leva em consideração não só o número de sequências alinhadas em determinado gene, como também o tamanho deste gene e o número total de sequências da amostra (Mortazavi *et al.*, 2008).

3.2 Avaliação da expressão gênica por *Real-Time* PCR

3.2.1 Coleta de amostras

As amostras para extração de RNA foram coletadas ao longo de um ensaio de fermentação da linhagem JAY270, realizado em triplicata. A fermentação foi conduzida em *erlenmeyers* de 125 ml, contendo 50 ml de xarope de cana-de-açúcar (8º BRIX, 0,01% de extrato de levedura), os quais foram mantidos sob baixa agitação (80rpm), a 30°C. A fermentação foi iniciada com uma alta densidade celular (OD₆₀₀ de 7), e amostras de 2ml foram retiradas nos pontos 0h, 5h, 10h, 24h e 30h de fermentação. As amostras foram imediatamente centrifugadas a 8000rpm, 4°C. O *pellet* de células foi lavado com água DEPC 0.01% e, em seguida, congelado em nitrogênio líquido, para posterior extração de RNA.

3.2.2 Extração de RNA

A extração de RNA das amostras foi realizada seguindo o protocolo descrito por Ausubel e colaboradores (1998). O *pellet* de células foi ressuscitado em 400µL de tampão TES (10mM Tris.HCl; 10mM EDTA pH8,0; 0,5% SDS). Em seguida, foram adicionados 400µL de fenol ácido. As amostras foram vortexadas por 10 segundos e incubadas a 65°C por 60 minutos, sendo agitadas ocasionalmente. Após este período, as amostras foram mantidas em banho de gelo por 5 minutos e, então, foram centrifugadas a 14000rpm, durante 5 minutos, a 4°C. A fase aquosa obtida foi

transferida para um novo tubo e tratada com 400µL de fenol ácido. As amostras foram submetidas a banho de gelo e, em seguida, centrifugadas. A fase aquosa foi transferida para novos tubos, nos quais foram adicionados 400µL de clorofórmio e álcool isoamílico (proporção 24:1). Os tubos foram vortexados e centrifugados novamente. A fase aquosa obtida foi precipitada com 40µL de acetato de sódio 3M pH5,2 e 1ml de etanol 100% gelado. As amostras foram centrifugadas a 14000rpm, por 5 minutos, a 4°C. Em seguida, o *pellet* foi lavado com etanol 70% e ressuspendido em 40µL de água DEPC 0,01% gelada.

3.2.3 Gel de agarose desnaturante para verificar a integridade do RNA

Inicialmente, todo o aparato necessário para a preparação do gel de agarose foi tratado com peróxido de hidrogênio 40%, por 30 minutos. Foram solubilizados 0,75g de agarose em 42,5ml de água DEPC 0,01%. Após o resfriamento, foram adicionados 5 ml de MOPS 10x e 2,5ml de formaldeído. Em seguida, o líquido foi vertido em uma barca, previamente tratada com peróxido de hidrogênio.

Um total de 1µg de RNA foi diluído em 24µL de tampão de carregamento (5µL de MOPS 10x, 9µL de formaldeído, 25µL de formamida, 2µL Loading Buffer -RNA, 8µL de água DEPC 0,01% e 1µL de brometo de etídio). A solução foi desnaturada a 65°C por 10 minutos, seguida da imediata incubação a 4°C, por 2 minutos. As amostras foram aplicadas no gel, que foi mantido em cuba contendo tampão MOPS 1X, por cerca de 1 hora a uma voltagem de 90 V. Foi verificada a integridade das amostras, que posteriormente foram quantificadas em *Nanodrop 2000 (Thermo Scientific)*.

3.2.4 Síntese de cDNA

Inicialmente, as amostras foram tratadas com DNase (*Promega*), utilizando-se o protocolo fornecido pelo fabricante, detalhado na tabela abaixo.

Tabela 2. Condições utilizadas no tratamento das amostras de RNA com DNase

Reagente	Volume (μL)
RNA ($1\mu\text{g}/\mu\text{L}$)	1
RQ1 DNase Buffer 10x	1
RQ1 RNase-free DNase	1
H2O DEPC 0.01%	7

As amostras foram mantidas a 37°C durante 30 minutos para permitir a ação da enzima, em seguida, foi adicionado em cada tubo $1\mu\text{L}$ de RQ1 *Stop Solution* para parar as reações. A enzima foi inativada através de incubação a 65°C , por 10 minutos. Para comprovar a ausência de DNA genômico nas amostras, o RNA tratado de todas as condições foi utilizado como molde em uma reação de PCR-RT com *primers* específicos para amplificação do gene 1947.

O RNA total das amostras ($1\mu\text{g}$) foi utilizado como molde para a síntese de cDNA utilizando-se a enzima *SuperscriptII* (Invitrogen), de acordo com as informações do fabricante. A cada amostra de $1\mu\text{g}$ de RNA foram adicionados $1\mu\text{L}$ de dNTP (10mM de cada base), 150ng de *random primers* e o volume de água DEPC necessário para um completar $12\mu\text{L}$ de reação. Os tubos foram aquecidos a 65°C por 5 minutos e esfriados rapidamente em banho de gelo. Em seguida, foram adicionados $4\mu\text{L}$ de *First Strand Buffer 5X*, $2\mu\text{L}$ de $0,1\text{M}$ DTT e $1\mu\text{L}$ de RNase OUT e as amostras foram incubadas a 25°C por 2 minutos. Após este período, foi adicionado $1\mu\text{L}$ da enzima *SuperscriptII*. Imediatamente, as amostras foram submetidas a um ciclo de 25°C por 10 minutos, 42°C por 50 minutos, seguido da inativação da enzima a 70°C por, 15 minutos.

3.2.5 Real-time PCR

O *Real-time PCR* foi realizado no aparelho *StepOne Plus* utilizando o *kit SYBR® Green PCR Master Mix* da Applied Biosystem. A Tabela 3 contém a reação padrão empregada na amplificação desses genes.

Tabela 3. Protocolo padrão empregado nas reações de *Real-Time* PCR

Reagente	Volume (μ L)
SYBER Green PCR Master Mix	8
<i>Primer forward</i>	1
<i>Primer reverse</i>	1
cDNA	1
H ₂ O	5

O programa de corrida empregado nas reações foi: desnaturação (95°C por 10 minutos), amplificação e quantificação (40 ciclos de 95°C por 15 segundos, 56°C por 15 segundos, 60°C por 60 segundos), avaliação da *melting curve* (95°C por 15 segundos, 60°C por 1 minuto e 95°C por 15 segundos). A Tabela 4 fornece informações adicionais acerca dos genes analisados e dos *primers* utilizados neste experimento.

Tabela 4. Identificação dos genes e *primers* utilizados nos experimentos de Real-time PCR.

Gene	Amplicon (pb)	Sequência dos <i>primers</i> 5'-3'	
		<i>Forward</i>	<i>Reverse</i>
1947	97	GACACACAAGGCATAGGT	TTCGGGACCGTCACATTAAG
5218	110	CCGTTCCCTCACCAAGATATG	GTAGCTTATCAGTCCAGACA
5219	95	ATATATCAGCAGAGGACCAG	TGTCCCCCATTAAAATCGGA
5220	106	GCAGAGAGAGACAGTCCT	GGCTTTAACTTTGTAGACTC
<i>ACT1</i>	95	CTTTCAACGTTCCAGCCTTC	ATCACCGGAATCCAAAACAA
<i>ALG9</i>	98	GGTGAACAATTACACAGCTC	CGTCCCGTACATACATTC

A expressão de cada gene foi avaliada em relação ao ponto de coleta de 5 horas de fermentação, e os genes *ACT1* e *ALG9* foram utilizados como controles endógenos. O uso destes genes como controles para *S. cerevisiae* é indicado pela literatura (Teste et al. 2009). A adequação do uso dos mesmos para as condições utilizadas nesse experimento também foi confirmada, visto a pequena variação em expressão

observada ao longo da fermentação. O valor de *threshold* estabelecido para a análise de todas as amostras foi de 0,8. Os dados foram analisados por meio do método da eficiência (Pfaffl 2001), que leva em consideração a eficiência dos *primers* utilizados em cada reação para o cálculo da expressão relativa de cada gene.

3.3 Estimativa do número de cópias dos 21 genes novos da linhagem JAY291

A estimativa do número de cópias dos 21 genes em estudo foi realizada por meio de novos dados de sequenciamento gerados para quatro haploides derivados da linhagem JAY270 (JAY289, JAY290, JAY291 e JAY292). A estimativa foi baseada no número de *reads* alinhados em cada gene, levando-se em consideração o tamanho do gene (pb) e o tamanho da biblioteca. A fórmula utilizada para o cálculo do número estimado de cópias de cada gene está descrita abaixo.

$$\text{N}^{\circ} \text{ de cópias} = \frac{\text{Total de reads alinhados ao gene}}{\text{Tamanho do gene (pb) X N}^{\circ} \text{ reads da biblioteca}}$$

Os valores obtidos foram comparados com os valores observados para uma região do cromossomo IX, sabidamente presente em uma cópia no genoma do haploide JAY291, e em duas cópias no genoma do diplóide JAY270. A Figura 1, apresenta um esquema simplificado de valores reais utilizados para estimar o número de cópias do gene 5218 na linhagem JAY291. O mesmo cálculo foi feito para os outros genes em estudo e para as diferentes linhagens analisadas.

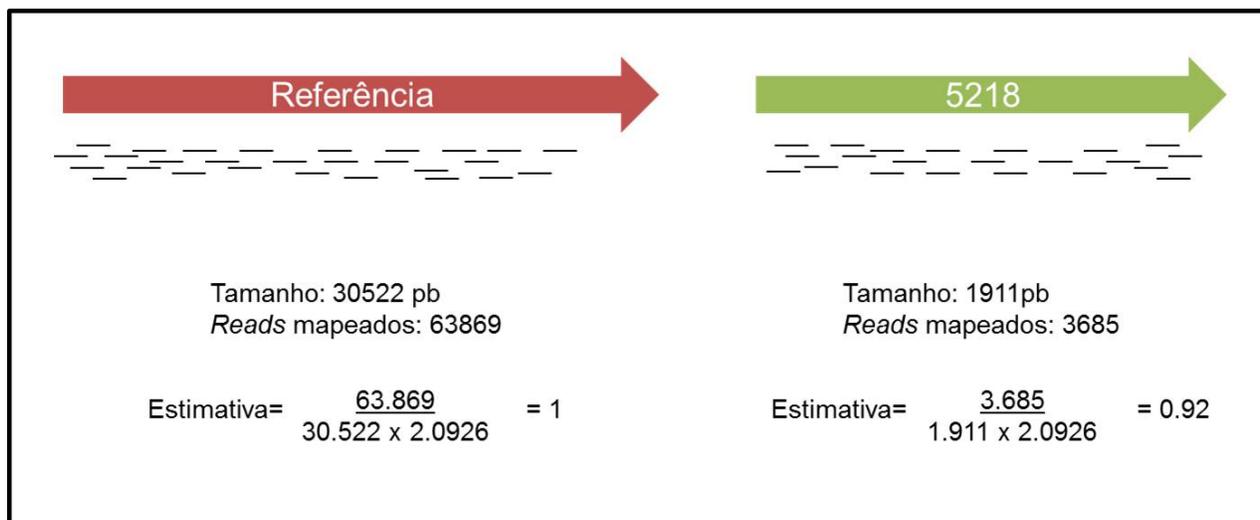


Figura 1. Esquema exemplificando o cálculo realizado para estimar o número de cópias do gene 5218 na linhagem JAY291. Este gene apresenta um valor estimado muito próximo de um, assim como obtido para a sequência de referência presente em uma cópia no genoma desta linhagem.

3.4 Sequenciamento do locus do gene 1947 nas linhagens JAY291 e JAY292

O DNA das linhagens JAY291 e JAY292 utilizado para as reações de sequenciamento foi extraído seguindo o protocolo descrito por Ausubel e colaboradores (1998). As amostras, contendo 320ng de DNA e 10pmol do *primer* de interesse, foram enviadas para sequenciamento na *Proteomics and Metabolomics Facility*, localizada na *Colorado State University*. O sistema ABI 3130xL *Genetic Analyzer* é utilizado pela *facility* para sequenciamento e as amostras são preparadas com o *kit* BigDye Terminator v3.1, utilizando o polímero POP7.

O sequenciamento do *gap* entre o *contig178*, que contém o gene 1947, e o *contig467*, que contém o gene *THI21*, foi realizado seguindo o modelo *primer walking*, esquematizado na figura abaixo.

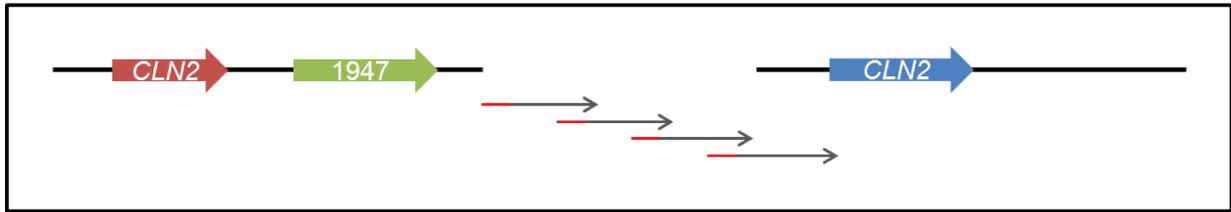


Figura 2. Esquema do modelo de sequenciamento *primer walking*, empregado para a elucidação da sequência flanqueadora do gene 1947. As setas em cinza representam as sequências obtidas após cada rodada de sequenciamento e as extremidades em vermelho representam os *primers* utilizados.

3.5 Esporulação de diplóides e dissecação de tétrades

Os métodos de esporulação e dissecação de tétrades foram conduzidos de acordo com os protocolos de Rose et al. (1990). A indução da esporulação de células diplóides da linhagem contendo a deleção de duas cópias do gene 1947 foi realizada através do plaqueamento em meio de esporulação (1% acetato de potássio, 0.1% extrato de levedura, 0.05% glicose, 2% ágar), com incubação de cerca aproximadamente 7 dias.

Após esse período, uma pequena quantidade de células foi ressuspensa em 400µL de tampão de micromanipulação (Sorbitol 2M, Tris.Cl 1M - pH7,5, NaH₂PO₄ 0,5M, EDTA 0,5M - pH8,0, água) a fim de estender a viabilidade dos esporos. Adicionou-se 4µL de β-mercaptoetanol seguida, após breve agitação, da adição de 20µL de Liticase 1mg/ml (*Sigma-Aldrich*). A solução com células foi mantida em estufa a 30°C por 50 minutos. Em seguida, 15µL da solução foram aplicados em placa de meio YPD e as tétrades foram dissecadas no micromanipulador Nikon Eclipse E400. Para a análise do padrão de segregação do gene 1947, os esporos obtidos foram plaqueados em placas YPD contendo os antibióticos geneticina (200mg/L) ou higromicina (300mg/L).

3.6 Obtenção de linhagens mutantes

3.6.1 Deleções do gene 1947 e do *cluster3*

A deleção de toda a sequência do *cluster3* e do gene 1947 foi realizada em diferentes isolados da linhagem haploide JAY291. Para o gene 1947 realizou-se, adicionalmente, a sua deleção na linhagem diplóide JAY270. A deleção dessas regiões foi obtida através da inserção de cassetes integrativos via recombinação homóloga. Esses cassetes possuem, nas suas extremidades, 42 pares de bases de homologia com as regiões flanqueadoras dos genes-alvo. Além disso, estes cassetes carregam genes de seleção que conferem resistência aos antibióticos geneticina ou higromicina, respectivamente, *KanMX4* e *Hph*, que permitem a seleção de linhagens com deleção de uma ou duas cópias do gene de interesse. Na figura abaixo, pode-se observar um esquema dos plasmídeos pFA6KanMx4 e pAG32, utilizados como molde para amplificação dos genes *KanMx4* e *Hph*, assim como uma representação do mecanismo geral de deleção via-recombinação homóloga empregado nas transformações das linhagens haploides JAY291 (Figura 3, A) e da linhagem diplóide JAY270 (Figura 3, B).

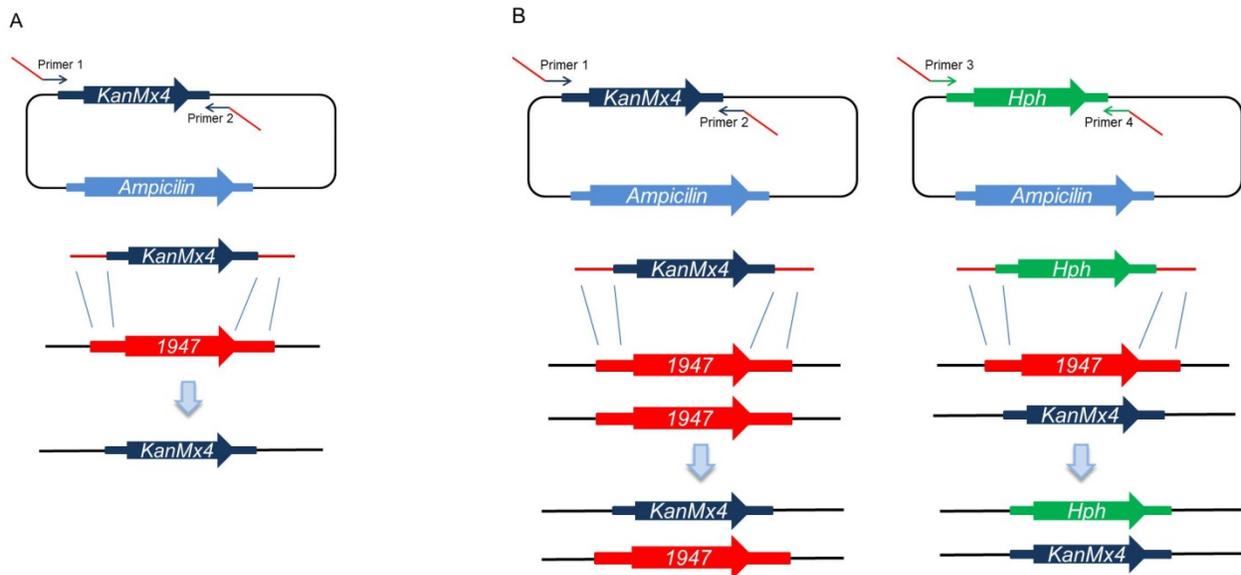


Figura 3. Representação esquemática dos plasmídeos *pFA6KanMX4* e *pAG32* e do mecanismo geral de recombinação homóloga responsável pela deleção dos genes de interesse, utilizando-se como exemplo a deleção do gene 1947. A) Representação dos passos seguidos para deleção dos genes de interesse nas linhagens haploides. B) Representação dos passos necessários para deleção do gene 1947 na linhagem diplóide JAY270. As linhas em vermelho apresentadas em cada *primer* representam as regiões de homologia com as regiões flanqueadoras do gene de interesse.

3.6.2 Inserção do gene 1947 nas linhagens RM11-1a e S288c

O cassete utilizado na transformação das linhagens selvagens foi obtido através da ligação dos genes *KanMx4* e 1947, por meio de um *overlapping* PCR (Ho et al. 1989). Adicionalmente, nas extremidades do cassete foram inseridas caudas de 42 pares de bases de homologia com o centrômero V (CENV) das linhagens RM11-1a e S288c.

Inicialmente, amplificou-se o gene *KanMx4* utilizando-se um *primer forward*, que anela na região *upstream* do gene e contém uma cauda de homologia à região do CENV de cada linhagem, e um *primer reverse*, que anela na região *downstream* do gene e contém uma cauda de homologia com a região *upstream* do gene 1947. O gene 1947 foi amplificado utilizando-se um *primer forward*, que anela na região *upstream* do gene e apresenta uma cauda de homologia à região *downstream* do gene *KanMx4*, e um *primer reverse*, que anela na região *downstream* do gene e contém uma cauda de homologia ao CENV. O *primer forward* utilizado anela em uma região cerca de 840pb anterior ao início do gene 1947, de modo a garantir a inserção da região promotora do gene. Em seguida, os dois fragmentos foram unidos através de uma reação de PCR utilizando os dois cassetes como molde e os dois *primers* que contêm as caudas com homologia ao CENV. Segue abaixo um esquema das reações realizadas.

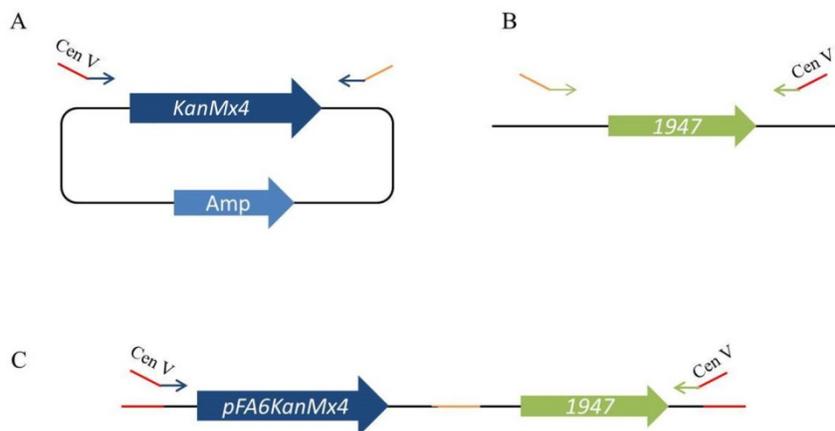


Figura 4. Reações empregadas na construção do cassete contendo os genes *KanMx4* e 1947. (A e B) Reação de amplificação do gene *KanMx4* e 1947 utilizando-se *primers* contendo caudas de homologia com o centrômero V das linhagens RM11-1a e S288c. (C) Reação realizada para ligar os genes *KanMx4* e 1947. O cassete obtido foi utilizado nos experimentos de transformação genética das linhagens RM11-1a e S288c.

3.7 Transformação de leveduras

A etapa de preparação das células competentes foi baseada no protocolo de Ausubel (Ausubel et al. 1998) e as leveduras foram transformadas pelo método de Acetato de Lítio (Gietz et al. 1995).

O preparo das células competentes foi iniciado através do pré-inóculo das células em 5 ml de meio líquido YPD (meio completo composto por 1% de extrato de levedura, 2% de peptona e 2% de glicose). Em seguida, este volume foi adicionado à 200 ml de meio líquido YPD e a cultura foi mantida em rotação 250 rpm, a 30°C, até a fase de crescimento exponencial das células (aproximadamente 4h). As células foram centrifugadas, lavadas com água destilada e ressuspensas em 1ml de solução 1X TE/LiAc. Em tubos de 1,5ml foram adicionados 12,5µL do DNA de interesse e 100µg de esperma de salmão previamente denaturado durante 10 minutos a 30°C. Em seguida, foram adicionados 100 µL de células competentes e 600µL de uma solução 40% PEG/1X LiAc (50% PEG; 10X LiAc; água). Após homogeneização, as células foram incubadas a 30°C durante 30 minutos, sob agitação de 250 rpm. Após este período, foram adicionados 70µL de DMSO e a solução foi misturada por inversão.

Na etapa seguinte, as células foram aquecidas a 42°C por 15 minutos e imediatamente transferidas para banho de gelo por 2 minutos. Após a adição de 700µL de água, a solução foi centrifugada por 30 segundos (13000 rpm, temperatura ambiente).

O sobrenadante foi descartado e as células ressuspensas com a água restante no tubo. Em seguida, as células foram plaqueadas em meio sólido YPD (meio completo com adição de 2% de ágar) e incubadas a 30°C. Após 24 horas, foram feitas réplicas em placas contendo o antibiótico de interesse e a cultura foi incubada a 30°C por 2 a 4 dias até a formação de colônias.

3.8 Extração de DNA total de leveduras e confirmação dos transformantes

A confirmação da deleção ou inserção dos genes foi realizada em duas etapas. A primeira etapa consistiu na seleção das colônias transformantes em meio YPD contendo o antibiótico geneticina (200mg/L) ou higromicina (300mg/L). Em seguida, o DNA total das colônias selecionadas no meio seletivo com antibiótico foi extraído, seguindo o protocolo de Ausubel e colaboradores (1998), e utilizado como molde nas reações de PCR empregadas para confirmação da deleção ou inserção de cada gene/*cluster* gênico. Os *primers* utilizados anelam na região upstream do gene alvo e no interior do gene de resistência a antibiótico. Adicionalmente, para confirmar a correta inserção do gene 1947 nas linhagens S288c e RM11-1a realizou-se o sequenciamento da sequência completa do gene para as colônias selecionadas. Culturas permanentes das linhagens com uma e duas cópias deletadas de cada gene foram estocadas a -80°C (Tabela 8).

3.9 Curvas de crescimento

O pré-inóculo de cada cultura foi feito em meio YPD e mantido por 24 horas em *shaker* a 30°C, 250rpm. Após esse período, a densidade óptica (OD₆₀₀) de todas as amostras foi aferida e calculou-se o volume necessário de cultura para um inóculo de 50ml (YPD 2% glicose) com OD inicial 0,1. Os testes foram realizados em triplicata e mantidos a 30°C, 250rpm. O crescimento das linhagens foi avaliado em intervalos de 1 a 2 horas através da aferência da densidade óptica (OD₆₀₀) das culturas por leitura em espectrofotômetro Ultrospec 2000 UV/VIS (*Pharmacia Biotech*).

3.10 Teste de tolerância a etanol

Este ensaio foi realizado seguindo o mesmo procedimento experimental geral adotado para as curvas de crescimento (ver seção 3.9). Neste caso, no entanto, os

inóculos foram feitos em meio YPD contendo etanol a uma concentração final de 8%. O crescimento das linhagens foi avaliado em intervalos de 2 a 6 horas através da aferência da densidade óptica (OD_{600}) das culturas por leitura no espectrofotômetro Ultrospec 2000 UV/VIS (*Pharmacia Biotech*).

3.11 Teste de tolerância a estresse osmótico

O pré-inóculo de cada cultura foi feito em meio YPD e mantido por 24 horas em shaker a 30°C, 250rpm. Após esse período, aferiu-se a OD do pré-inóculo e calculou-se o volume necessário para iniciar um inóculo de 5ml de meio YPD com OD inicial de 0,1. O inóculo foi mantido por 6 horas em shaker a 250rpm, 30°C. Em seguida, foram realizadas diluições seriadas da cultura até 10^{-5} . Utilizando-se pipeta multicanal, foram plaqueados 3 μ L de cada diluição de cada linhagem em placas de meio YPD contendo 2%, 10% e 18% de glicose e 1M de NaCl. As placas foram mantidas por cerca de 30 horas em estufa a 30°C até a avaliação do crescimento. Os experimentos foram conduzidos em triplicata.

3.12 Teste de tolerância a temperaturas elevadas

Este ensaio foi realizado em triplicata seguindo o mesmo procedimento experimental geral adotado para a avaliação da tolerância das linhagens ao estresse osmótico (ver seção 3.11). Neste caso, no entanto, as diluições seriadas de células foram plaqueadas em YPD 2% de glicose. As placas foram mantidas a 30°C, 37°C ou 41°C, por um período de 30 a 72 horas, ao longo do qual o crescimento das colônias foi avaliado.

3.13 Ensaios de competição

As linhagens mutantes foram co-cultivadas com as respectivas linhagens parentais selvagens em ensaios de competição de 15 ou 20 ciclos. As culturas foram

crescidas em xarope de cana-de-açúcar (18° BRIX, 0.01% de extrato de levedura), sob agitação de 250rpm a 30°C. Inicialmente, todas as linhagens foram inoculadas em 5mL de xarope e mantidas por 24h para garantir a saturação de crescimento. Em seguida, 1% de células foi transferida para novo meio, a fim de induzir uma adaptação das mesmas antes do início da competição. Visando o ajuste da taxa inicial de células da linhagem parental e de células das linhagens mutantes para aproximadamente 1 (50% de cada) um mesmo volume de inóculo saturado de células selvagens e mutantes foi adicionado ao novo meio de cultura. A cada 24h (um ciclo), 0,02% da cultura foram transferidas para um novo meio para crescimento e fermentação (devido ao baixo crescimento da linhagem S288c, 1% dessa cultura foi transferida para novo meio). As mudanças na população celular foram monitoradas periodicamente a cada 5 ciclos. Para isso, plaqueou-se uma diluição de cada cultura em meio YPD e após um período entre 24 e 48 horas, foram feitas réplicas em placas de YPD suplementadas com o antibiótico geneticina (200mg/mL). Um esquema geral do ensaio pode ser observado na Figura 5. A contagem das colônias foi realizada nas placas contendo geneticina. Colônias resistentes, constituídas por células mutantes, apresentam crescimento normal. A identificação das colônias sensíveis, constituídas por células parentais, é possível devido ao baixo crescimento das colônias, que se apresentam na forma de pequenas marcas na placa de cultura.

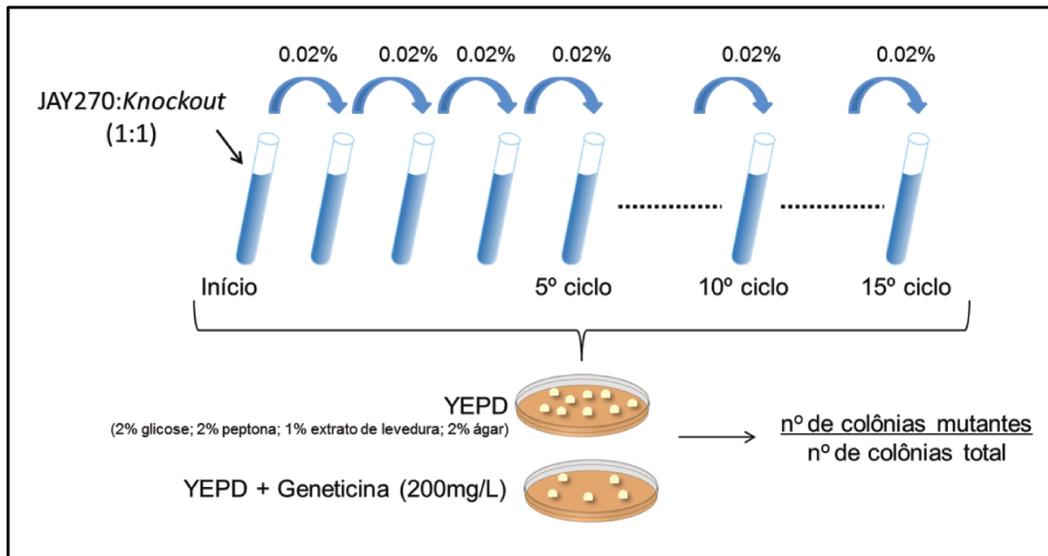


Figura 5. Esquema geral dos experimentos de competição entre as linhagens mutantes produzidas e suas respectivas linhagens parentais.

3.14 Ensaios fermentativos

O pré-inóculo foi obtido pela adição de 2mL de cultura permanente de cada linhagem a 200mL de meio YPD e foi mantido por 24 horas, a 30°C sob agitação de 250rpm. O inóculo foi feito em 50ml de meio YPD com OD inicial 0,1, constituindo o ciclo zero da fermentação. Em seguida, foram realizados cinco ciclos fermentativos de 12 horas de fermentação. Ao final de cada um dos quatro primeiros ciclos, as células foram centrifugadas e ressuspensas em 50ml de meio novo. Ao início do quinto ciclo, a densidade óptica (OD_{600}) dos inóculos foi novamente ajustada para 0,1, no qual foram retiradas amostras para avaliação da cinética fermentativa.

3.15 Análise dos metabólitos produzidos durante a fermentação

Durante o quinto ciclo de fermentação foram coletadas amostras nos pontos 0h, 3h, 6h, 9h e 12h, para análise da cinética fermentativa por Cromatografia Líquida de

Alta Eficiência (HPLC). As amostras foram congeladas imediatamente em nitrogênio líquido e estocadas a -20°C. A presença e concentração do etanol, glicose e glicerol nas amostras provenientes dos testes fermentativos foram verificadas por HPLC. As amostras foram previamente centrifugadas e filtradas (filtro Millipore 0,22µm). Em seguida, foram injetadas em cromatógrafo Alliance-HT (*Waters*) com a coluna Aminex HPX 87H (*BioRad*) com ácido sulfúrico a 2 mM como fase móvel, à temperatura de 35°C e com fluxo de 0,6 mL/min. A detecção dos compostos foi realizada por Índice de Refração (detector IR 2414, *Waters*) e a quantificação dos compostos foi possível pela comparação com curvas de concentração padrão de glicose, etanol e glicerol.

3.16 Cálculo dos parâmetros de desempenho fermentativo

Com a análise quantitativa dos substratos e produtos da fermentação, o desempenho dos processos foi avaliado através dos seguintes parâmetros:

- **Rendimento:** Massa de produto formado por unidade de massa consumida, expresso em porcentagem (Doran 1995).

$$\text{Rendimento \% (m/m)} = \frac{\text{MF produto} - \text{MI produto}}{\text{MI substrato} - \text{MF substrato}} \times 100$$

No qual,

MF produto: massa final do produto analisado, expresso em gramas.

MI produto: massa inicial do produto analisado, expresso em gramas.

MI substrato: massa inicial do substrato utilizado, expresso em gramas.

MF substrato: massa final do substrato utilizado, expresso em gramas.

- **Produtividade Volumétrica:** massa de produto formado por unidade de volume, por unidade de tempo.

$$\text{Produtividade Volumétrica (g/L.h)} = \frac{\text{MF produto} - \text{MI produto}}{\text{V.t}}$$

No qual,

MF produto: massa final do produto analisado, expresso em gramas.

MI produto: massa inicial do produto analisado, expresso em gramas.

V: Volume final da fermentação, expresso em Litros.

t: tempo de cultura, expresso em horas.

4 RESULTADOS E DISCUSSÃO

4.1 Caracterização genômica dos genes novos de JAY291

4.1.1 Organização dos genes putativos em clusters

Após a publicação do genoma da linhagem JAY291, novos sequenciamentos foram realizados culminando no aumento da cobertura do genoma, o que permitiu uma montagem mais completa do que a realizada anteriormente no trabalho de Argueso e colaboradores (2009). Como parte do presente trabalho, uma nova análise comparativa da montagem atual de JAY291 contra o genoma de S288c foi realizada, e encontrou-se um total de 21 sequências gênicas não identificadas no genoma da linhagem S288c.

A predição dos genes novos da linhagem JAY291 revelou que os mesmos encontram-se distribuídos em cinco *scaffolds*, sendo que a maioria dos genes situa-se agrupada em regiões específicas de cada *scaffold*. Este agrupamento de *ORFs* que perderam homologia com sequências da linhagem de laboratório S288c também foi observado em outros casos, por exemplo, as linhagens de vinho AWRI796, cujo genoma apresenta uma região de 45kb contendo 19 novas *ORFs* preditas, e EC1118, citada anteriormente (Borneman et al. 2011). A Tabela 5 resume a identificação geral

dos genes novos da linhagem JAY291, identificando a posição de cada gene nos *scaffolds*.

Tabela 5. Dados gerais para identificação dos 21 genes presentes na linhagem JAY291 e ausentes em S288c

Cluster	ID do gene	Nº de acesso do gene	Nº de acesso do scaffold	Posição do gene dentro do scaffold	Orientação da fita
1***	23	EEU07467	ACFL01000080.1	22.901 – 24.412	Negativa
	24	*	ACFL01000080.1	18.557 – 22.420	Positiva
	25	EEU07466	ACFL01000080.1	16.534 – 18.495	Negativa
	1396	*	ACFL01000080.1	9.298 – 10.479	Negativa
	C1Q_05653	EEU07465	ACFL01000080.1	13.749 – 15.049	Positiva
2	876	*	ACFL01000301.1	16.275 – 17.366	Negativa
	877	EEU05584	ACFL01000301.1	13.799 - 15.400	Positiva
	878	EEU05583	ACFL01000301.1	10.423 - 12.357	Negativa
	879	*	ACFL01000301.1	7.959 - 9.599	Negativa
	880	EEU05582	ACFL01000301.1	6.350 - 7.216	Positiva
	881	EEU05581	ACFL01000301.1	3.648 - 5.483	Negativa
	882	EEU06104	ACFL01000301.1	532 - 1.389	Negativa
	883	*	**		Positiva
3	5218	EEU04262	ACFL01000448.1	22.247 - 24.157	Negativa
	5219	EEU04263	ACFL01000448.1	27.114 - 28.784	Negativa
	5220	EEU04264	ACFL01000448.1	29.364 - 30.998	Positiva
4	5323	*	ACFL01000453.1	52.354 - 53.592	Negativa
	5324	EEU04190	ACFL01000453.1	53.646 - 54.446	Negativa
-	5206	*	ACFL01000447.1	4.196 - 3.834	Negativa
-	5207	*	ACFL01000447.1	7.592 - 8.281	Positiva
-	1947	EEU05823	ACFL01000264.1	34.161 - 35.279	Negativa

* Genes identificados após realização da nova montagem, portanto, não possuem número de acesso no EMBL-EBI.

** Scaffold identificado após realização da nova montagem, não possui número de acesso no NCBI.

*** Grupo de genes conservados citado no trabalho de Borneman *et al.* (2011).

4.1.2 Identidade dos genes putativos

A anotação das sequências foi realizada através da análise pelo BLASTx (Altschul et al. 1997). A Tabela 6 apresenta a identidade e *e-value* observados para o primeiro organismo que apresenta uma função associada a sequência de interesse.

Tabela 6. Anotação de 21 genes presentes na linhagem JAY291 e ausentes na linhagem S288c.

ID do gene	Top hit	Função hipotética	Identidade (%)	E-value
23	<i>Saccharomyces cerevisiae</i> EC1118	Transportador	99	0.0
24	<i>Saccharomyces cerevisiae</i> EC1118	Oxoprolinase	100	0.0
25	<i>Zygosaccharomyces bailii</i>	Fator transcricional	100	0.0
1396	<i>Saccharomyces cerevisiae</i> RM11-1a	Flo11 (floculina de superfície celular)	100	3E-138
C1Q_05653	<i>Saccharomyces cerevisiae</i> RM11-1a	Fator de transcrição	100	0.0
876	<i>S. cerevisiae</i> x <i>S. kudriavzevii</i> VIN7	Adh6 (álcool desidrogenase de cadeia média)	49	4.00E-118
877	<i>Debaryomyces hansenii</i> CBS767	Transportador	52	3.00E-163
878	<i>Meyerozyma guilliermondii</i> ATCC 6260	Fator de transcrição	35	2.00E-99
879	<i>Saccharomyces cerevisiae</i> YJM789	Transportador de membrana	59	0.0
880	<i>Yersinia frederiksenii</i>	Arginase	58	6.00E-112
881	<i>Scheffersomyces stipitis</i>	Metabolismo de tiamina	42	3.00E-135
882	<i>Lachancea thermotolerans</i> CBS 6340	Metiltransferase	52	9E-103
883	<i>Saccharomyces paradoxus</i>	Alpha-galactosidase	96	0.0
5218	<i>S. cerevisiae</i> x <i>S. kudriavzevii</i> VIN7	<i>Sit1p</i> (Sideróforo)	89	0.0
5219	<i>S. cerevisiae</i> x <i>S. kudriavzevii</i> VIN7	Fator de transcrição	86	0.0
5220	<i>S. cerevisiae</i> x <i>S. kudriavzevii</i> VIN7	Transportador	90	0.0
5323	<i>Saccharomyces cerevisiae</i> FostersO	Função desconhecida	56	2.00E-23
5324	<i>S. cerevisiae</i> x <i>S. kudriavzevii</i> VIN7	Função desconhecida	80	2.00E-74
5206	<i>S. cerevisiae</i> AWRI796	Função desconhecida	100	1.00E-19
5207	<i>Saccharomyces cerevisiae</i> RM11-1a	Acetiltransferase (MPR1)	100	3.00E-168
1947	<i>Saccharomyces cerevisiae</i> FostersB	Função desconhecida	99	0.0

* As linhas pontilhadas na Tabela separam os *clusters* gênicos e os genes não agrupados que estão presentes na linhagem JAY291. Como será detalhado nas próximas seções, no presente trabalho, o *cluster* contendo os genes 5218, 5219 e 5220, assim como o gene 1947 foram estudados com maior profundidade.

Estes dados sugerem que alguns dos genes novos apresentam grande potencial para desempenhar funções importantes durante o processo fermentativo, como metabolismo de nitrogênio, transporte celular e resistência a estresses. Além disso, os genes de função desconhecida, fatores de transcrição e permeases também são de grande interesse. O efeito da deleção de alguns desses genes foi avaliado e será discutido nas sessões seguintes.

4.1.3 Expressão gênica dos genes novos de JAY291 em condições industriais

Os dados de expressão gênica em condições industriais dos 21 genes novos da linhagem JAY291 foram obtidos através do trabalho de doutorado de Osmar Vaz de Carvalho Netto, intitulado “Fisiologia da fermentação alcoólica: análise da expressão gênica de uma linhagem industrial de *Saccharomyces cerevisiae* durante o processo fermentativo”. No referido trabalho, amostras foram coletadas de uma dorna que foi alimentada até um ponto entre 4 e 9 horas. Por este motivo, a análise quantitativa dos açúcares presentes ao longo da fermentação tornou-se complexa. A quantidade de

sacarose observada é relativamente baixa, pois a mesma é rapidamente convertida à frutose e glicose. Adicionalmente, a quantidade de glicose decai rapidamente e é inferior à de frutose pelo fato de seu consumo ser favorecido em relação ao consumo de frutose. Com resultado, é possível observar uma concentração crescente de frutose até o ponto em que a dorna deixa de ser alimentada. A análise global de expressão gênica da linhagem realizada no referido trabalho, revelou que o pico da fermentação ocorre entre os pontos de 9 e 12 horas de fermentação. Nesses pontos também foi observado uma maior concentração de etanol nas amostras de fermentação. A Figura 6 apresenta as concentrações de açúcares (Figura 6, A) e de etanol (Figura 6, B) presentes nas amostras coletadas em cada ponto da fermentação (dados fornecidos pelo pesquisador Osmar Vaz de Carvalho Netto).

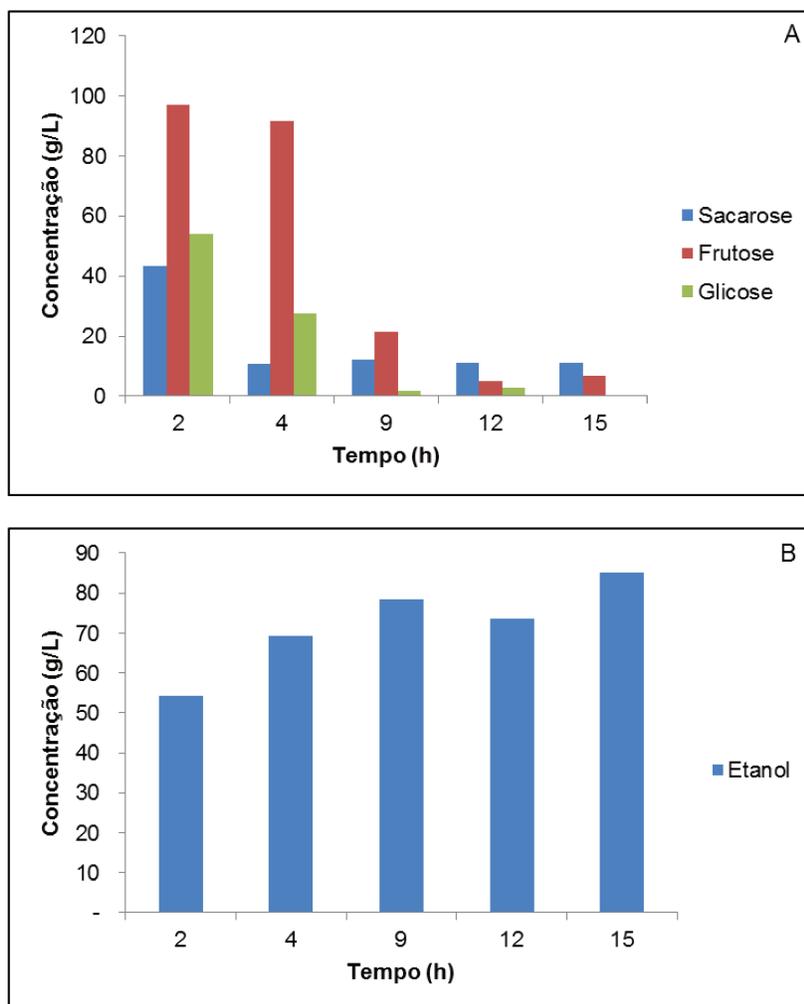
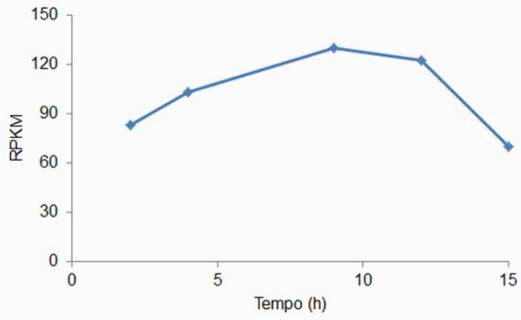


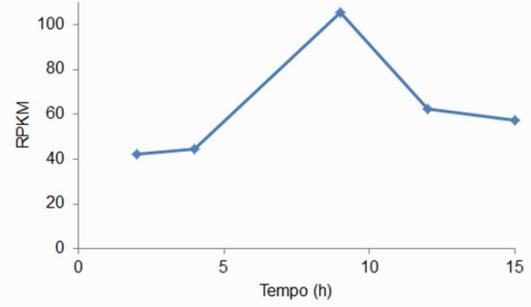
Figura 6. Perfil fermentativo observado na Usina Nova América (Maracaí-SP), em uma fermentação sob condições típicas (células não floculadas, temperatura da dorna de 34°C e mosto com 17º BRIX. A) Cinética da fermentação, representando o consumo dos açúcares sacarose, frutose e glicose. B) Cinética da fermentação, representando a produção de etanol ao longo da fermentação em função de sua concentração nas amostras coletadas. Dados adaptados do trabalho de Osmar Vaz de Carvalho Netto, intitulado “Fisiologia da fermentação alcoólica: análise da expressão gênica de uma linhagem industrial de *Saccharomyces cerevisiae* durante o processo fermentativo”,

Estes dados dão suporte para a análise da expressão destes genes em cada uma das amostras coletadas ao longo da fermentação industrial. Os valores de expressão são apresentados pela medida de RPKM (*Reads PerKilobase of the transcript per Million mapped reads*) e podem ser visualizados na Figura 7.

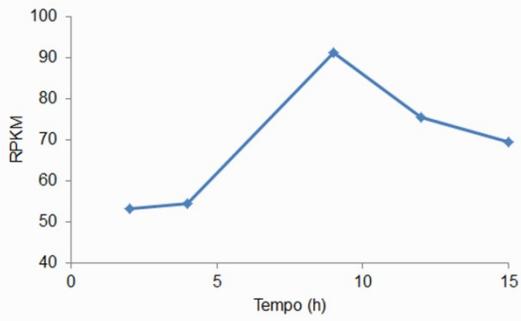
Gene 23



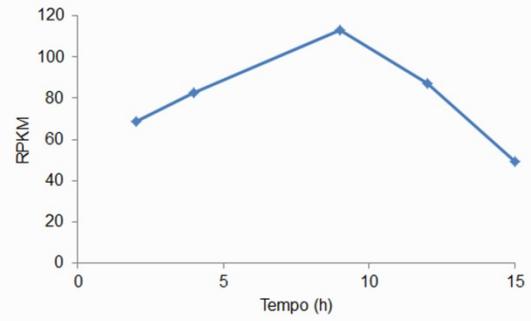
Gene 24



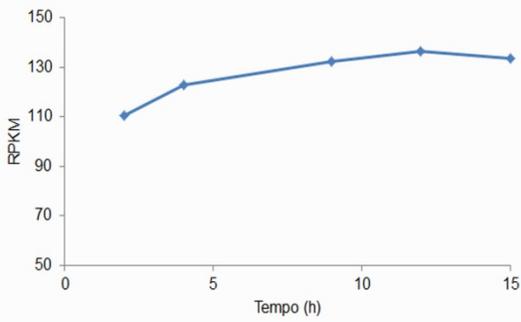
Gene 25



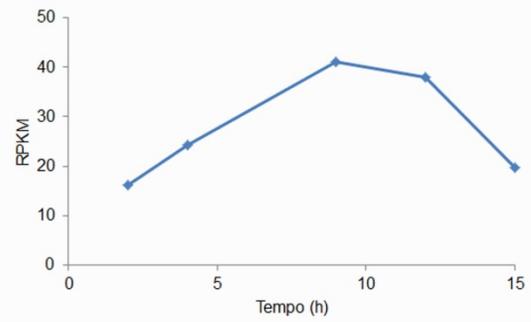
Gene 1396



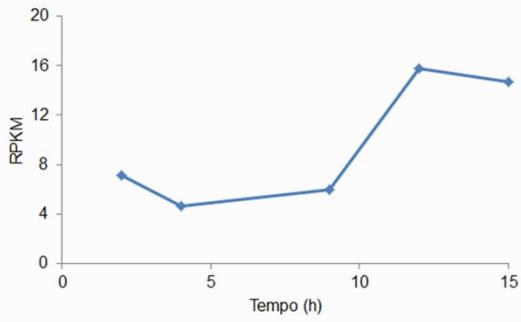
Gene C1Q_05653



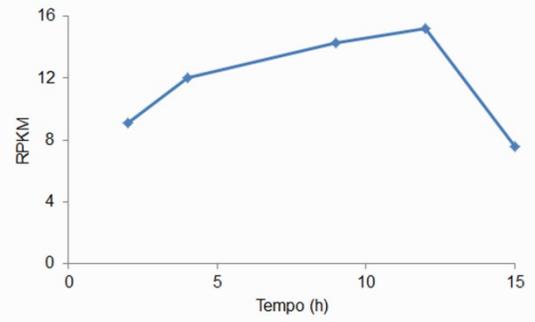
Gene 876



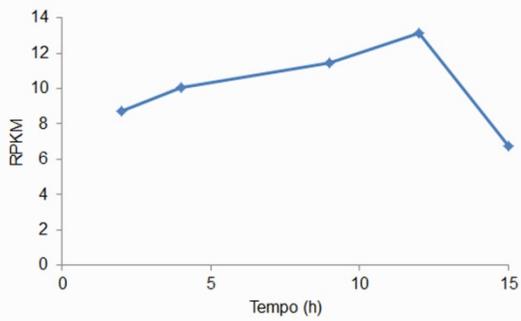
Gene 877



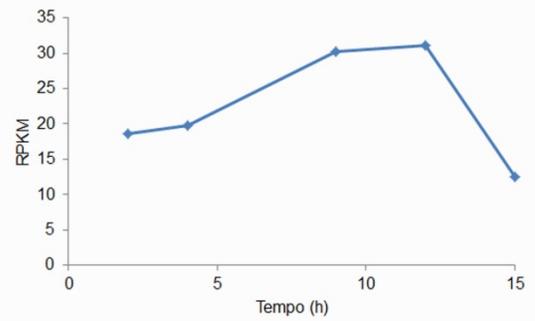
Gene 878



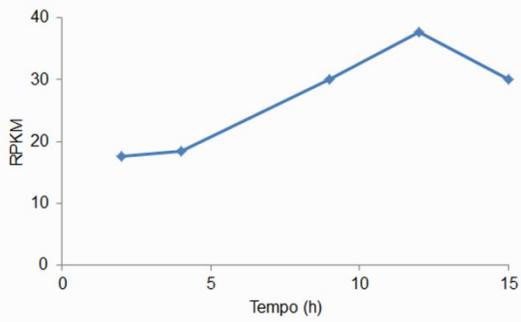
Gene 879



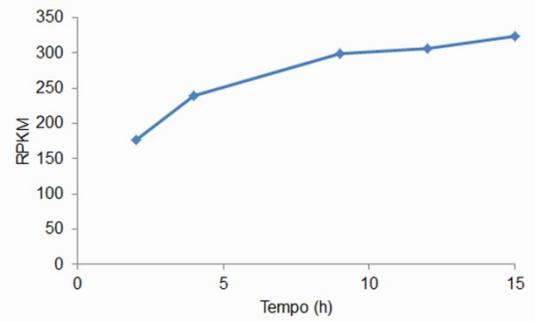
Gene 880



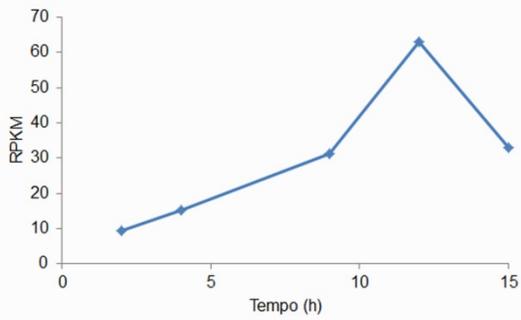
Gene 881



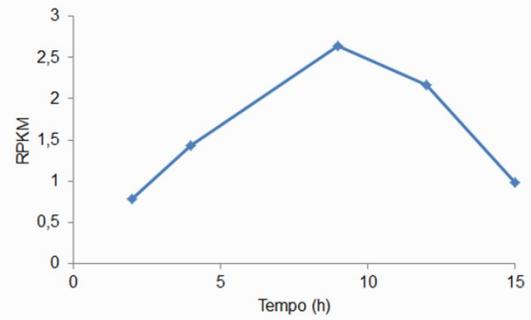
Gene 882



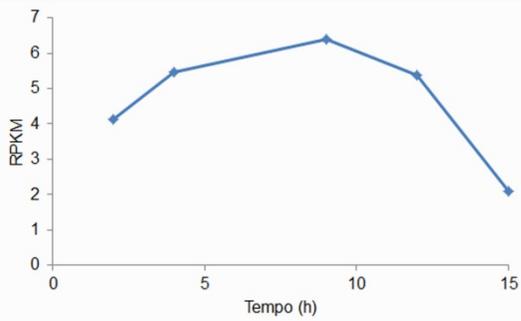
Gene 883



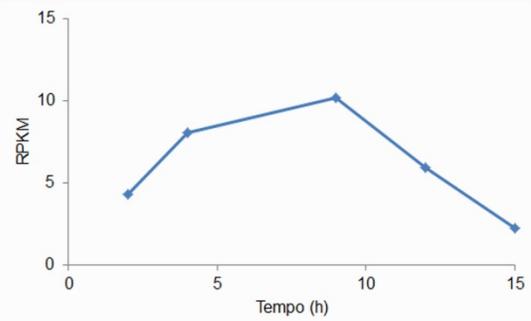
Gene 5218



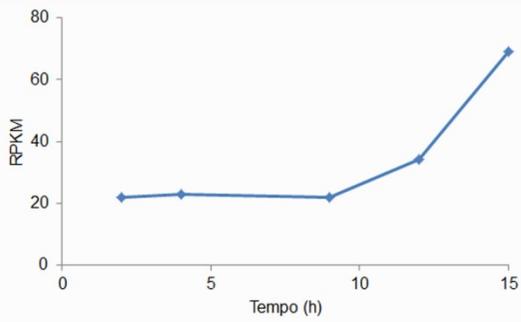
Gene 5219



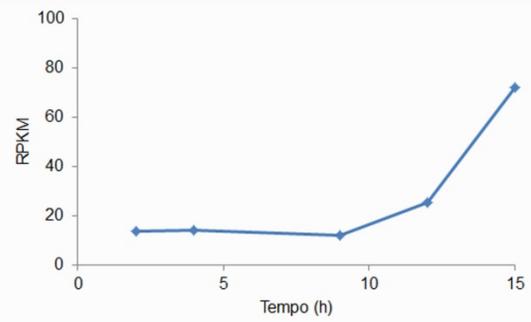
Gene 5220



Gene 5323



Gene 5324



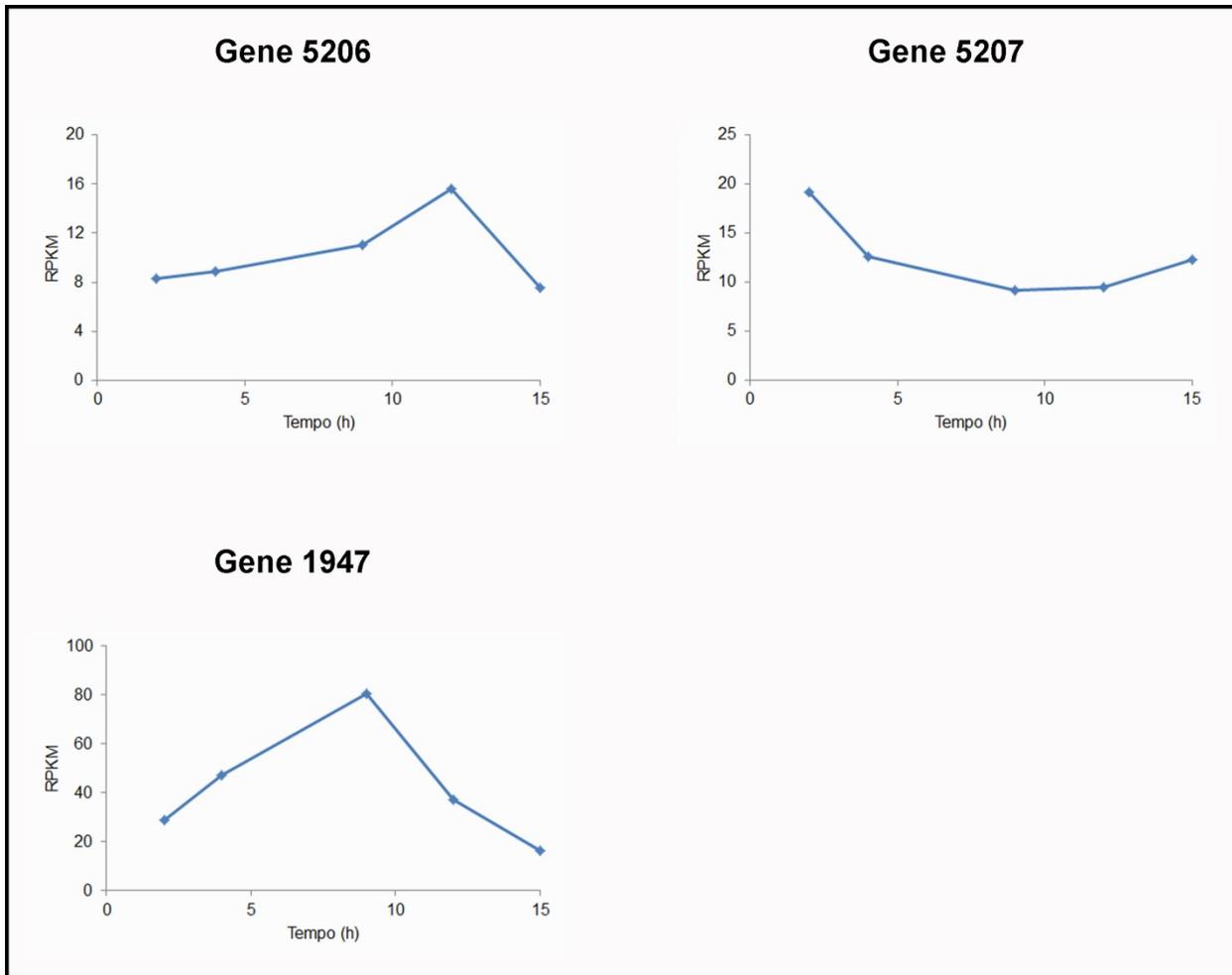


Figura 7. Dados de expressão gênica dos 21 genes novos da linhagem JAY291 ao longo de fermentação conduzida na Usina Nova América.

A análise da expressão gênica global de PE-2 nestas condições revelou que 88% dos genes da linhagem apresentam valor de RPKM entre 0 e 130 e o valor de expressão mediano para todos os genes da linhagem é 19 (Tese de doutorado – Osmar Vaz de Carvalho Netto “Fisiologia da fermentação alcoólica: análise da expressão gênica de uma linhagem industrial de *Saccharomyces cerevisiae* durante o processo fermentativo”). Portanto, a comparação com os valores de RPKM dos 21 genes novos de JAY291 revela que os mesmos apresentam valores de expressão próximos ao encontrado para a maioria dos genes da linhagem PE-2.

A expressão diferencial de alguns desses genes ao longo da fermentação dá suporte a hipótese de que os mesmos geram proteínas funcionais, potencialmente

relacionadas à excelente capacidade fermentativa da linhagem PE-2. É possível observar quatro diferentes perfis de expressão para estes genes. Os genes 23, 24, 25, 1396, 876, 5218, 5219, 5220 e 1947 apresentam um pico de expressão no ponto intermediário da fermentação industrial. Um aumento gradual na expressão, com decaimento no ponto final da fermentação, é observado para os genes 877, 878, 879, 880, 881, 883 e 5206. Os dois genes contidos no *cluster4* (5323 e 5324), apresentam um perfil similar, com uma expressão constante nos pontos iniciais da fermentação e um drástico aumento nos pontos finais. Os genes C1Q_05653 e 882 também apresentaram um aumento gradual no valor de expressão ao longo do processo fermentativo, porém, a expressão inicial destes genes já se apresenta elevada desde o início da fermentação. O gene 5207 apresentou um perfil não observado para outros genes, exibindo uma redução da expressão na fase intermediária do processo. Esses perfis de expressão poderiam justificar funções essenciais para uma fermentação industrial, como tolerância a estresses, que estão mais acentuados na fase final da fermentação, ou a funções importantes durante todo processo e relacionadas ao desempenho fermentativo e robustez da linhagem. Dessa forma, o conjunto desses dados deu suporte ao interesse no estudo funcional de alguns desses genes específicos da linhagem industrial JAY291/PE-2.

4.1.4 Screening do efeito da deleção dos genes novos em linhagens diplóides

Inicialmente, uma coleção de linhagens *knockout*, derivadas de JAY270, com a deleção dos 21 genes novos em estudo foi construída. Considerando-se que cada gene estaria presente em duas cópias no genoma da linhagem diplóide, realizou-se a deleção de duas cópias de cada um dos genes isolados (5206, 5208 e 1947) e dos *clusters* inteiros (*cluster1*, *cluster2*, *cluster3* e *cluster4*), através da substituição dos mesmos pelos genes de resistência a geneticina e higromicina. A deleção destes genes nas linhagens diplóides, derivadas de JAY270, foi confirmada através de seleção em meio com antibióticos e também através de reações de PCR, utilizando-se *primers* que anelam nas regiões *upstream* aos genes de interesse e na região interna dos genes de resistência.

Um primeiro *screening* do efeito da deleção destes genes foi obtido por meio de ensaios de competição direta contra a linhagem parental JAY270. Este ensaio é efetivo na detecção de variações fenotípicas de uma linhagem em relação a outra, pois essas diferenças se acumulam durante os ciclos fermentativos, que apresentam condições de estresse elevado. A tendência geral da variação das populações de células mutadas durante os ciclos fermentativos pode ser observada na Figura 8.

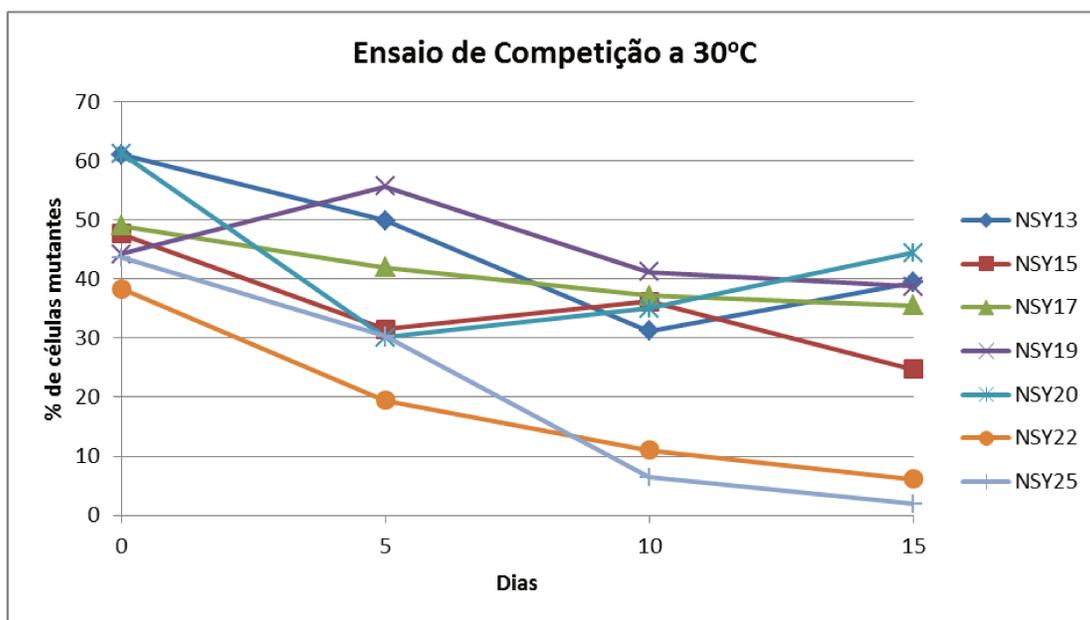


Figura 8. Variação na população de células das 7 linhagens *knockout* derivadas do diploide JAY270, durante 15 ciclos de competição a 30°C. O ensaio foi conduzido em xarope de cana-de-açúcar (18° BRIX, contendo 0.01% de extrato de levedura). No ciclo inicial, são misturadas células das linhagens *knockout* e selvagem (JAY270) visando uma proporção aproximada de 1:1. A população de cada tipo celular é monitorada ao longo dos ciclos de acordo com os métodos estabelecidos na seção 3.13. Identificação das linhagens: NSY13 (deleção do *cluster2*); NSY15 (deleção do *cluster1*); NSY17 (deleção do gene 5206); NSY19 (deleção do *cluster4*); NSY20 (deleção do gene 5207); NSY22 (deleção do *cluster3*); NSY25 (deleção do gene 1947).

Este ensaio forneceu indícios de que a deleção do gene 1947 e do *cluster3* afetou, respectivamente, o fenótipo das linhagens NSY25 e NSY22. Como pode ser observado na Figura 8, as linhagens mutantes para estes genes apresentaram uma drástica queda na população celular ao longo dos ciclos de fermentação, sugerindo uma menor persistência e robustez das mesmas com relação à parental JAY270. A partir desses resultados, ensaios semelhantes foram conduzidos com réplicas

biológicas para estas linhagens, visando a confirmação dos fenótipos observados. No entanto, o experimento realizado com a réplica biológica da linhagem NSY25 gerou resultados inconsistentes com o primeiro ensaio, ou seja, o fenótipo observado no teste inicial para a linhagem *knockout* para o gene 1947 não foi observado em sua réplica biológica, que foi gerada de forma independente. Esse resultado sugeriu que o efeito observado inicialmente para a linhagem NSY25 não estava correlacionado com a deleção do gene 1947. É possível que durante o processo de transformação, ou durante os ciclos de fermentação, algum evento de rearranjo cromossômico tenha ocorrido, culminando com a perda de heteroziguidade ou com a inativação de alguma região importante do genoma, gerando o efeito observado. O mesmo poderia ter ocorrido para a linhagem com deleção do *cluster3*.

Diante desses dados, foi feito um remodelamento do projeto. A fim de se diminuir as possibilidades de mascaramento do efeito da deleção dos genes em estudo, devido a mecanismos de rearranjos ou outros fatores, optou-se pela seleção de um número reduzido de genes para avaliação fenotípica de linhagens *knockouts* derivadas do haploide JAY291. A avaliação do efeito da deleção desses genes na linhagem haploide foi considerada mais robusta, por esta apresentar apenas um conjunto cromossômico.

Adicionalmente, os dados obtidos preliminarmente evidenciaram a necessidade de uma investigação mais criteriosa desses genes antes do início da construção das linhagens mutantes. É comum no genoma da linhagem PE-2, a presença de genes com uma amplificação no número de cópias ou de genes em hemizigose, o que dificulta a construção de linhagens *knockout* (Argueso et al. 2009). Por este motivo, realizou-se uma estimativa do número de cópias dos 21 genes novos no genoma das linhagens JAY270 e seus derivados haploides para a escolha dos genes que seriam estudados em maior profundidade. O cálculo foi baseado na contagem do número total de *reads* alinhados por par de bases de cada gene, tomando-se como referência um segmento do cromossomo IX sabidamente presente em duas cópias no genoma da linhagem JAY270 e em uma cópia no genoma do haploide JAY291 (Tabela 7).

Tabela 7. Estimativa do número de cópias dos genes novos de JAY291 baseada no total de *reads* por pb em um segmento do cromossomo IX.

Grupo ID	Gene ID	Comprimento de referência (pb)	JAY270	JAY289	JAY290	JAY291	JAY292
Sequência de referência	Chr09 50 to 80.5 kb	30522	2.00	1.00	1.00	1.00	1.00
Grupo 1	23	1512	3.66	1.65	2.41	1.88	2.22
	24	3864	3.14	1.57	2.44	1.78	2.38
	25	1962	3.27	1.56	2.14	1.81	2.04
	1396	1182	3.23	1.98	2.24	1.71	2.71
	C1Q_05653	1281	3.36	1.63	2.28	1.85	2.47
Grupo 2	876	1092	1.03	0.00	1.13	0.97	0.01
	877	1602	1.03	0.00	0.94	0.95	0.01
	878	1935	0.81	0.00	0.91	0.93	0.00
	879	1641	1.07	0.01	0.90	0.93	0.01
	880	867	0.97	0.00	1.09	0.97	0.01
	881	1836	0.79	0.00	1.13	0.91	0.00
	882	858	2.84	0.55	1.74	1.90	0.97
	883	1416	1.07	0.06	1.30	1.02	0.07
Grupo 3	5218	1911	1.03	0.00	1.01	0.92	0.01
	5219	1671	0.83	0.00	0.84	0.89	0.01
	5220	1635	1.02	0.00	0.94	0.92	0.00
Grupo 4	5323	1239	1.58	0.96	1.15	0.88	0.92
	5324	801	1.64	0.62	0.65	0.87	0.52
-	1947	1119	1.80	0.97	1.31	0.86	1.53
-	5206	363	3.18	2.79	1.19	1.23	2.83
-	5207	690	1.53	0.59	1.00	0.97	0.72

Como é possível observar na tabela acima, os genes em estudo podem estar presentes em até três cópias no genoma da linhagem diplóide JAY270. A amplificação no número de cópias de alguns desses genes pode ser um indício de que as condições encontradas durante a fermentação industrial favorecem a retenção dos mesmos. Esta análise revelou também que alguns genes apresentam apenas uma cópia no genoma da linhagem diplóide. Não está claro como foi possível a confirmação da deleção, através de reações de PCR, de duas cópias desses genes nas linhagens mutantes construídas inicialmente a partir do diploide selvagem JAY270. A inconsistência entre esses resultados e a incerteza com relação ao genótipo dessas linhagens nos levou a

suspender o uso das mesmas no estudo funcional destes genes, utilizando-se para este fim novas linhagens derivadas do haploide JAY291.

Apesar de *a priori* todos os 21 genes em estudo apresentarem potencial para desempenhar funções importantes para a linhagem PE-2 em condições industriais, seria inviável uma nova avaliação do efeito dos 21 genes em linhagens haploides. Além disso, o procedimento experimental para construção de linhagens *knockout* empregado neste trabalho é facilitado para genes presentes em no máximo duas cópias no genoma. Desse modo, com base nos resultados obtidos no ensaio de competição detalhado na Figura 8, optou-se por averiguar se a deleção do *cluster3* e do gene 1947 poderia, de fato, afetar o fenótipo da linhagem. Portanto, estes genes que apresentam apenas uma cópia no isolado haploide JAY291 (ver Tabela 7) foram escolhidos para análise em maior detalhe ao longo do presente trabalho de mestrado. Os testes realizados inicialmente com diplóides tornaram clara a necessidade de uma maior caracterização e conhecimento dos genes envolvidos neste tipo de estudo, portanto, antes de iniciar a construção de novas linhagens mutantes, realizou-se a confirmação da presença e localização dos genes selecionados no genoma da linhagem JAY291, confirmação da expressão dos mesmos durante a fermentação em condições de laboratório e análise de similaridade de sequência com outras linhagens.

4.2 Caracterização funcional de quatro genes novos de JAY291

4.2.1 Caracterização geral do *cluster3*

A fim de analisar a conservação dos genes 5218, 5219 e 5220 em diferentes linhagens de *S. cerevisiae* realizou-se uma busca de similaridade dessas sequências contra todas as linhagens com sequências depositadas no banco de dados “*Saccharomyces genome database*” (SGD – www.yeastgenome.org).

Esta análise revelou que os genes contidos no *cluster3* estão conservados principalmente em linhagens isoladas de ambientes naturais, como a EC9-8, PW-5, T7, Y10 e a YJM269. Interessantemente, de forma similar ao observado em PE-2, notou-se que estes genes estão conservados de forma agrupada também para estas linhagens,

perfazendo uma sequência total de 8751pb em JAY291. A presença desses genes também foi confirmada para a linhagem industrial de bioetanol BG-1, através da amplificação dos mesmos por PCR utilizando-se *primers* que flanqueiam cada um dos genes do *cluster*. Este resultado sugere que estes genes podem estar associados com características importantes para sobrevivência de linhagens expostas a situações de competição com outros micro-organismos, como as condições encontradas em ambientes naturais e em dornas de produção de bioetanol. Isso é reforçado pela ausência desses genes em todas as linhagens laboratoriais com sequências depositadas no banco de dados do SGD, as quais são cultivadas constantemente em condições ideais de crescimento.

Além disso, as análises realizadas por meio da ferramenta BLASTx (Altschul et al. 1997), que podem ser visualizadas na Tabela 6, fornecem mais um indício de que estes genes poderiam ser importantes para a performance da linhagem PE-2 durante a fermentação. Observou-se que o gene 5219 apresenta similaridade com sequências de permeases e, portanto, pode estar associado ao transporte de diferentes moléculas para dentro da célula. O gene 5218 apresenta similaridade com sequências de sideróforos, que são quelantes específicos para íons de ferro (Fe^{+3}). Essas moléculas são secretadas por diversos organismos com a finalidade de mobilizar íons de ferro para dentro da célula e estoca-los de forma atóxica (Hider and Kong 2010). A levedura *S. cerevisiae* não é capaz de sintetizar sideróforos, mas é capaz de utilizar os produzidos por outras espécies, para isso, ela apresenta genes relacionados à retenção dos sideróforos junto à parede celular e genes responsáveis pela redução do íon férrico a ferroso, que é transportado para dentro da célula (Haas 2003). Sabe-se que o ferro é essencial para o crescimento e proliferação da maioria dos micro-organismos, dessa forma, o sucesso na obtenção deste nutriente poderia estar relacionado à melhor competitividade da linhagem PE-2 em dornas de fermentação. O gene 5220 apresenta similaridade com proteínas de função desconhecida, o que torna seu estudo interessante do ponto de vista científico, já que o mesmo é potencialmente importante para o fenótipo da linhagem, porém ainda não foi descrito. Visando uma melhor elucidação do papel desses genes, linhagens derivadas de JAY291 *knockouts*

para toda a sequência do *cluster3* foram construídas e avaliadas em diferentes testes fenotípicos.

Antes de dar início à construção de mutantes knockout para o *cluster3* confirmou-se a localização dos genes 5218, 5219 e 5220 na linhagem JAY291, através da amplificação da sequência dos mesmos em reações de PCR utilizando-se primers que anelam nas regiões flanqueadoras de cada gene (Figura 9, C). A posição relativa dos genes dentro do cluster também foi verificada por meio de um conjunto de reações de PCR utilizando primers que anelam em diferentes regiões de cada gene. As regiões amplificadas são detalhadas na Figura 9 (A, B).

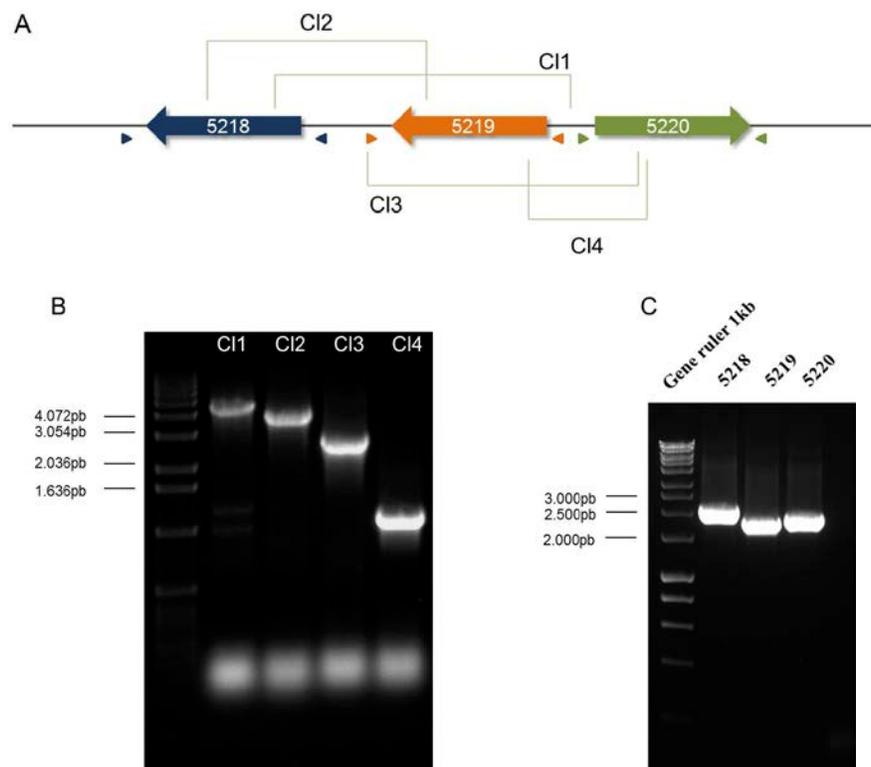


Figura 9. Confirmação da localização dos genes 5218, 5219 e 5220, que constituem o *cluster 3*, na linhagem JAY291. A) Esquema ilustrando as regiões do *cluster3* que foram amplificadas para confirmar a localização, posição relativa e tamanho dos genes. B) Gel de eletroforese confirmando a posição relativa dos genes do *cluster3* na ordem 5218, 5219 e 5220. A identificação de cada fragmento mostrado no gel está detalhada na parte A desta figura. C) Gel de eletroforese confirmando a localização e tamanho dos genes do *cluster3*. Um esquema dos primers utilizados para amplificar cada gene está representado por cabeças de seta mostrado na parte A desta figura.

Antes do início dos testes funcionais em laboratório, foram avaliadas as expressões dos genes contidos no *cluster3* sob condições de laboratório, por meio da técnica de *Real-Time* PCR. Os dados obtidos podem ser visualizados na Figura 10.

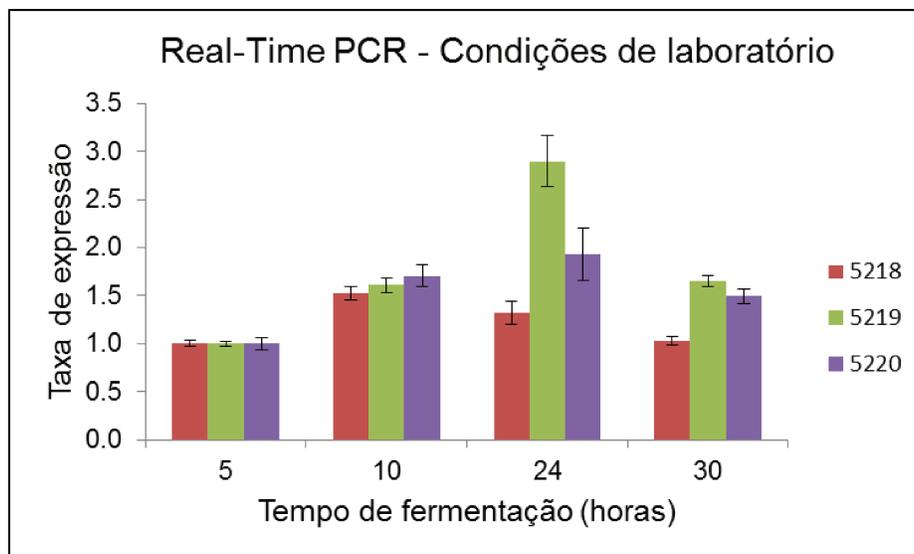


Figura 10. Avaliação da expressão dos genes 5218, 5219 e 5220, por meio da técnica de *Real-Time* PCR. As amostras para extração de RNA foram coletadas nos pontos 5h, 10h, 24h e 30h de um ensaio de fermentação da linhagem JAY270. A fermentação foi conduzida a 30°C em xarope de cana-de-açúcar (8° BRIX, 0,01% de extrato de levedura), a uma agitação de 80rpm.

A variação da expressão desses genes em condições de laboratório foi inferior em comparação aos valores observados em condições industriais, nos experimentos de RNA-Seq (Figura 7). Para as condições de laboratório, as variações entre os pontos de maior e menor expressão para os genes 5218, 5219 e 5220 foram, respectivamente, 1.52, 2.89 e 1.92. As variações encontradas para estes nas condições industriais foram, respectivamente, 3.35, 3.03 e 4.44. Essa diferença pode ser resultado das variações decorrentes dos diferentes tipos de técnicas e condições empregadas para avaliação da expressão gênica. No entanto, o fato da expressão desses genes sofrer uma maior variação sob condições industriais suporta a hipótese de que estes genes estão relacionados e tem importância nessas condições.

4.2.2 Caracterização geral do gene 1947

A distribuição do gene 1947 entre as diferentes linhagens de *S. cerevisiae* também foi avaliada através da busca deste gene no genoma das linhagens com sequências depositadas banco de dados do SGD. A sequência deste gene está

conservada em outras duas linhagens industriais utilizadas na produção de cerveja, denominadas FosterO e FosterB. Além disso, a presença deste gene foi verificada por amplificação por PCR nas linhagens BG-1, CAT-1 e SA-1, três linhagens industriais brasileiras utilizadas para produção de bioetanol. O fato de este grupo de linhagens, empregadas em processos industriais de história relativamente recente, compartilharem este gene sugere que o mesmo pode estar envolvido em mecanismos relacionados à elevada performance fermentativa, ou ainda, à sobrevivência a situações de competição e/ou estresse celular encontradas no ambiente industrial. Para todas as linhagens que apresentam o gene 1947 foi possível identificar que o mesmo está localizado em um mesmo locus no cromossomo XVI, entre os genes *THI21* e *CLN2*. Este fato sugere que este gene foi perdido por outras linhagens durante o processo evolutivo da espécie *S. cerevisiae*, já que diferentes eventos de inserção de um gene em um mesmo ponto do genoma são extremamente improváveis.

Apesar de a conservação do gene 1947 entre diferentes linhagens industriais sugerir um possível envolvimento do mesmo com funções importantes para sobrevivência em condições de estresse, o mesmo não apresenta similaridade com outros domínios ou proteínas de função conhecida (Tabela 6), o que evidencia a necessidade de mais estudos visando a caracterização do mesmo.

A distribuição do gene 1947 na espécie *S. cerevisiae* gerou o interesse na investigação de evidências do processo evolutivo de perda deste gene entre as diferentes linhagens da espécie. Com este intuito, o locus em que o gene 1947 se insere foi analisado, e observou-se que a sequência imediatamente após a região terminadora do gene está incompleta no genoma da linhagem JAY291 depositado no banco de dados do NCBI.

Na busca de um indício do tamanho da região desconhecida pelo genoma público, amplificou-se a sequência presente entre o fim do *contig* que contém o gene 1947 e o *contig* seguinte utilizando-se *primers* que anelam na região terminal do gene 1947 e na região anterior ao gene *THI21*, que se situa adjacente ao gene 1947. Foi verificado que o locus em que este gene se insere é heterozigoto na linhagem JAY270, o que é evidenciado pelos diferentes padrões de bandas encontrados para esta linhagem (Figura 11).

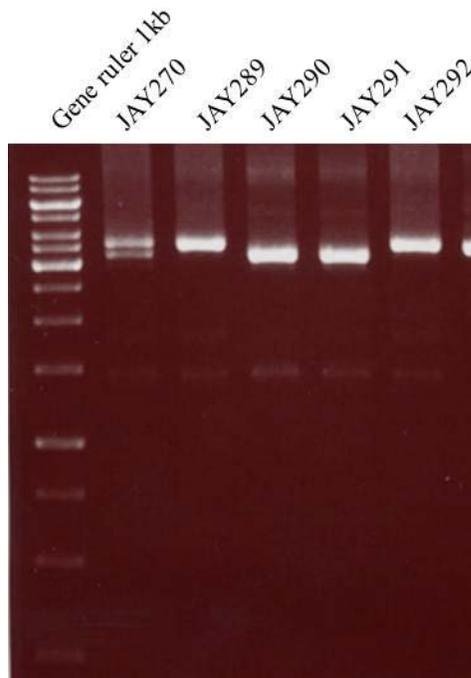


Figura 11. Gel de eletroforese confirmando a heterozigosidade da região imediatamente posterior ao gene 1947 na linhagem JAY270. A imagem evidencia também que esta região difere em tamanho entre diferentes linhagens da espécie *S. cerevisiae*.

Com o objetivo de investigar essa diferença estrutural no locus do gene 1947, realizou-se o sequenciamento no modelo *primer walking* do *gap* existente entre o *contig* que contém este gene e o *contig* adjacente. Foi sequenciado um total de 3.137 pares de bases do genoma da linhagem JAY291 e 3.365 pares de bases do genoma da linhagem JAY292. O alinhamento das sequências derivadas da montagem final dos *reads* de sequenciamento permitiu a comprovação de um sítio heterozigoto entre os esporos de JAY270. Essa região consiste em 343pb presentes no genoma da linhagem JAY292, mas ausentes em JAY291 (Figura 12). Uma análise de similaridade desta sequência de 343pb contra o genoma da linhagem referência S288c foi realizada utilizando-se a ferramenta BLASTn do banco de dados SGD. Observou-se que esta sequência apresenta 98% de similaridade com um elemento denominado *delta*. Na Figura 12, também é possível observar parte das sequências dos genes *CLN2* e *THI21*, assim como regiões repetitivas idênticas de 200pb, anotadas como r1, comuns às duas linhagens. As sequências r1 também apresentam homologia com elementos *delta*.

Estes elementos constituem longas repetições terminais que flanqueiam os elementos de transposição das classes Ty1 e Ty2. A presença desses elementos repetitivos neste locus sugere uma grande plasticidade na região flanqueadora ao gene 1947, que pode estar relacionada aos eventos de perda desse gene entre diversas linhagens de *S. cerevisiae*. A presença de elementos delta isolados é comum no genoma de *S. cerevisiae* e sugere que os mesmos foram derivados de eventos de recombinação entre elementos *delta* que flanqueiam elementos Ty completos (Mieczkowski, Lemoine, and Petes 2006).



Figura 12. Alinhamento entre as sequências correspondentes ao locus do gene 1947 das linhagens JAY291 e JAY292. Na sequência consenso é possível observar regiões com sequências idênticas entre as duas linhagens em verde. Na montagem final de cada linhagem, a sequência de nucleotídeos está representada em preto e a anotação de parte do gene *CLN2*, gene 1947 e parte do gene *THI21* estão representadas em verde. A seta em vermelho aponta para a região de 342pb presente no genoma da linhagem JAY292, mas ausente no genoma de JAY291. As setas com anotação *r1* representam regiões repetitivas idênticas entre as linhagens.

A expressão do gene 1947 também foi confirmada sob condições de laboratório, por meio da técnica de *Real-Time* PCR. Os dados obtidos podem ser visualizados na Figura 13.

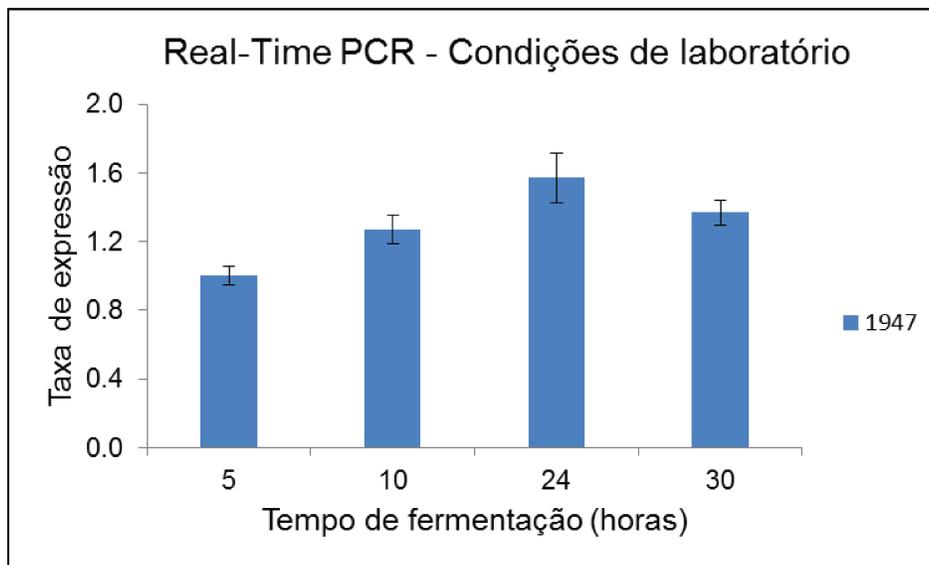


Figura 13. Avaliação da expressão do gene 1947, por meio da técnica de *Real-Time* PCR. As amostras para extração de RNA foram coletadas nos pontos 5h, 10h, 24h e 30h de um ensaio de fermentação da linhagem JAY270. A fermentação foi conduzida a 30°C em xarope de cana-de-açúcar (8° BRIX, 0,01% de extrato de levedura), a uma agitação de 80rpm.

Como pode ser observado na Figura 13, o perfil de expressão do gene 1947 obtido através da técnica de *Real-Time* PCR é similar ao perfil observado nos experimentos de *RNA-Seq* (Figura 7), comprovando que o mesmo apresenta variação de sua expressão durante o processo fermentativo e fornecendo um maior suporte a um possível papel deste gene para o fenótipo fermentativo da linhagem.

Adicionalmente, o gene 1947 foi caracterizado quanto à posição relativa de suas cópias no genoma da linhagem diploide JAY270. A Figura 14 apresenta um exemplo do padrão fenotípico de resistência aos antibióticos geneticina e higromicina observado nos esporos derivados de tétrades da linhagem que contém a deleção de duas cópias do gene 1947. Nenhum dos esporos avaliados apresentou resistência a ambos antibióticos. Igualmente, nenhum esporo apresentou-se sensível a mais de um

antibiótico, o que permite inferir que as cópias deste gene estão inseridas de forma alélica, ou ligada, no genoma da linhagem diplóide.

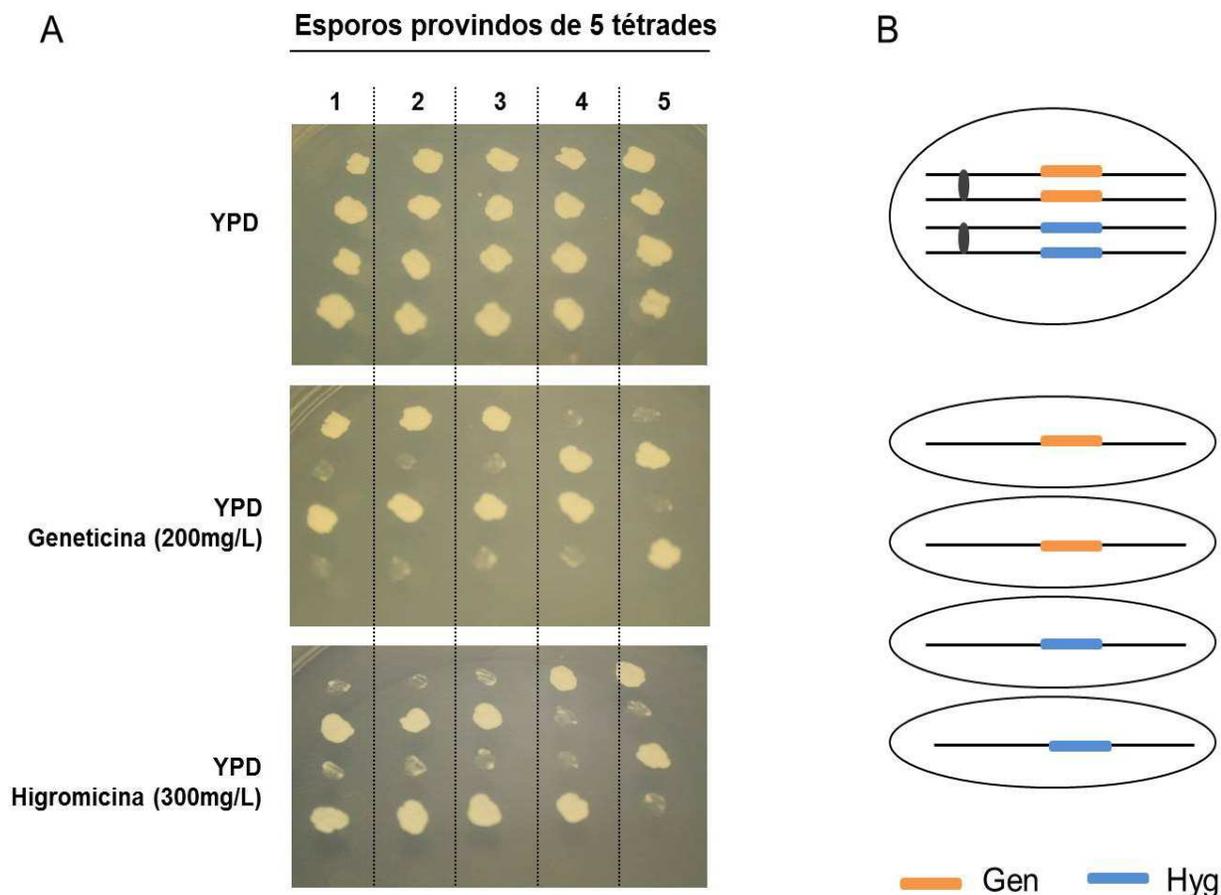


Figura 14. Ensaio para avaliar segregação do gene 1947. A) Esporos providos de 5 tétrades da linhagem NSY43 foram dissecados e plaqueados em meio de cultivo YPD e no mesmo meio de cultura suplementado com geneticina e higromicina. B) Esquema proposto para posicionamento do gene 1947 no genoma da linhagem JAY270, evidenciando as inserções alélicas dos genes *KanMx4* (Gen) e *Hph* (Hyg), previamente inseridos no locus do gene 1947.

Foram avaliados esporos de um total de 14 tétrades derivadas da linhagem que apresenta deleção de duas cópias do gene 1947. O padrão de resistência aos antibióticos geneticina e higromicina dos esporos indica que não houve recombinação entre os genes *KanMx4* e *Hph* em nenhum dos esporos analisados. Isto sugere que os genes *KanMx4* e *Hph*, que substituíram as duas cópias do gene 1947 na linhagem, são provavelmente inserções alélicas.

4.2.3 Obtenção de linhagens mutantes

Visando uma avaliação do efeito dos genes contidos no *cluster3* e do gene 1947 no fenótipo da linhagem PE-2, foram realizados ensaios de deleção dos mesmos no organismo haploide JAY291. Adicionalmente, foi realizada uma nova deleção das duas cópias do gene 1947 na linhagem diplóide JAY270. Como a linhagem PE-2 apresenta uma elevada eficiência fermentativa e alta resistência a fatores de estresse, a identificação de interferências nessas características causadas pela deleção destes genes poderia ser imperceptível. Por esse motivo, como uma forma adicional de avaliar os potenciais efeitos relacionados a este gene, realizou-se a inserção de uma cópia do mesmo no genoma das linhagens S288c e RM11-1a, a linhagem referência utilizada para estudos em laboratório e uma linhagem isolada de campos de vinho, respectivamente.

Como estratégia para seleção dos transformantes contendo a inserção do gene 1947, foi realizada a ligação do mesmo ao gene de resistência a geneticina, *KanMx4*, através de uma reação de *overlapping* PCR (Ho et al. 1989). A confirmação da construção dos cassetes pode ser observada na Figura 15.

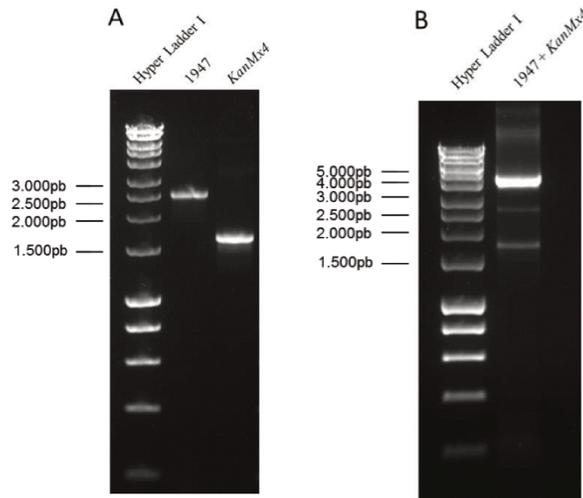


Figura 15. Fragmentos amplificados nas reações de *overlapping* PCR. A) Fragmentos correspondentes ao gene 1947, correspondente a banda de 2454pb, e *KanMx4*, correspondente a banda de 1631pb, com suas respectivas caudas de homologia. B) Fragmento de 4049pb obtido após a reação de junção do gene 1947 e *KanMx4*.

A confirmação da inserção da sequência correta do gene 1947 nas linhagens RM11-1a e S288c foi realizada através de sequenciamento. Não foi observada nenhuma mutação na sequência do gene inserido nas cepas construídas. A identificação geral das linhagens geradas neste trabalho pode ser observada na Tabela 8.

Tabela 8. Identificação e genótipo das linhagens produzidas por deleção ou inserção dos genes em estudo em diferentes linhagens de *S. cerevisiae* (colocar todas as demais linhagens).

ID da linhagem	Genótipo	Descrição
NSY24	JAY270 1947/1947Δ:: <i>KanMX4</i>	JAY270, 1 cópia 1947 deletada
NSY27	JAY291 Cluster3Δ:: <i>KanMX4</i>	JAY291, 1 cópia <i>cluster3</i> deletada
NSY28	JAY291 Cluster3Δ:: <i>KanMX4</i>	JAY291, 1 cópia <i>cluster3</i> deletada
NSY29	JAY291 1947Δ:: <i>KanMX4</i>	JAY291, 1 cópia 1947 deletada
NSY30	JAY291 1947Δ:: <i>KanMX4</i>	JAY291, 1 cópia 1947 deletada
NSY35	RM11-1a:: <i>KanMX4</i> ; 1947	RM11-1a, inserção do gene 1947
NSY36	RM11-1a:: <i>KanMX4</i> ; 1947	RM11-1a, inserção do gene 1947
NSY37	S288c:: <i>KanMX4</i> ; 1947	S288c, inserção do gene 1947
NSY38	S288c:: <i>KanMX4</i> ; 1947	S288c, inserção do gene 1947
NSY39	JAY291:: <i>KanMX4</i>	Controle: JAY291- inserção <i>KanMx4</i>
NSY40	JAY291:: <i>KanMX4</i>	Controle: JAY291- inserção <i>KanMx4</i>
NSY43	NSY24 1947Δ:: <i>Hph</i> /1947Δ:: <i>KanMX4</i>	JAY270, 2 cópias 1947 deletadas
NSY45	JAY270 1947:: <i>KanMX4</i>	Controle: JAY270 – inserção <i>KanMx4</i>
NSY46	JAY270 1947:: <i>KanMX4</i>	Controle: JAY270 – inserção <i>KanMx4</i>
NSY47	JAY270 1947/1947Δ:: <i>KanMX4</i>	JAY270, 1 cópia 1947 deletada
NSY48	JAY270 1947/1947Δ:: <i>KanMX4</i>	JAY270, 1 cópia 1947 deletada
NSY49	NSY45 :: <i>KanMx4</i> , <i>Hph</i>	Controle: JAY270 – inserção <i>KanMx4</i> e <i>Hph</i>
NSY50	NSY46 :: <i>KanMx4</i> , <i>Hph</i>	Controle: JAY270 – inserção <i>KanMx4</i> e <i>Hph</i>
NSY51	NSY47 1947Δ:: <i>Hph</i> /1947Δ:: <i>KanMX4</i>	JAY270, 2 cópias 1947 deletadas

4.2.4 Ensaio fenotípicos com as linhagens mutantes construídas

A partir da obtenção da coleção de linhagens mutantes para os genes de interesse e da caracterização geral dos mesmos iniciou-se uma busca por um fenótipo associado a estes genes. Em geral, a avaliação fenotípica de mutantes pode fornecer informações importantes acerca da função dos genes mutados. Esses mutantes podem capazes de crescer normalmente sob condições permissivas, mas podem se apresentar inviáveis ou com crescimento prejudicado sob circunstâncias restritivas. Os

fenótipos, geralmente, estão associados a uma maior sensibilidade à determinada condição, mas também podem se relacionar à maior resistência a outras condições (Hampsey 1997). Dessa forma, visando encontrar uma associação entre fenótipo e função dos genes em estudo, as linhagens mutantes e parentais selvagens foram submetidas a uma série de testes fenotípicos, detalhados nas sessões a seguir.

4.2.5 Avaliação do crescimento das linhagens *knockout*

Inicialmente, foi investigado se a deleção dos genes de interesse afetou a taxa de crescimento das linhagens mutantes. O crescimento das linhagens *knockout* e das parentais selvagens JAY270 e JAY291 foi avaliado através da aferência da densidade óptica (OD_{600}) das culturas ao longo de uma fermentação em meio YPD com 2% de glicose. Como pode ser observado na Figura 16, as linhagens mutantes construídas apresentaram uma curva de crescimento similar às linhagens selvagens, sugerindo que os genes 5218, 5219, 5220 e 1947 não afetam significativamente o crescimento da linhagem nas condições avaliadas.

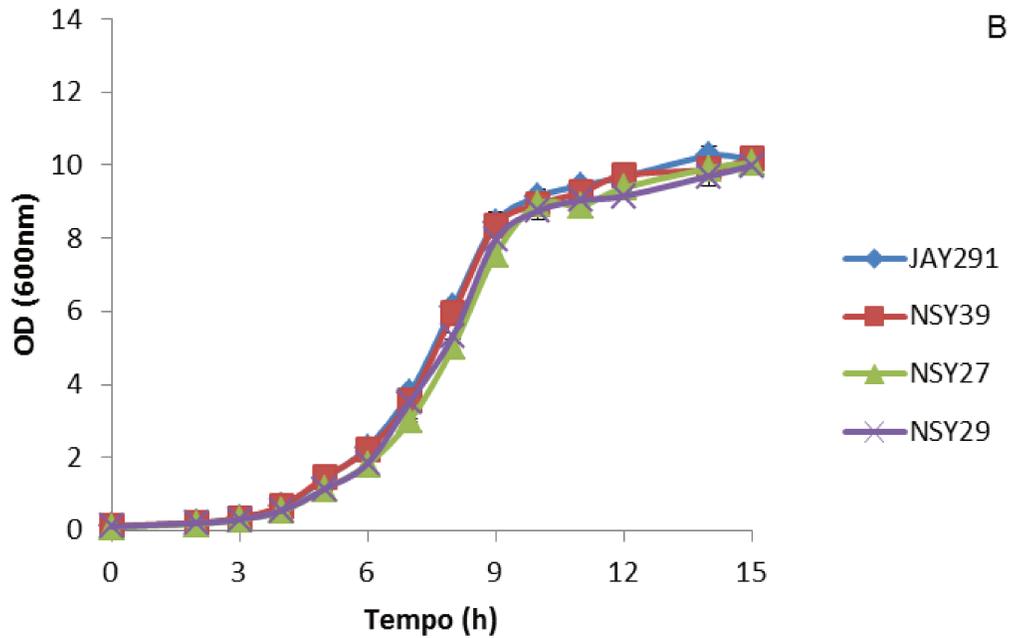
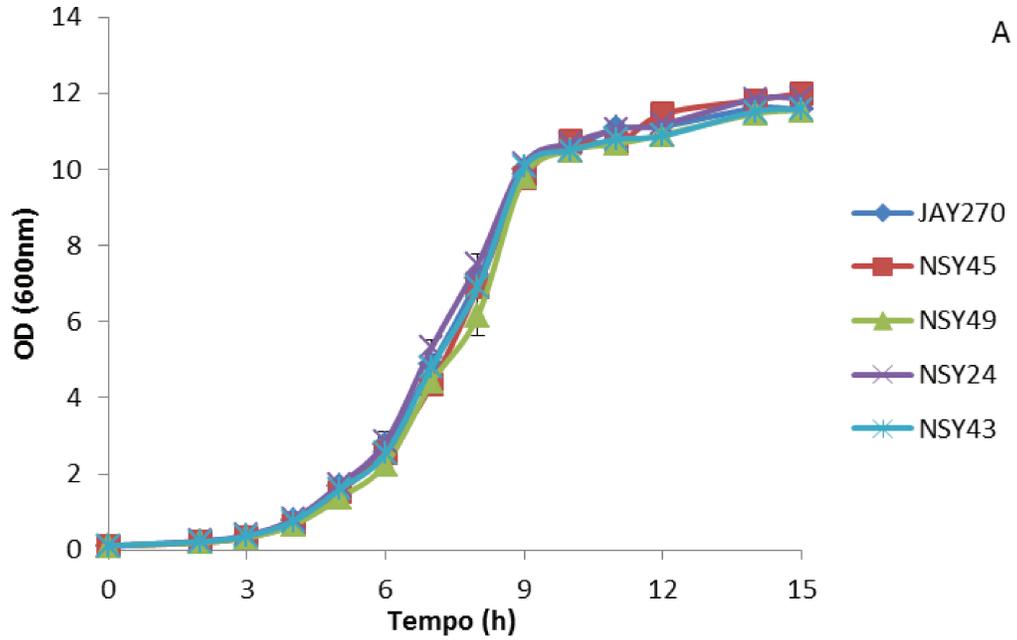


Figura 16. Curvas de crescimento das linhagens *knockout* construídas. A) Curvas de crescimento das linhagens diplóides. B) Curvas de crescimento das linhagens haplóides.

4.2.6 Ensaio de tolerância a estresses

Em condições industriais, um dos fatores capazes de gerar maior estresse às células é a alta concentração de etanol ao longo da fermentação (Basso et al. 2008; Argueso et al. 2009). Por este motivo, é constante a busca por genes relacionados à maior tolerância celular a este composto. As linhagens obtidas foram testadas quanto à tolerância a etanol, através do acompanhamento do crescimento das culturas em meio YPD, suplementado com 8% de etanol (Figura 17). Como é possível observar, em comparação com os dados apresentados na Figura 16, o crescimento das linhagens mutantes e selvagens sob alta concentração de etanol foi afetado tanto com relação ao valor de saturação final das culturas quanto à taxa de crescimento. No entanto, não foi notado um crescimento diferencial das linhagens mutantes em comparação às selvagens, que pudesse estar relacionado à deleção dos genes em estudo. Esse resultado indica que estes genes não contribuiriam para a proteção das células contra o efeito citotóxico do etanol.

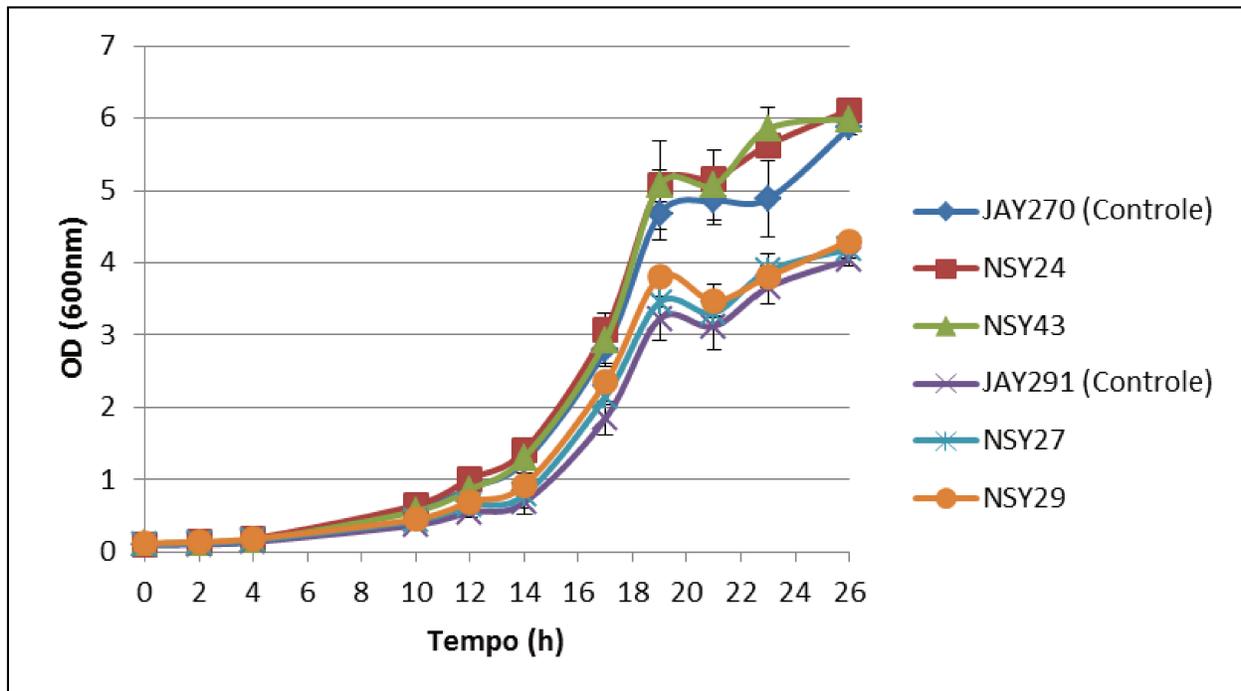


Figura 17. Curvas de crescimento em meio YPD suplementado com 8% de etanol.

Os genes em estudo também foram analisados quanto a uma possível função relacionada com a tolerância à elevadas concentrações osmóticas. Para este objetivo, foi comparado o crescimento em colônia das linhagens em meio sólido contendo 2%, 10% e 18% de glicose. Como uma forma adicional da avaliação do efeito do estresse osmótico sobre o crescimento das linhagens, estas também foram submetidas ao crescimento em meio sólido suplementado com NaCl, na concentração final de 1M.

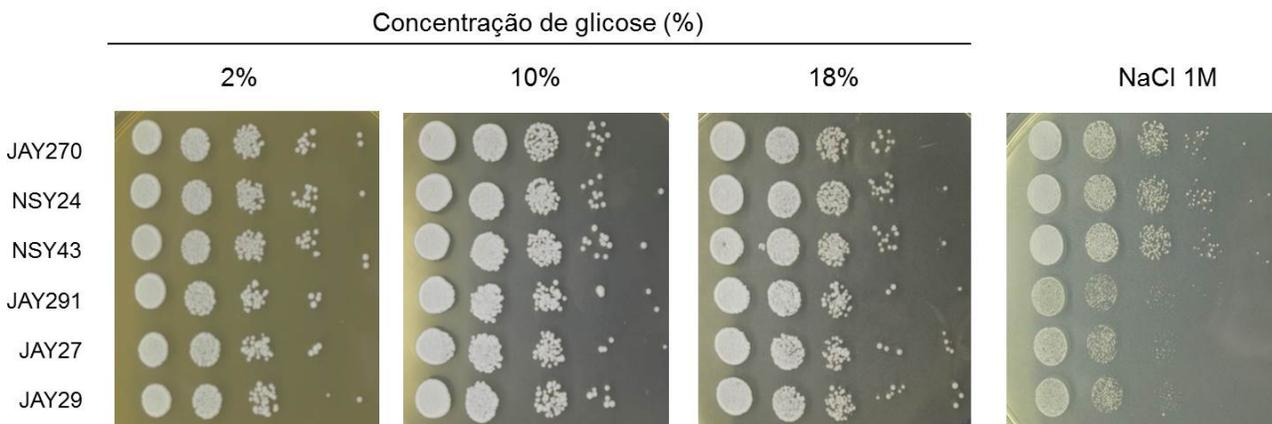


Figura 18. Análise fenotípica comparativa do efeito do estresse osmótico no crescimento de colônias das linhagens *knockout* produzidas e suas respectivas linhagens parentais. Diluições seriadas de 10 vezes de culturas saturadas foram plaqueadas em meio YPD contendo diferentes concentrações de glicose. O crescimento também foi avaliado em meio YPD suplementado com NaCl 1M.

A Figura 18 evidencia que não há correlação aparente entre a deleção dos genes em estudo e a tolerância ao estresse osmótico. Isto é comprovado pela comparação do número e do crescimento das colônias das linhagens mutantes, em relação as suas respectivas parentais selvagens.

O efeito de elevadas temperaturas foi avaliado em um ensaio semelhante ao conduzido na investigação dos efeitos do estresse osmótico. Neste caso, diluições seriadas de culturas de todas as linhagens *knockout* e das selvagens JAY270 e JAY291 foram realizadas em meio YPD 2% de glicose e, então, as placas foram mantidas nas temperaturas 30°C, 37°C e 41°C. Para as condições testadas não foi possível relacionar os genes contidos no *cluster3* e o gene 1947 a um efeito protetor

aos danos causados por temperaturas elevadas. Como pode ser observado na Figura 19, as linhagens mutantes apresentaram um crescimento, avaliado pelo número de colônias, semelhante às linhagens selvagens em todas as temperaturas testadas.

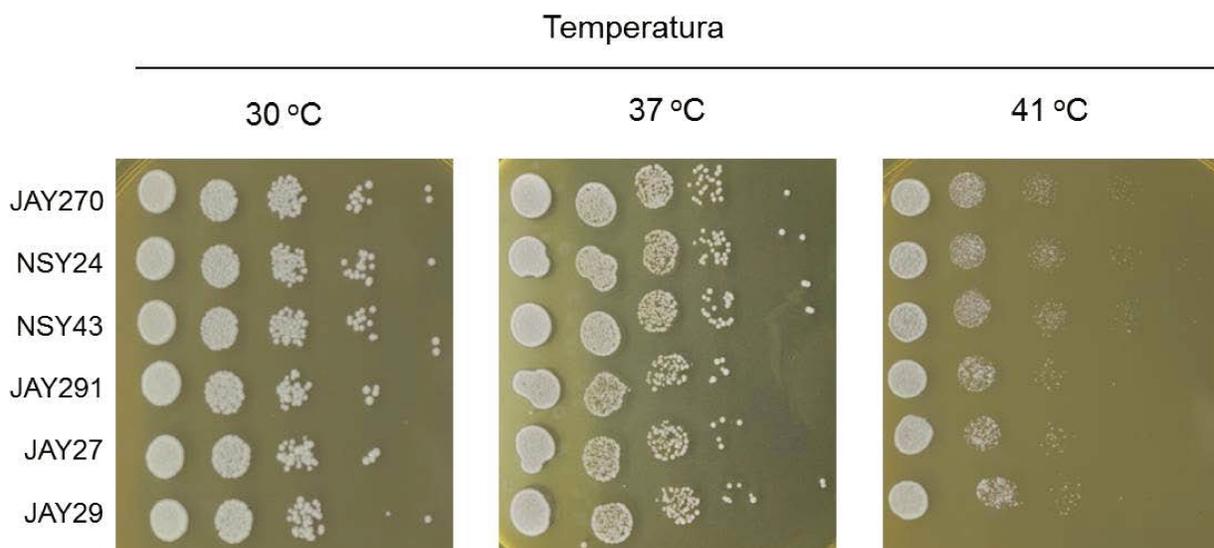


Figura 19. Análise fenotípica comparativa do efeito de altas temperaturas no crescimento de colônias das linhagens *knockout* produzidas e de suas respectivas linhagens parentais. Diluições seriadas de 10 vezes de culturas saturadas foram plaqueadas em meio YPD contendo 2% de glicose. Em seguida, as placas foram mantidas sob diferentes temperaturas.

4.2.7 Ensaio de competição

Visando avaliar a relação dos genes em estudo com a capacidade competitiva de PE-2 foram realizados ensaios de competição direta das linhagens mutantes contra suas respectivas linhagens parentais. Desvios na proporção inicial de células de 1:1 são observados quando a população de células de uma linhagem se sobrepõe à outra, indicando uma maior capacidade competitiva durante os ciclos de fermentação. Dessa forma, este ensaio é muito efetivo na detecção de variações na adaptação de uma linhagem em relação a outra, pois essas diferenças se acumulam durante os ciclos fermentativos, que apresentam condições de estresse elevado. A Figura 20 apresenta a tendência geral da variação da população de células das linhagens NSY39 (controle), NSY27 e NSY29, com relação à parental selvagem JAY291 durante os ciclos fermentativos.

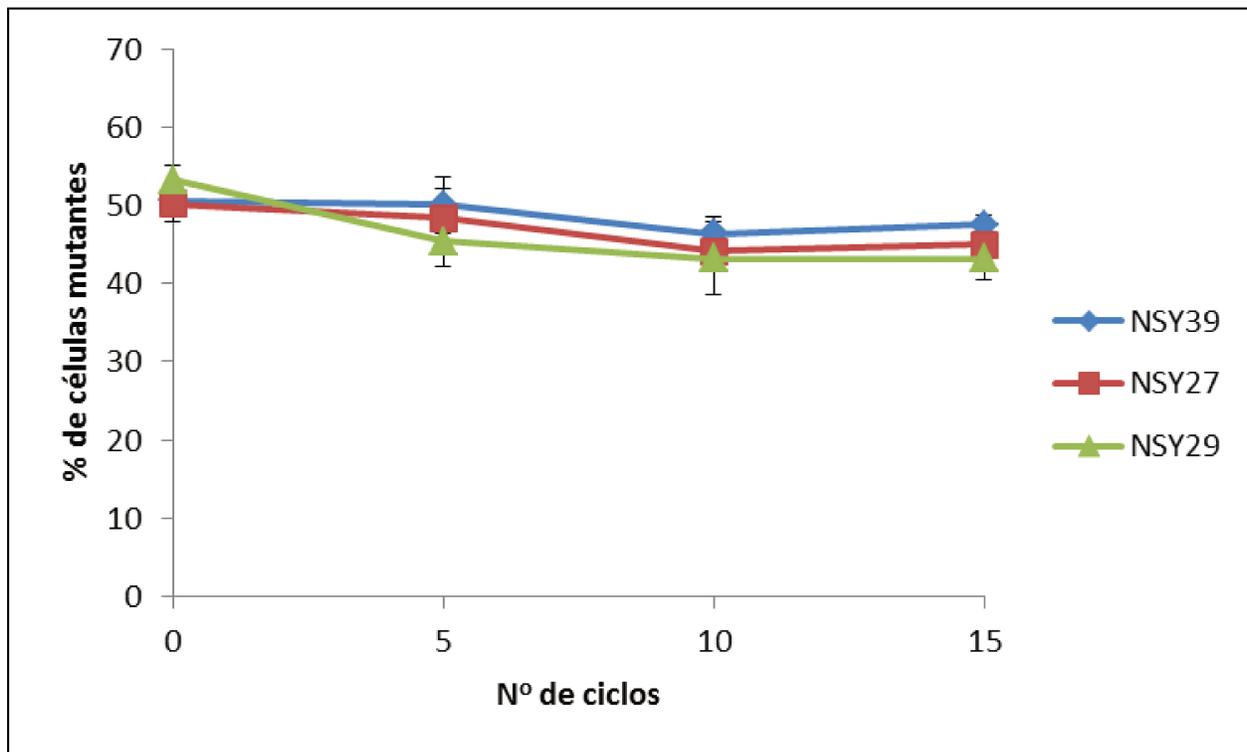


Figura 20. Variação na população de células das linhagens mutantes durante 15 ciclos de competição com a linhagem parental selvagem JAY291. O ensaio foi conduzido em xarope de cana-de-açúcar (18° BRIX, contendo 0.01% de extrato de levedura). No ciclo inicial, são misturadas células das linhagens haploides *knockout*, contendo a deleção do *cluster3* (NSY27) e do gene 1947 (NSY29), e selvagem haploide (JAY291) visando uma proporção aproximada de 1:1. Foi incluída também uma competição controle entre a linhagem haploide selvagem JAY291 e uma linhagem derivada de JAY291 contendo a inserção do gene *KanMx4*. A população de cada tipo celular é monitorada ao longo dos ciclos de acordo com os métodos estabelecidos na seção 3.13.

Como se pode observar no gráfico acima, ao longo de 15 ciclos de competição não foram notadas diferenças significativas na proporção de células das linhagens *knockout* para o *cluster3* e para o gene 1947, respectivamente, NSY27 e NSY29. A tendência geral da variação da população celular dessas linhagens é semelhante à linhagem controle, NSY39, que contém apenas a inserção do gene *KanMx4*. Portanto, conclui-se que a deleção do gene 1947 e do *cluster 3* na linhagem JAY291 não prejudicou a capacidade competitiva das linhagens mutantes em relação a linhagem selvagem.

Ensaio de competição realizados anteriormente, competindo linhagens diplóides *knockout* para esses genes contra a parental selvagem JAY270, revelaram uma drástica diminuição na população de células mutadas ao longo de 15 ciclos de

competição. Desse modo, o resultado obtido com mutantes haploides não era esperado. No entanto, nos testes de competição realizados previamente, o efeito da deleção desses genes na linhagem diplóide JAY270 pode ter sido mascarado por um efeito de perda de heterozigosidade das linhagens mutantes ao longo de inúmeros ciclos de crescimento. A linhagem JAY270, assim como outras linhagens da mesma espécie, é conhecida por apresentar uma elevada taxa de rearranjos, principalmente nas regiões distais de seus cromossomos (Argueso et al. 2009; Dunn et al. 2012). É possível que esses rearranjos resultem em uma redução no *fitness* das linhagens *knockout* de JAY270, devido à perda de heterozigosidade no diplóide (Sellis et al. 2011; Goldstein 1992). Esses eventos poderiam explicar os fenótipos observados nos testes realizados anteriormente de competição com diplóides. Dessa forma, os experimentos realizados com mutantes da linhagem haploide JAY291 são mais robustos, visto que rearranjos entre cromossomos homólogos não estão presentes nesta linhagem, sugerindo que os genes em estudo não tem uma importância na fisiologia das linhagens sob as condições testadas.

Apesar dos resultados observados nos testes de competição com haploides, o fato de a linhagem PE-2 apresentar um desempenho fermentativo extremamente elevado e características de alta tolerância aos estresses decorrentes da fermentação, poderia tornar imperceptível interferências nessas características causadas pela deleção individual dos genes em estudo. Por outro lado, a linhagem padrão de laboratório S288c apresenta um desempenho fermentativo significativamente inferior ao de PE-2, assim como uma maior sensibilidade às condições de estresse encontradas durante o processo industrial (Argueso et al. 2009). Desse modo, pequenas alterações na fisiologia dessa linhagem durante a fermentação podem ser notadas com maior facilidade. Por esse motivo, como uma forma adicional de investigar uma possível relação do gene 1947 com o desempenho fermentativo da linhagem PE-2, realizou-se a inserção do mesmo na linhagem padrão de laboratório S288c. Este gene também foi inserido na linhagem isolada de campos de vinho RM11-1a, que se encontrava disponível na coleção de nosso laboratório. As linhagens construídas foram submetidas a testes de competição com suas respectivas parentais selvagens, a fim de se comparar o desempenho de ambas em competição (Figura 21). Esse teste foi

escolhido por ser eficiente na detecção de variações no *fitness* das linhagens mutantes frente a diversos fatores de estresse, comumente encontrados no ambiente industrial.

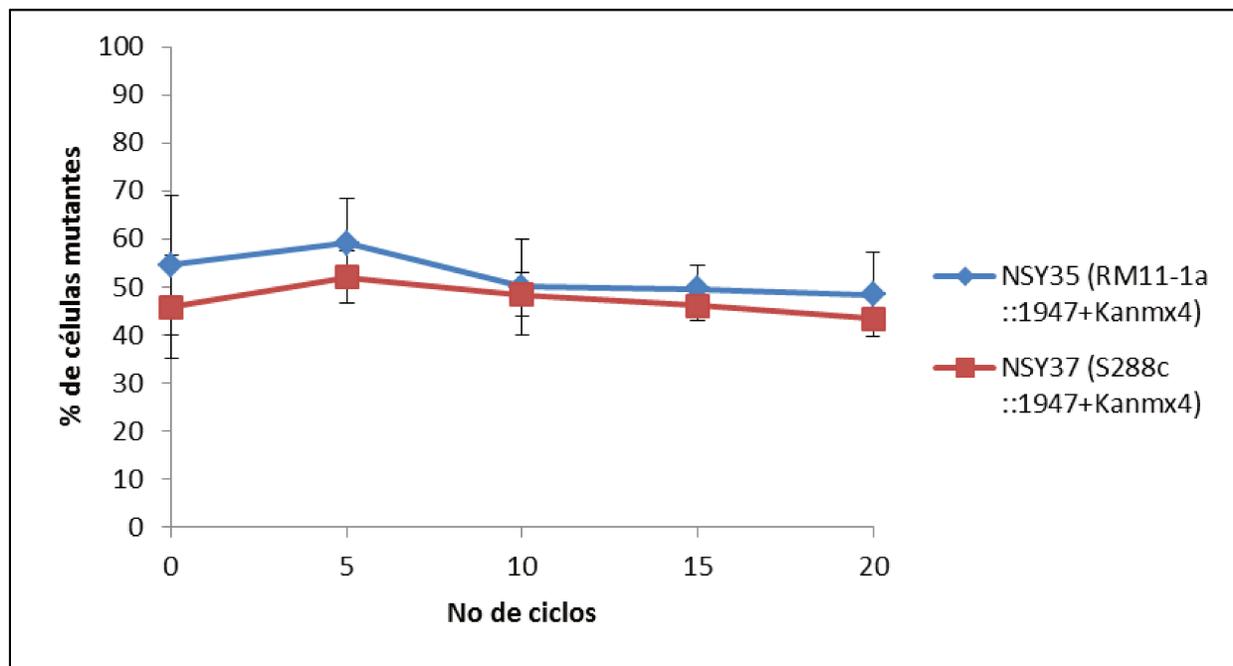


Figura 21. Variação na população de células das linhagens mutantes durante 20 ciclos de competição com suas respectivas linhagens parentais. O ensaio foi conduzido em xarope de cana-de-açúcar (18° BRIX, contendo 0.01% de extrato de levedura). No ciclo inicial, são misturadas células das linhagens mutantes, contendo a inserção dos genes *KanMx4* e 1947, e das respectivas linhagens parentais (S288c ou RM11-1a) visando uma proporção aproximada de 1:1. A população de cada tipo celular é monitorada ao longo dos ciclos de acordo com os métodos estabelecidos na seção 3.13.

Os resultados obtidos fornecem uma maior evidência de que o gene 1947 não apresenta relação com a capacidade competitiva da linhagem PE-2. Como pode ser observado na Figura 21, a inserção do gene 1947 nas linhagens RM11-1a e S288c não conferiu uma vantagem competitiva à essas linhagens sob as condições estabelecidas nos ensaios realizados. Este resultado, no entanto, não é suficiente para refutar totalmente uma possível importância deste gene durante o processo fermentativo. As divergências existentes entre o genoma das linhagens RM11-1a e S288c em relação ao genoma de linhagens industriais que possuem este gene podem ter influência na função do mesmo. Portanto, é possível que a falta deste *background* genético impeça a expressão deste gene e, conseqüentemente, sua ação.

4.2.8 Ensaios fermentativos

Foram realizados ensaios de fermentação visando avaliar se a deleção dos genes em estudo poderia interferir no rendimento e na produtividade de etanol das linhagens *knockout*, derivadas de JAY291 (Figura 22).

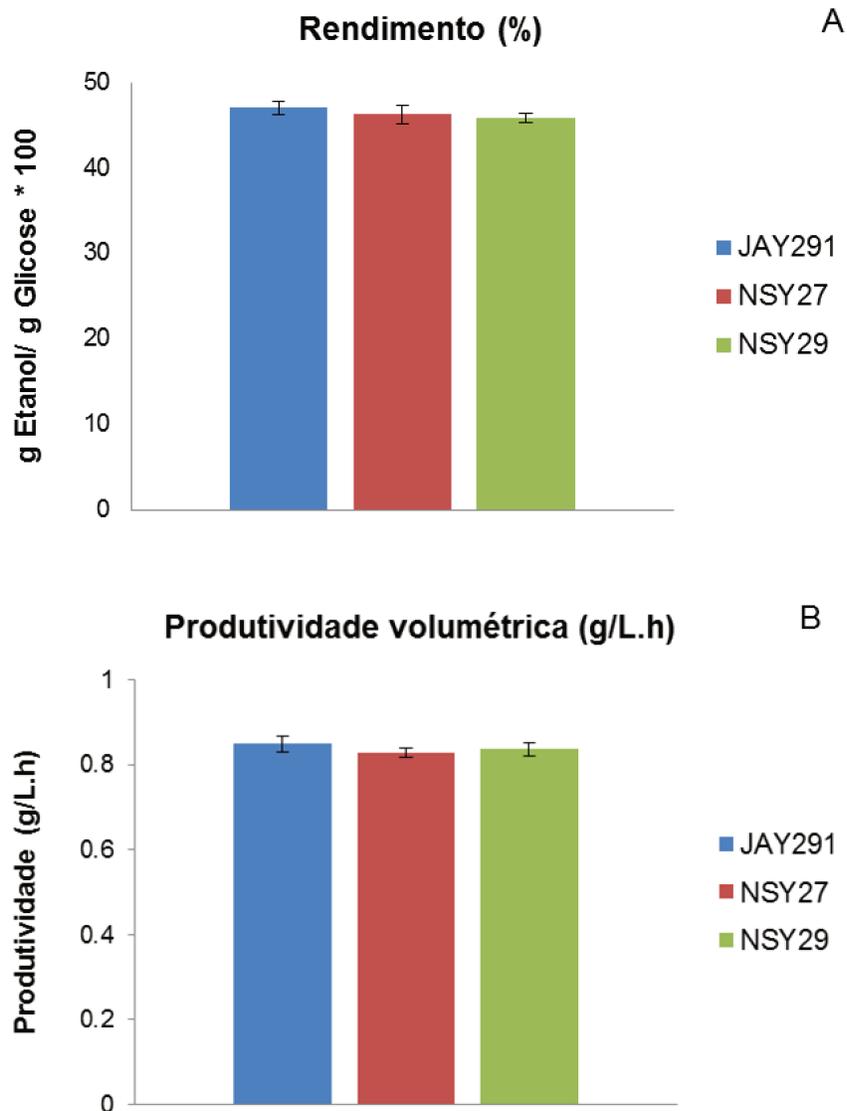


Figura 22. Desempenho fermentativo da linhagem selvagem JAY291 e das linhagens *knockout* para o *cluster3* e para o gene 1947, respectivamente, NSY27 e NSY29. As amostras foram coletadas ao final do 5º ciclo de um ensaio de fermentação conduzido em meio YPD 2% glicose. A) Comparação do rendimento de etanol atingido pelas diferentes linhagens durante o ensaio fermentativo. B) Comparação da produtividade volumétrica fermentativa das linhagens em estudo.

Como é possível observar na Figura 22, as linhagens NSY27 e SNY29 não apresentaram diferença significativa em relação a linhagem controle NSY39, no rendimento ou na produtividade volumétrica, sob as condições testadas neste experimento. Este resultado sugere que estes genes não apresentam relação com a via da fermentação alcoólica ou com a captação e utilização de açúcar. Cabe ressaltar, no entanto, que este resultado foi obtido para as condições experimentais específicas deste ensaio. Apesar de fornecer informações preliminares sobre o desempenho fermentativo destas linhagens, as condições deste ensaio diferem drasticamente das observadas em condições industriais, e podem ter desviado o metabolismo fermentativo da levedura para outras vias, como crescimento e formação de biomassa. É bastante conhecida a influência das condições de crescimento nos parâmetros de desempenho fermentativo, como produtividade volumétrica e rendimento em massa. Dessa forma, é possível que estes genes afetem esses parâmetros em condições que diferem das utilizadas neste experimento, como maior estresse ósmótico, competição com outros micro-organismos, temperaturas flutuantes, entre outros. Por esse motivo, pretendemos realizar novos ensaios para avaliar esses parâmetros utilizando outras condições, como xarope de cana-de-açúcar como substrato fermentativo e lavagens ácidas durante os ciclos de fermentação.

5 CONCLUSÃO

A partir dos experimentos conduzidos nesse trabalho, não foram observadas interferências no fenótipo das linhagens mutantes para os genes contidos no *cluster3* e para o gene 1947 que pudessem estar relacionadas a deleção ou inserção destes genes. Dessa forma, não foi possível correlacionar estes genes a uma função específica. Apesar disso, este trabalho apresenta fortes indícios de uma potencial função desses genes na fisiologia da linhagem PE-2 exposta a condições industriais, como por exemplo, o fato de esses genes estarem conservados em outras linhagens industriais e isolados naturais da espécie *Saccharomyces cerevisiae* e, adicionalmente, por apresentarem expressão diferencial ao longo de fermentações conduzidas em

usina. É possível que os testes realizados não tenham sido eficientes na reprodução das condições em que estes genes são necessários, principalmente, pelo fato de os experimentos conduzidos em laboratório diferirem drasticamente das condições encontradas no ambiente industrial. Portanto, a elucidação do papel destes genes requer uma quantidade maior de ensaios, e idealmente, uma avaliação do fenótipo das linhagens construídas no ambiente industrial. Dessa forma, propomos a realização de um conjunto mais extensivo de testes fenotípicos como perspectivas futuras na continuidade deste projeto. Devido à inviabilidade de testar essas linhagens no ambiente industrial, a realização de testes que reproduzem mais proximamente essas condições se fazem necessários. Nesse sentido, poderiam ser inseridos no conjunto de testes fenotípicos, ensaios de competição das linhagens *knockout* com outros micro-organismos comumente encontrados como contaminantes em dornas de fermentação. Adicionalmente, a análise do rendimento e produtividade dessas linhagens durante ciclos de fermentação em xarope ou melaço de cana-de-açúcar concentrados, incluindo-se lavagens ácidas também seriam interessantes, por simular as condições encontradas nas usinas.

Além disso, a caracterização geral dos 21 genes específicos da linhagem industrial JAY291, revelou que todos apresentam potencial para estarem relacionados com as características de interesse dessa linhagem. Dessa forma, a caracterização detalhada e estudo funcional dos outros 17 genes novos encontrados é de grande interesse. Como foi demonstrado pelas análises iniciais conduzidas, muitos genes que não puderam ser enfocados nesse trabalho, apresentam grande potencial para estarem envolvidos nas características desejáveis da linhagem PE-2.

Em suma, os dados gerados por este trabalho deverão servir como base para futuros estudos focados na elucidação do papel desses genes, ausentes no genoma da linhagem de referência S288c, visando tanto a geração de um conhecimento básico acerca da biologia e evolução da espécie *S. cerevisiae*, como a identificação de potenciais alvos para o melhoramento genético da linhagem PE-2.

6 REFERÊNCIAS BIBLIOGRÁFICAS

Akao, T., I. Yashiro, A. Hosoyama, H. Kitagaki, H. Horikawa, D. Watanabe, R. Akada, et al. 2011. "Whole-genome Sequencing of Sake Yeast *Saccharomyces Cerevisiae* Kyokai No. 7." *DNA Research: an International Journal for Rapid Publication of Reports on Genes and Genomes* 18 (6) (December): 423–34.

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.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. "Gapped BLAST and PSI-BLAST: a New Generation of Protein Database Search Programs" 25 (17): 3389–3402.

Amorim, H.V, M.L. Lopes, J.V. de Castro Oliveira, M. S. Buckeridge, and G.H. Goldman. 2011. "Scientific Challenges of Bioethanol Production in Brazil." *Applied Microbiology and Biotechnology* 91 (5) (September): 1267–75.

Argueso, Juan Lucas, Marcelo F Carazzolle, Piotr a Mieczkowski, Fabiana M Duarte, Osmar V C Netto, Silvia K Missawa, Felipe Galzerani, et al. 2009. "Genome Structure of a *Saccharomyces Cerevisiae* Strain Widely Used in Bioethanol Production." *Genome Research* 19 (12) (December): 2258–70.

Babrzadeh, F., R. Jalili, C. Wang, S. Shokralla, S. Pierce, A. Robinson-Mosher, P. Nyren, et al. 2012. "Whole-genome Sequencing of the Efficient Industrial Fuel-ethanol Fermentative *Saccharomyces Cerevisiae* Strain CAT-1." *Molecular Genetics and Genomics: MGG* 287 (6) (June): 485–94.

Basso, L.C., H.V. Amorim, A.J. Oliveira, and M.L. Lopes. 2008. "Yeast Selection for Fuel Ethanol Production in Brazil." *FEMS Yeast Research* 8 (7) (November): 1155–63.

Borneman, A. R., B. A. Desany, D. Riches, J. P. Affourtit, A. H. Forgan, I. S. Pretorius, M. Egholm, and P. J. Chambers. 2011. "Whole-genome Comparison Reveals Novel Genetic Elements That Characterize the Genome of Industrial Strains of *Saccharomyces Cerevisiae*." *PLoS Genetics* 7 (2) (January).

Doran, PM. 1995. *Bioprocess Engineering Principles*. London: Academic Press Limited.

Dunn, B., C. Richter, D. J. Kvitek, T. Pugh, and G. Sherlock. 2012. "Analysis of the *Saccharomyces Cerevisiae* Pan-genome Reveals a Pool of Copy Number Variants Distributed in Diverse Yeast Strains from Differing Industrial Environments": 908–924.

- Engel, S. R., and J. M. Cherry. 2013. "The New Modern Era of Yeast Genomics: Community Sequencing and the Resulting Annotation of Multiple *Saccharomyces Cerevisiae* Strains at the *Saccharomyces* Genome Database." *Database* 2013 (March 13).
- Gietz, R. D., R. H. Schiestl, A. R. Willems, and R. A. Woods. 1995. "Studies on the Transformation of Intact Yeast Cells by the LiAc/SS-DNA/PEG Procedure." *Yeast (Chichester, England)* 11 (4) (April 15): 355–60.
- Goldemberg, José. 2008. "The Brazilian Biofuels Industry." *Biotechnology for Biofuels* 1 (1) (January).
- Goldstein, D B. 1992. "Heterozygote Advantage and the Evolution of a Dominant Diploid Phase." *Genetics* 132 (4) (December): 1195–8.
- Haas, H. 2003. "Molecular Genetics of Fungal Siderophore Biosynthesis and Uptake: The Role of Siderophores in Iron Uptake and Storage." *Applied Microbiology and Biotechnology* 62 (4) (September): 316–30.
- Hampsey, M. 1997. "A Review of Phenotypes in *Saccharomyces Cerevisiae*." *Yeast (Chichester, England)* 13 (12) (September 30): 1099–133.
- Hider, R. C., and X. Kong. 2010. "Chemistry and Biology of Siderophores." *Natural Product Reports* 27 (5) (May): 637–57.
- Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease. 1989. "Site-directed Mutagenesis by Overlap Extension Using the Polymerase Chain Reaction." *Gene* 77 (1) (April 15): 51–9.
- Jacques, K.A., T.P. Lyons, and D.R. Kelsall. 2003. *The Alcohol Textbook*. 4th ed.
- Kowalska, E., and A. Kozik. 2008. "The Genes and Enzymes Involved in the Biosynthesis of Thiamin and Thiamin Diphosphate in Yeasts." *Cellular & Molecular Biology Letters* 13 (2) (January): 271–82.
- Kowalska, E., M. Kujda, N. Wolak, and A. Kozik. 2012. "Altered Expression and Activities of Enzymes Involved in Thiamine Diphosphate Biosynthesis in *Saccharomyces Cerevisiae* Under Oxidative and Osmotic Stress." *FEMS Yeast Research* 12 (5) (August): 534–46.
- Mieczkowski, P. A., F. J. Lemoine, and T. D. Petes. 2006. "Recombination Between Retrotransposons as a Source of Chromosome Rearrangements in the Yeast *Saccharomyces Cerevisiae*." *DNA Repair* 5 (9-10) (September 8): 1010–20.
- Mortazavi, A.; Williams, B. A.; McCue, K.; Schaeffer, L.; Wold, B. 2008. "Mapping and quantifying mammalian transcriptomes by rna-seq". *Nat Methods*, 5(7):621.

Novo, M., F. Bigey, E. Beyne, V. Galeote, F. Gavory, S. Mallet, B. Cambon, et al. 2009. "Eukaryote-to-eukaryote Gene Transfer Events Revealed by the Genome Sequence of the Wine Yeast *Saccharomyces Cerevisiae* EC1118." *Proceedings of the National Academy of Sciences of the United States of America* 106 (38) (September 22): 16333–8.

Pérez-Ortín, José, Amparo Querol, Sergi Puig, and Eladio Barrio. 2002. "Molecular Characterization of a Chromosomal Rearrangement Involved in the Adaptive Evolution of Yeast Strains": 1533–1539.

Pfaffl, M. W. 2001. "A New Mathematical Model for Relative Quantification in Real-time RT-PCR." *Nucleic Acids Research* 29 (9) (May 1).

Rose MD, Winston F, Hieter P. 1990. *Methods in yeast genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. "Heterozygote Advantage as a Natural Consequence of Adaptation in Diploids." *Proceedings of the National Academy of Sciences of the United States of America* 108 (51) (December 20): 20666–71.

Da Silva-Filho, Eurípedes Alves, Scheila Karina Brito dos Santos, Alecsandra Do Monte Resende, José Otamar Falcão de Moraes, Marcos Antonio de Moraes, and Diogo Ardaillon Simões. 2005. "Yeast Population Dynamics of Industrial Fuel-ethanol Fermentation Process Assessed by PCR-fingerprinting." *Antonie Van Leeuwenhoek* 88 (1) (July): 13–23. doi:10.1007/s10482-004-7283-8. <http://www.ncbi.nlm.nih.gov/pubmed/15928973>.

Sivakumar, Ganapathy, Daniel R. Vail, Jianfeng Xu, David M. Burner, Jackson O. Lay, Xumeng Ge, and Pamela J. Weathers. 2010. "Bioethanol and Biodiesel: Alternative Liquid Fuels for Future Generations." *Engineering in Life Sciences* 10 (1) (February): 8–18. doi:10.1002/elsc.200900061. <http://doi.wiley.com/10.1002/elsc.200900061>.

Stambuk, Boris U, Barbara Dunn, Sergio L Alves, Eduarda H Duval, and Gavin Sherlock. 2009. "Industrial Fuel Ethanol Yeasts Contain Adaptive Copy Number Changes in Genes Involved in Vitamin B1 and B6 Biosynthesis": 2271–2278. doi:10.1101/gr.094276.109.19.

Teste, Marie-Ange, Manon Duquenne, Jean M François, and Jean-Luc Parrou. 2009. "Validation of Reference Genes for Quantitative Expression Analysis by Real-time RT-PCR in *Saccharomyces Cerevisiae*." *BMC Molecular Biology* 10 (January): 99. doi:10.1186/1471-2199-10-99. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2776018&tool=pmcentrez&rendertype=abstract>.

Wang, Zhong, Mark Gerstein, and Michael Snyder. 2009. "RNA-Seq: a Revolutionary Tool for Transcriptomics." *Nature Reviews. Genetics* 10 (1) (January): 57–63. doi:10.1038/nrg2484.

Weber, Christian, Alexander Farwick, Feline Benisch, Dawid Brat, Heiko Dietz, Thorsten Subtil, and Eckhard Boles. 2010. "Trends and Challenges in the Microbial Production of Lignocellulosic Bioalcohol Fuels." *Applied Microbiology and Biotechnology* 87 (4) (July): 1303–15. doi:10.1007/s00253-010-2707-z. <http://www.ncbi.nlm.nih.gov/pubmed/20535464>.

Wheals, a E, L C Basso, D M Alves, and H V Amorim. 1999. "Fuel Ethanol After 25 Years." *Trends in Biotechnology* 17 (12) (December): 482–7. <http://www.ncbi.nlm.nih.gov/pubmed/10557161>.