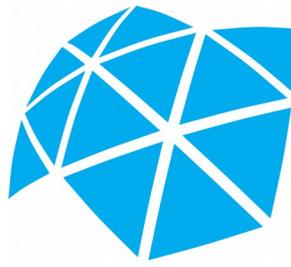


RESSALVA

Atendendo solicitação do(a) autor(a),
o texto completo desta tese será disponibilizado
somente a partir de 01/07/2021.

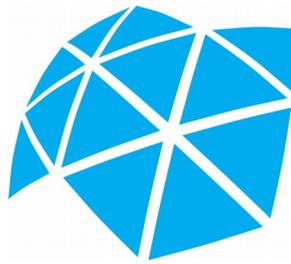


**Universidade Estadual Paulista “Júlio de Mesquita Filho”
Instituto de Biociências de Botucatu
Programa de Pós-Graduação em Biotecnologia**

**Identificação de assinaturas sistêmicas associadas à tolerância ao etanol em
linhagens de *Saccharomyces cerevisiae***

Ivan Rodrigo Wolf

**Botucatu – SP
2019**



**Universidade Estadual Paulista “Júlio de Mesquita Filho”
Instituto de Biociências de Botucatu
Programa de Pós-Graduação em Biotecnologia**

**Identificação de assinaturas sistêmicas associadas à tolerância ao etanol em
linhagens de *Saccharomyces cerevisiae***

**Doutorando: Ivan Rodrigo Wolf
Orientador: Dr. Guilherme Targino Valente**

Tese apresentada ao Programa de Pós-Graduação em Biotecnologia do Instituto de Biociências de Botucatu da Universidade Estadual Paulista “Júlio de Mesquita Filho”, para obtenção do título de doutor.

**Botucatu - SP
2019**

W853i

Wolf, Ivan Rodrigo

Identificação de assinaturas sistêmicas associadas à tolerância ao etanol em linhagens de *Saccharomyces cerevisiae* / Ivan Rodrigo Wolf. -- Botucatu, 2019

118 p.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências, Botucatu
Orientador: Guilherme Targino Valente

1. Biologia de Sistemas. 2. *Saccharomyces cerevisiae*. 3. Bioinformática. 4. Biologia celular. 5.

Tolerância ao Etanol. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências, Botucatu. Dados fornecidos pelo autor(a).

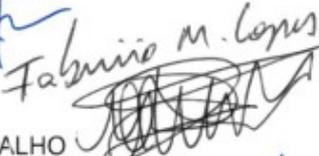
Essa ficha não pode ser modificada.

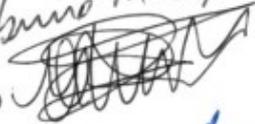
ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE IVAN RODRIGO WOLF, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA, DO INSTITUTO DE BIOCIÊNCIAS - CÂMPUS DE BOTUCATU.

Aos 01 dias do mês de julho do ano de 2019, às 14:00 horas, no(a) Sala B - Central de Aulas do IBB, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Prof. Dr. GUILHERME TARGINO VALENTE - Orientador(a) do(a) Departamento de Engenharia de Bioprocessos e Biotecnologia / Faculdade de Ciências Agrônômicas de Botucatu - UNESP, Prof. Dr. RAFAEL PLANA SIMÕES do(a) Departamento de Engenharia de Bioprocessos e Biotecnologia / Faculdade de Ciências Agrônômicas de Botucatu - UNESP, Prof. Dr. FABRÍCIO MARTINS LOPES do(a) Departamento Acadêmico de Computação / Universidade Tecnológica Federal do Paraná, Prof. Dr. ROBSON FRANCISCO CARVALHO do(a) Departamento de Morfologia / Instituto de Biociências de Botucatu - UNESP, Professor Associado PAULO EDUARDO MARTINS RIBOLLA do(a) Departamento de Parasitologia / Instituto de Biociências - IB - Unesp - Câmpus de Botucatu, sob a presidência do primeiro, a fim de proceder a arguição pública da TESE DE DOUTORADO de IVAN RODRIGO WOLF, intitulada **Identificação de assinaturas sistêmicas associadas à tolerância ao etanol em linhagens de *Saccharomyces cerevisiae***. Após a exposição, o discente foi arguido oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: APROVADO _____. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.


Prof. Dr. GUILHERME TARGINO VALENTE

Prof. Dr. RAFAEL PLANA SIMÕES 

Prof. Dr. FABRÍCIO MARTINS LOPES 

Prof. Dr. ROBSON FRANCISCO CARVALHO 

Professor Associado PAULO EDUARDO MARTINS RIBOLLA 

“Eu sei que você acredita que entende o que acha que eu disse, mas não tenho certeza se você entende que o que você ouviu não é o que eu quis dizer.”

- Uma das infinitas frases exibidas no Gaussian, um programa para cálculo de estrutura eletrônica.

Dedico este trabalho a meu pai, minha mãe, minha família e aos meus amigos,
estes ajudaram a constituir os alicerces do que sou.

Agradecimentos

Agradeço à *Deus*, pois nas horas mais escuras uma luz sempre brilhou na forma de um familiar, um amigo, o trecho de um texto ou um vídeo, um besouro de ponta cabeça na calçada... enfim, em todo e qualquer sinal de esperança. Hoje sinto sua presença naqueles que estão a minha volta.

Agradeço à *Vida* e a *Morte*, à primeira gosta de rir da nossa cara e de nos fazer tropeçar, mas sempre nos ajuda a levantar e nos empurra para seguir adiante, mais longe e mais determinados; à segunda sempre está ao nosso lado com um olhar de ternura e um sorriso reconfortante e alegre que nos lembram constantemente de darmos o nosso melhor durante essa jornada. Elas são os símbolos de início e fim que nos ensinam que passamos por diversos ciclos e que tudo tende a melhorar, terminar bem e (de preferência) em uma grande bagunça.

Agradeço aos meus pais *Ivan Wolf* e *Inês Maristela Palamar Wolf*, pela vida maravilhosa que me proporcionaram e por me ensinarem os bons princípios que me orientam fundamentalmente. Lembrar de vocês me reconforta e preenche meu coração, por isso sei que nunca vou estar sozinho. Espero conseguir passar a diante pelo menos um pouco do amor e carinho que recebi de vocês. Se os pais incomodam os filhos, só tenho a agradecer o incomodo que recebi, pois sei que são a prova do quanto vocês me amam e me querem bem. Graças a vocês não tenho medo da vida, da dor, do trabalho e do que vier pela frente. Amo vocês!

Agradeço à minha vó Jaci do Rocio Palamar, pois ao longo dos anos que se passaram depois do seu falecimento notei que o caminho que segui me levou a compreender melhor a doença que te afligi. Esse caminho também, pouco a pouco, me forneceu as ferramentas para ajudar a, quem sabe um dia, encontrar a cura que teria permitido você estar junto de nós. Se hoje estou conseguindo o título de doutor, grande influência veio dessa vontade de ter salvo você. Te amo onde quer que você esteja.

Agradeço à algumas pessoas que são como irmãos e irmãs para mim: *Jefferson Andrei Silva*, *Marina Krasniak*, *Aline Bonk*, *Camila Borão*, *Danyella Nunes*, *Joni Marcos Lopes*, *Luiz Ricardo Fantim*, *Oliver Lie Bueno*, *Everton Souza*. Por cada um de vocês tenho amor, carinho, amizade, gratidão e admiração. Mesmo depois de muitos anos terem se passado nós continuamos os mesmos. Vocês já me viram com alegria e com raiva, feliz e triste, melancólico e otimista, altruísta e egoísta, flexível e teimoso, legal e insuportável. Vocês viram o que tenho de melhor e também o que tenho de pior. Ainda assim, me recebem com um sorriso no rosto, algo para beber, conversam comigo até tarde (quando não até o raiar do sol), riem comigo e combinam o próximo encontro. Vocês sempre serão pessoas especiais para mim. Sinto cada um de vocês vivos dentro do meu coração, e assim espero viver no coração de vocês.

Agradeço aos meus primos e primas *Ederson Nogueira dos Santos*, *Tháise Nogueira dos Santos*, *Edimar Wolf* e *Edicler Wolf*. Durante muito tempo vocês foram meus irmãos e irmãs, assim não me senti um filho único. Com vocês aprendi a desenhar um pouco melhor, ler um pouco mais e a usar minha imaginação. Obrigado por todas as aventuras, acampamentos, viagens, conversas e os finais de semana brincando com Lego, jogando RPG de mesa, Xadrez, Megaman, Zelda, Mario, Mario Party, GTA, Resident Evil, Doom, Grandia, Crash Bandicoot (...) em noitadas regadas à guloseimas variadas e muitas risadas.

Agradeço ao meu orientador *Guilherme Targino Valente*, você me colocou diante de desafios que exauriram minhas forças, acabaram com minha mente, abalaram minha fé e que eu realmente achei que não fosse capaz de superar. Mas sei que você fez isso sabendo que eu estaria preparado para superar cada etapa e cada desafio. Essas superações profissionais também se devem ao fato de que você estava lá para me orientar, me ajudar, e me mostrar que com foco e competência posso resolver qualquer problema! Você me ensinou a ser um profissional melhor do que eu era. E, de quebra, se mostrou uma pessoa extremamente sábia, me deu conselhos para a vida e ainda foi humilde de me escutar quando eu tinha uma contraposição. Você é um grande professor, orientador e profissional. Obrigado!

Agradeço à *Amanda Piveta Schnepfer*, que chegou de mansinho e me cativou com seu jeito singular e um sorriso que está no seu olhar. Você me mostrou como fazer fogo com duas pedras. Por sua causa o sabor, a cor e o cheiro do vinho nunca mais serão os mesmos. Obrigado!

Agradeço à todo o pessoal do SBGL (*Systems Biology and Genomic Laboratory – FCA – UNESP de Botucatu*) e LGI (*Laboratório de Genômica Integrativa – IBB - UNESP de Botucatu*), em especial a *Lauana Fogaça de Almeida, Rodrigo de Oliveira Almeida, Luiz Henrique Cardozo, Camila Cristina de Oliveira, Lucas Lazari, Lucas Farinazzo, Guilherme Luz, Eric Kazuo Kawagoe, Diego Marques, Jordana Oliveira, Adauto Cardoso, Rafael Takahiro Nakajima, Luiz Augusto Bovolenta e Érica Ramos*. Obrigado pelas muitas horas de trabalho, pelas conversas no laboratório, pelas festas, pelas dicas, pelas idas ao bar, pela ajuda, pelos favores feitos e pelos favores cobrados. Tudo que vi, vivi e aprendi com cada um de vocês vêm a acrescentar algo positivo para melhorar a vida que tenho. Espero que possamos nos ver de novo para compartilhar experiências e histórias. Obrigado!

Agradeço aos professores *Rafael Plana Simões, César Martins, Alexandre Dal Pai, José Luiz Rybarczyk Filho* e as professoras *Rejane Maria Tommasini Grotto, Carla dos Santos Riccardi, Ivana Cesarino, Valéria Cristina Rodrigues Sarnighausen, Sarita Candida Rabelo e Lilian Cristina Pereira*. Aprendi muito com suas lições profissionais e conversas casuais. Vocês ajudaram a expandir os meus horizontes.

Agradeço aos professores da *Universidade Estadual de Londrina - UEL*. Em especial aos meus ex-orientadores *Laurival Antônio Vilas Bôas e Rogério Fernandes de Souza*, obrigado pelos desafios, por jamais duvidarem da minha competência, pelas dicas, sugestões, críticas e conselhos. Vocês são uma grande inspiração profissional.

Agradeço também aos amigos e amigas *Natália Venturelli, Marcos Edgar Herkenhoff, Raquel Bozini Gallo, Jaqueline Fernanda Dionisio, Juceli Gonzalez Gouveia, Ana Paula Scaramal Ricieto, Carla Suzuki Altrão, Carolina Galdino Gumiero Ribeiro e Rachel Colauto Milanezi Aguiar* e à todos os outros cujo os nomes não cabem aqui. Vocês são pessoas incríveis. Que nossas amizades continuem através do tempo e do espaço. Tenho certeza que não vão faltar oportunidades para um churrasco, um almoço, um café, um barzinho... Que todos vocês tenham sucesso, felicidade e alegria em suas vidas, vocês merecem!

Agradeço todos os *docentes e técnicos* do Departamento de Bioprocessos e Biotecnologia da Faculdade de Ciências Agrônômicas, UNESP- Câmpus de Botucatu, em especial à *Raquel Francisco Rubio, José Eduardo Gomes Montanha, Leonardo Nazario de Moraes, Jackson Eliezer Neves Batalha, Marcelo Felipe Matias, Martin Kássio Leme*

da Silva, Firmo Sousa Campos. Vocês fazem o departamento funcionar, me ajudaram, me ensinaram e me trataram como um colega de trabalho, obrigado por tudo!

Agradeço aos *graduandos* do Departamento de Bioprocessos e Biotecnologia da Faculdade de Ciências Agrônômicas, UNESP- Câmpus de Botucatu. Eu entrava na sala para ensinar, mas depois de cada curso, palestra ou aula senti que havia aprendido mais do que vocês. A lembrança dos seus rostos e das experiências que trocamos vão me manter, para sempre, um jovem aprendiz.

Agradeço à *Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP - Botucatu*, pelo suporte, estrutura e bons professores que contribuíram imensamente para minha formação. E aproveito para agradecer ao *Programa de Pós-Graduação em Biotecnologia* do *Instituto de Biociências*, pela oportunidade e suporte. Aproveito também para agradecer aos *colegas da pós-graduação*, aqui seria injusto tentar citar qualquer nome; e também à secretaria da pós graduação, em especial ao *Davi Barcellos de Oliveira Müller* sempre solucionando os problemas burocráticos junto à secretaria desde o início do meu doutorado.

Agradeço à *família Sidol*, em especial à filha mais velha *Poli*. Nossas vidas agora seguem caminhos separados, mas vou levar comigo lembranças carinhosas e felizes de todos. Com vocês aprendi valorosas lições de vida. Acho que sem nossa convivência, muito provavelmente, eu não teria chegado até aqui. Obrigado pelos anos de boa companhia e amizade. Desejo à vocês tudo de bom.

Agradeço à *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)* pelo primeiro ano de auxílio e à *Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – Processo número 2015/19211-7)* pelo auxílio financeiro durante esses anos de muito trabalho e aprendizado do doutorado.

Resumo

A crescente demanda por combustíveis fósseis e as instabilidades econômicas relacionadas à sua utilização despertam o interesse no desenvolvimento de biocombustíveis tais como o bioetanol. O processo de produção mais comum do bioetanol é a tecnologia de primeira geração, no qual o organismo mais amplamente utilizado é a levedura *Saccharomyces cerevisiae*. No entanto, a alta concentração de etanol gera toxicidade para *S. cerevisiae* e constitui um dos fatores limitantes na produção deste bioetanol. Assim, a produção de linhagens mais resistentes a esse estressor constitui um ponto chave no desenvolvimento dos processos biotecnológicos relacionados ao bioetanol. Nesse contexto, análises sistêmicas são pouco empregadas para o entendimento do fenômeno de tolerância ao etanol, deixando a busca por genes candidatos algo bastante laborioso e custoso. Assim, as diferenças entre linhagens pouco tolerantes (LT) e altamente tolerantes (HT) ao etanol (EtOH) foram analisadas por meio de ferramentas de bioinformática como a biologia de sistemas e a análise de transcriptomas. Os resultados mostraram que as linhagens HT e LT utilizam inicialmente um sistema de resposta à estresse comum, mas que devido as mudanças na estrutura das redes metabólicas, regulatórias e de interação proteína-proteína de cada linhagem, mecanismos de resposta ao estresse por etanol distintos são ativados em cada grupo. As linhagens LT mantêm a homeostase da célula com o ciclo celular ativo, efetuando o ajuste dos metabolismos ao custo de que elementos transponíveis ficam ativos e agentes anti oxidantes importantes deixam de ser produzidos, permitindo que ocorram danos no DNA. As linhagens HT interrompem seu ciclo celular e boa parte do metabolismo, porém, elas mantêm a produção e reparo de proteínas para estabilizar as membranas celulares e o controle de espécies reativas de oxigênio (ROS) em conjunto com a metabolização do etanol. Uma vez que o EtOH está sendo utilizado para obtenção de carbono no ciclo do ácido cítrico, as linhagens HT resistem a concentrações maiores de etanol. Assim, o conhecimento aqui gerado poderá ser aproveitado para enriquecer as discussões futuras sobre o fenômeno estudado e o desenvolvimento biotecnológico.

Palavras-chave: Biologia de Sistemas, *Saccharomyces cerevisiae*, Bioinformática, Biologia celular, Tolerância ao Etanol

Abstract

The increasing demand and economic instabilities related to the use of fossil fuels raised interest in the development of biofuels such as bioethanol. The most common bioethanol production process is from the first generation technology, in which the most widely used organism is the yeast *Saccharomyces cerevisiae*. However, the high concentration of ethanol generates toxicity to *S. cerevisiae* and this is one of the limiting factors in the bioethanol production. Thus, the production of more resistant strains to this stressor is crucial to the development of biotechnological processes related to bioethanol. The systems analyzes are poorly explored to understand the ethanol tolerance phenomenon, leading the search for candidate genes labor intensive and expensive. Thus, the differences between low tolerant (LT) and high tolerant (HT) strains to ethanol (EtOH) were analyzed using bioinformatics tools like systems biology and transcriptome analysis. The results showed that HT and LT strains initially used a common stress response system but due to changes in the network structure of metabolic, regulatory, and protein-protein interactions of each strain, different mechanisms of stress response are activated in each group. LT strains maintain the cellular homeostasis activating cell cycle, and adjusting metabolisms, despite activation of transposable elements and disruption of important anti-oxidant agents production, allowing the occurrence of DNA damage. HT strains disrupt the cell cycle and most of metabolisms, but maintaining protein production and repair to stabilize cell membranes and control reactive oxygen species (ROS), working together with the ethanol metabolism. Moreover, HTs are using the EtOH to obtain carbon in the citric acid cycle, thus, it may be the reason of HT strains surpass high ethanol concentration. Altogether, the knowledge here generated may be used to enrich future discussions about the studied phenomenon and to provide biotechnological development.

Keywords: Systems Biology, *Saccharomyces cerevisiae*, Bioinformatics, Cell Biology, Ethanol Tolerance

Lista de Ilustrações

Figura 1: Produção de bioetanol (bilhões de litros) em 2012	19
Figura 2: Classificações do bioetanol como biocombustível	20
Figura 3: Filogenia de <i>Saccharomyces cerevisiae</i>	21
Figura 4: Efeito de várias concentrações de etanol no crescimento de <i>Saccharomyces cerevisiae</i>	22
Figura 5: Nuvem de palavras mostrando termos ontológicos relacionados à tolerância ao etanol em <i>S. cerevisiae</i>	24
Figura 6: Tipos de dados e técnicas de <i>Machine learning</i>	29
Figura 7: Desenho experimental	33
Figura 8: Representação do corte feito nos dados de crescimento para ajuste ao modelo	47
Figura 9: Geração das redes de controle e tratamento	50
Figura 10: Visão geral da análise metabolômica	54
Figura 11: Análise do conjunto referente às DAMs	55
Figura 12: Comparação quantitativa e qualitativa dos genes diferencialmente expressos entre as linhagens analisadas	57
Figura 13: Diagramas comparando genes e termos GO encontrados em diferentes estudos	58
Figura 14: Via do metabolismo de arginina e prolina	65
Figura 15: Modelo proposto para o fluxo de etanol no interior da célula da levedura	66
Figura 16: Genes diferencialmente expressos na provável via de tamponamento de etanol	67
Figura 17: Ensaio de viabilidade celular e contagem de células que acumulam ROS	68
Figura 18: Distribuição dos degrees para a rede controle e rede tratamento	72
Figura 19: Análise da função DDC normalizada para redes de controle e tratamento	73
Figura 20: Crescimento celular em meio com estresse máximo de etanol e enriquecido com espermidina	75

Lista de Tabelas

Tabela 1: descrição das linhagens estudadas	33
Tabela 2: Agrupamento das linhagens estudadas e suas tolerâncias máximas ao EtOH	52
Tabela 3: Resumo das métricas de montagem da linhagem BMA64-1A considerando 196 montagens	53
Tabela 4: Número de Genes Diferencialmente expressos por linhagem	56
Tabela 5: Genes diferencialmente expressos comumente relacionados a alguns estressores	60
Tabela 6: Processos biológicos obtidos pelo REVIGO em grupos HT e LT	62
Tabela 7: KEGG pathways significativas detectadas pelo método GAGE	63
Tabela 8: Comparação dos dados de ensaios ROS, H2AX e SDH	68
Tabela 9: Ensaio de diferença na densidade óptica	69
Tabela 10: Porcentagem de células acumulando danos ao DNA (ensaio utilizando o anticorpo H2AX)	70
Tabela 11: Diferença entre os quadrantes do ensaio de anexina	70
Tabela 12: Estatísticas gerais das redes específicas	71
Tabela 13: Diferença estatística em medidas topológicas dos nós	72
Tabela 14: Performance preditiva em diferentes matrizes	76

Lista de Siglas

ADH – Aldeído desidrogenase

ALD – Álcool desidrogenase

BC - *burtconstraint*

BET - *betweenness*

C - *coreness*

CLOS - *closeness*

CO₂ - Dióxido de carbono

DAM – *Differentially Abundant Metabolites* (Metabólitos diferencialmente abundantes)

ECC - *eccentricity*

EIG - *eigenvectors*

EtOH - etanol

FDR - *False discovery rate*

GO - *Gene ontology term*

GRN - *Gene regulatory network* (rede regulatória genética)

HS - *hub score*

HT - *High tolerant*

K - *degree*

Knn - *Average nearest neighbor degree*

Inc-p - rede de interação *long-non-coding RNAs* com proteínas

LT - *Low tolerant*

Matriz CT - Matriz controle/tratamento

Matriz HTLT - Matriz Altamente tolerante/pouco tolerante

NRBP - *National Bioresources Project*

PPI - *Protein-protein interaction network* (rede de interação proteína-proteína)

PR - *page rank*

ROS - *Reactive oxygen species* (espécies reativas de oxigênio)

S - *strength*

SC - *subgraph centrality*

SGD – *Saccharomyces genome database*

T – *transitivity*

TE – *transposable element* (elemento transponível)

YPD - *yeas extract peptone*

Sumário

1. Introdução	18
1.1. Biocombustíveis e o fenômeno de tolerância ao etanol	18
1.2. Tolerância ao etanol em <i>Saccharomyces cerevisiae</i>	21
1.3. Biologia de sistemas, transcriptomas e seus empregos no estudo de tolerância ao etanol	25
1.4. Hipóteses	30
2. Objetivos	31
2.1. Objetivos geral	31
2.2. Objetivos específicos	31
3. Materiais e métodos	32
3.1. Experimentos de tolerância máxima ao etanol	32
3.2. Extração de biomoléculas, controle de qualidade e quantificação	34
3.3. Ensaio de biologia celular	36
3.3.1. Quantificação de células com danos ao DNA e de células com acumulação de espécies reativas de oxigênio	37
3.3.2. Ensaio de Succinato Desidrogenase	38
3.3.3. Viabilidade Celular	39
3.4. Sequenciamento, montagem e anotação do genoma	40
3.5. Análise de RNA-Seq e metabolomas: expressão diferencial , Enriquecimento de GOs e anotações KEGG	43
3.6. Tratamento com espermidina	45
3.7. Rede global integrada e análises estatísticas	47
3.8. Machine Learning aplicado a redes biológicas	48
4. Resultados	52
4.1. O nível de tolerância ao etanol	52
4.2. Montagem e anotação genômica	52
4.3. Análise de abundância diferencial de metabólitos	53
4.4. Análise de expressão diferencial	56
4.5. Citometria de fluxo e ensaio SDH	67
4.6. Rede global integrada e análises estatísticas	71
4.7. Experimentos de crescimento celular acrescidos de espermidina	74
4.8. Rede global integrada e análises de Machine Learning	76
5. Discussão	77
5.1. Remodelando o <i>diauxic shift</i> : mecanismo fundamental para a tolerância ao etanol	77
5.2. A espermidina tem um papel importante para a tolerância ao etanol	80
5.3. O estresse por etanol ocasiona um expressivo rewiring das redes: os módulos basais tendem a ser os que se mantêm mais ativos	82
5.4. Rede global integrada e análises de Machine Learning	83
5.5. Visão geral da tolerância ao etanol em <i>S. cerevisiae</i>	84
6. Conclusões	86
7. Referências	87

Apêndices

1	Tabela completa com Termos do REVIGO	99
2	Genes Diferencialmente Expressos e Metabólitos Diferencialmente Abundantes	102
3	Estudo Piloto de Machine Learning aplicado a redes Biológicas	103
4	Projetos realizados durante o doutorado	113

1. Introdução

1.1. Biocombustíveis e o fenômeno de tolerância ao etanol

O crescimento das demandas energéticas acompanham o crescimento da população mundial, e atualmente as fontes primárias de energia se baseiam em combustíveis fósseis não renováveis (GUPTA; VERMA, 2015). Os combustíveis fósseis são a fonte de aproximadamente 80% da energia consumida no mundo, conseqüentemente, o preço e a crescente demanda destes combustíveis afeta diretamente a economia global (CHAKRABORTY et al., 2012). Este fato desperta o interesse no desenvolvimento do mercado de biocombustíveis em vários países, com o objetivo de amenizar os efeitos econômicos causados pelo preço dos combustíveis fósseis, mitigar as emissões de CO₂ e atingir as metas estabelecidas em acordos internacionais sobre o meio ambiente (MUSSATTO et al., 2010).

Os biocombustíveis constituem uma forma sustentável de obtenção de energia e podem ser produzidos a partir de fontes agrárias e biomassa de plantas, como o trigo, beterraba e milho; e de alguns resíduos biológicos, como palha, madeira e mesmo alguns resíduos domésticos (DEMIRBAS, 2017). Dentre os biocombustíveis, o bioetanol está entre os mais importantes por ser caracterizado como um dos mais promissores substitutos da gasolina no mercado de combustíveis para transporte (CHAKRABORTY et al., 2012; GUPTA; VERMA, 2015).

A importância do bioetanol se torna mais clara ao observar que os Estados Unidos da América e o Brasil são os dois maiores produtores de bioetanol e estima-se que serão os dois principais consumidores deste biocombustível em 2024 (GUPTA; VERMA, 2015; DEMIRBAS, 2017) (**Figura 1**).

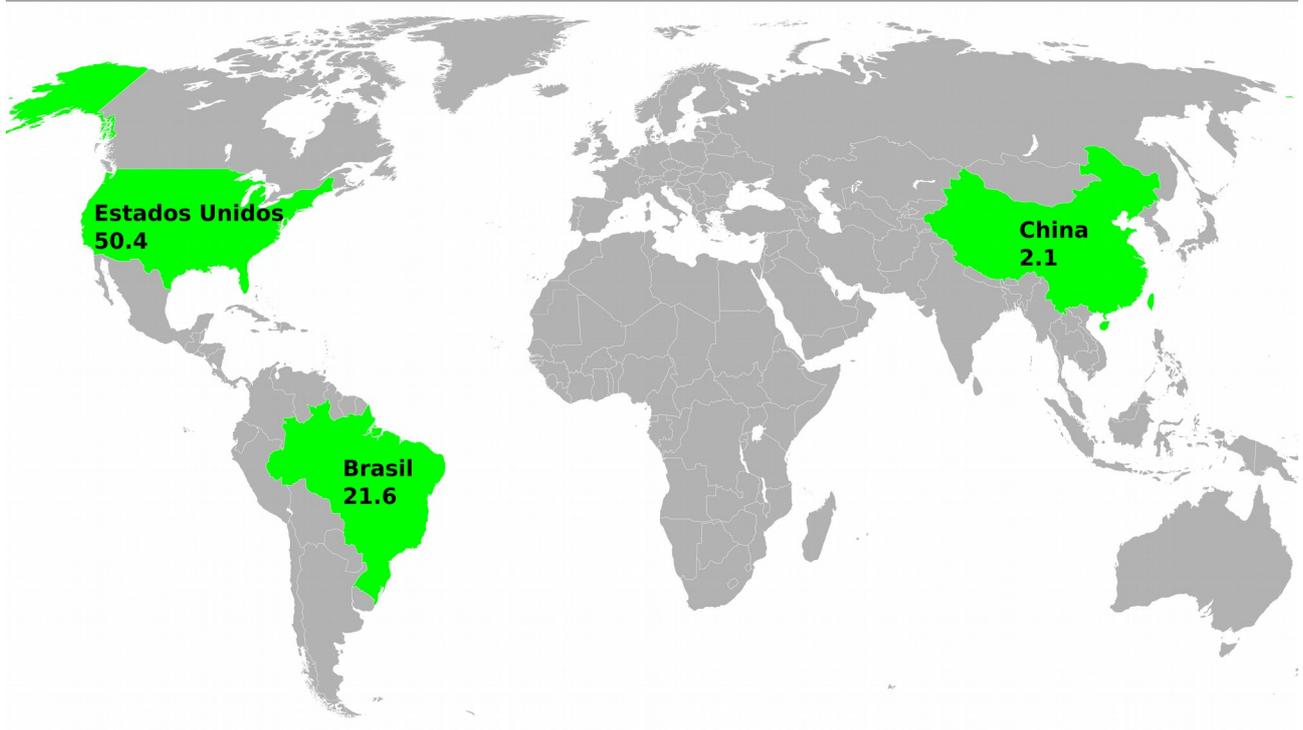


Figura 1: Produção de bioetanol (bilhões de litros) em 2012. Fonte (GUPTA; VERMA, 2015).

O bioetanol é um biocombustível secundário obtido principalmente pelas tecnologias de primeira geração, a fermentação alcoólica (**Figura 2**). Neste processo, geralmente os açúcares de seis carbonos de sementes, grãos ou da cana-de-açúcar são convertidos em etanol por enzimas da levedura (NIGAM; SINGH, 2011).

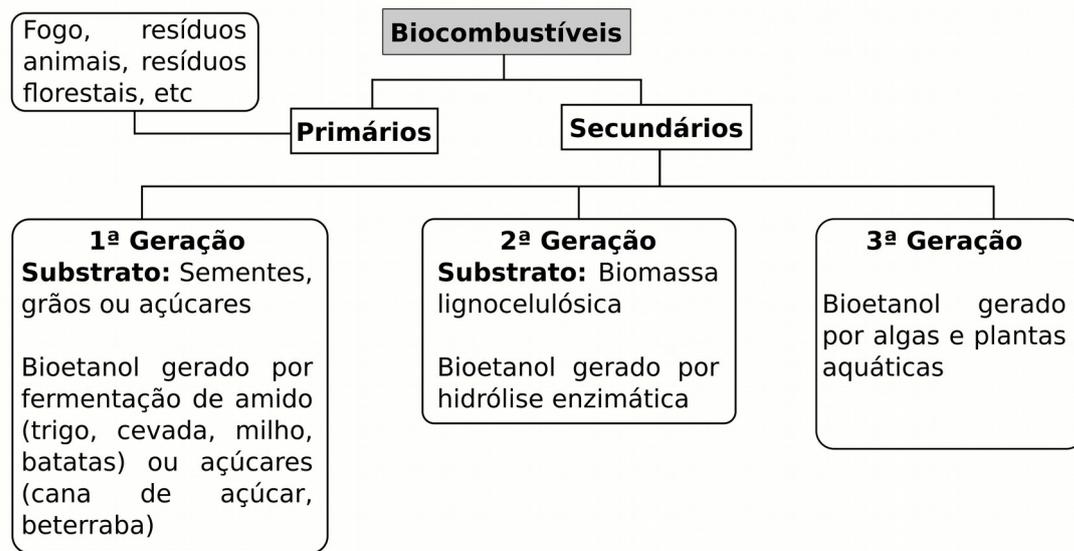


Figura 2: Classificações do bioetanol como biocombustível. Fonte: (NIGAM; SINGH, 2011).

A levedura *Saccharomyces cerevisiae* possui diversas linhagens laboratoriais e depósitos em bancos de dados, pertence ao filo Ascomycota, subfilo Saccharomycotina e ao gênero *Saccharomyces* (**Figura 3A-C**). Esta levedura destaca-se como os organismos mais utilizados para processos industriais de produção de bioetanol por apresentar alta capacidade fermentativa e de crescimento em condições anaeróbias, possuir alta tolerância às condições industriais, tais como temperatura e pH, e também apresentar uma tolerância relativamente alta ao etanol e a outros inibidores (MUSSATTO et al., 2010; DEMEKE et al., 2013).

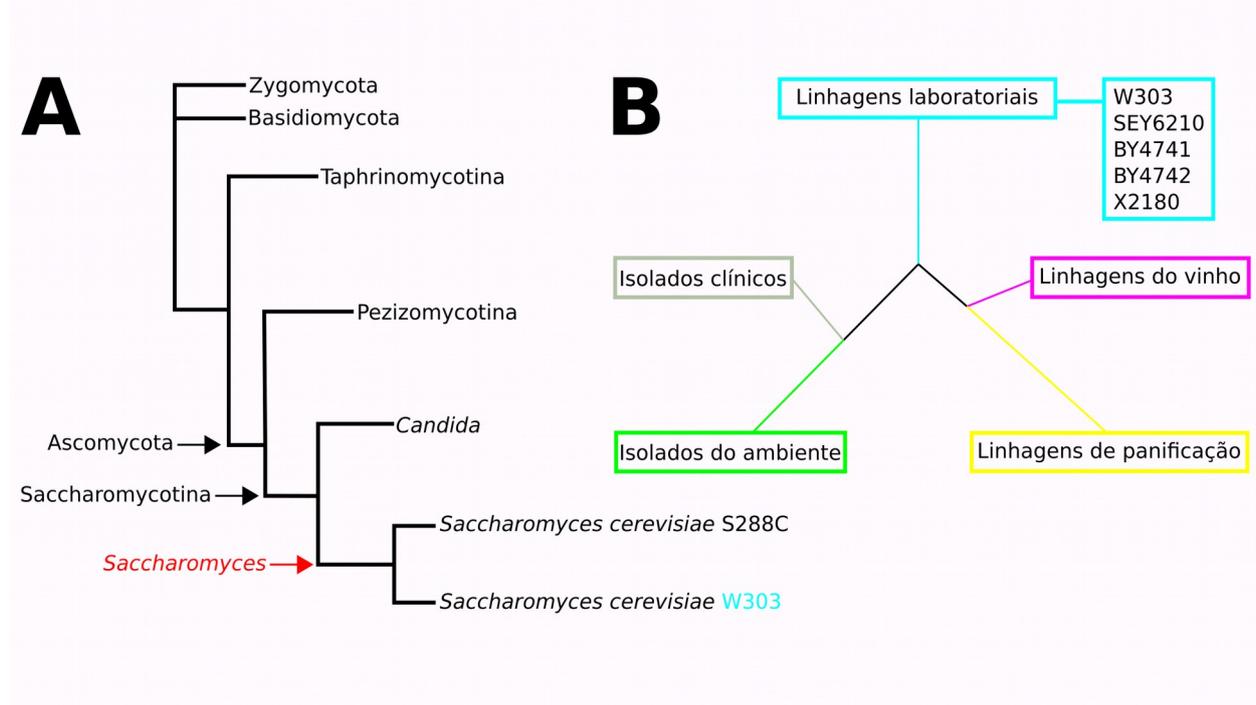


Figura 3: Filogenia de *Saccharomyces cerevisiae*. **A:** Localização de *S. cerevisiae* em Saccharomycetales, Vermelho: Clado de *Saccharomyces*; Azul: Linhagem laboratorial W303 de *S. cerevisiae*; **B:** Filogenia geral de linhagens do *Saccharomyces* Genome Database. Modificado de: (BUTLER, 2010; SONG et al., 2015; GALLONE et al., 2016).

1.2. Tolerância ao etanol em *Saccharomyces cerevisiae*

Durante a produção de bioetanol a levedura é exposta a alguns tipos de estresses destacando-se o osmótico, por aquecimento, o oxidativo e o aumento da concentração de etanol (AUESUKAREE, 2017). Apesar de *S. cerevisiae* ser comumente utilizada na produção de bioetanol, a alta concentração deste composto compromete a sobrevivência da levedura afetando sua taxa de crescimento (STANLEY et al., 2010a) (**Figura 4**).

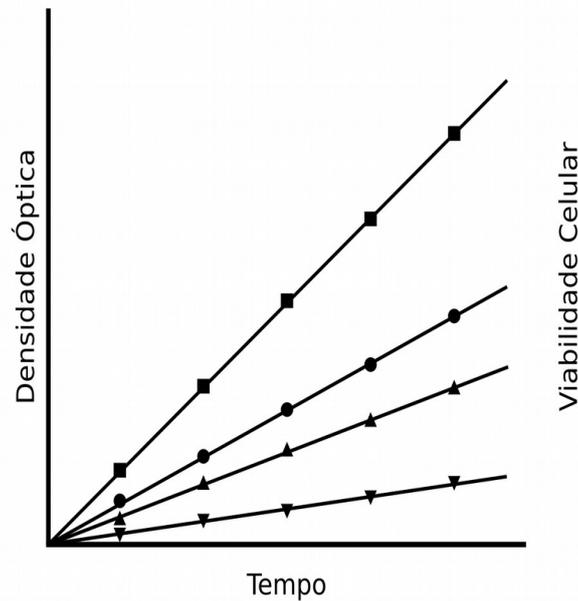


Figura 4: Efeito de várias concentrações de etanol no crescimento de *Saccharomyces cerevisiae*. As linhas de tendência exemplificam o comportamento das células quando inoculadas em meio definido apenas (■) ou em meio definido contendo 5% v/v (●), 7% v/v (▲) ou 10% v/v (▼) de etanol adicionado. Fonte: (CHANDLER et al., 2004).

Assim, o acúmulo de etanol constitui um dos estressores no processo de fermentação. Dentro das células, os efeitos comuns da toxicidade por etanol incluem a inibição da divisão celular, diminuição do seu volume, aumento da taxa de morte e redução da vitalidade das células (JIA; ZHANG; LI, 2010; STANLEY et al., 2010a, 2010b), afetando negativamente a produção em escala industrial. Conseqüentemente, para um maior custo-benefício na produção de bioetanol, são necessárias linhagens com maior tolerância à este composto, ou seja, que tenham um potencial de sobrevivência maior durante uma exposição crônica a esta substância (JIA; ZHANG; LI, 2010; STANLEY et al., 2010a; YU et al., 2012).

A exposição de *S. cerevisiae* ao estresse por etanol afeta primeiramente a integridade da membrana plasmática, alterando sua fluidez e permeabilidade às espécies iônicas levando a uma perda do potencial eletroquímico transmembrana e, em seguida,

acidificando a célula e seus vacúolos (CHANDLER et al., 2004; DING et al., 2009; MA; LIU, 2010; NAVARRO-TAPIA et al., 2016); desse modo, a homeostase intracelular encontra-se seriamente comprometida. Altas concentrações de etanol também podem causar disfunção e desnaturação de proteínas, afetar a absorção de glicose, maltose, amônia e aminoácidos e, finalmente, provocar o escape de nucleotídeos, aminoácidos e potássio das células (DING et al., 2009; MA; LIU, 2010). Análises de *microarray* confirmam que muitos destes sistemas e mecanismos sofrem ajustes em seus perfis de expressão frente ao estresse por etanol (ALEXANDRE et al., 2001; CHANDLER et al., 2004; STANLEY et al., 2010b).

Mais de 200 descrições ontológicas já foram identificadas como participantes na resposta ao estresse ou tolerância ao etanol, incluindo: 1- a biossíntese de moléculas tais como prolina, triptofano e proteínas; 2- o metabolismo de aminoácidos, nucleotídeos, lipídios, ergosterol e ácidos graxos; 3- e os sistemas de transporte, crescimento e ciclo celular (DING et al., 2009; MA; LIU, 2010; STANLEY et al., 2010b; NAVARRO-TAPIA et al., 2016).

Levando em consideração o interesse científico e econômico sobre a característica de tolerância ao etanol (DING et al., 2009), diversos trabalhos de engenharia genética buscam o desenvolvimento de linhagens de *S. cerevisiae* mais resistentes a este estressor (ALPER et al., 2006; TEIXEIRA et al., 2009; LEWIS et al., 2010; MUSSATTO et al., 2010; SWINNEN et al., 2012). Para isso, normalmente são utilizadas estratégias racionais ou randômicas que demandam de tempo e investimentos, tais como o *screening* de fenótipos, mutações e *knockout* aleatórios (JIA; ZHANG; LI, 2010).

Estudos como os acima citados, permitiram reportar que os genes relacionados à tolerância ao etanol normalmente estão associados a mais de uma função (MA; LIU,

cerevisiae. Proteínas codificadas por genes significativamente regulados positivamente estão localizadas na parede celular, membrana, núcleo, mitocôndria e citoplasma. Proteínas de choque térmico (HSPs) são principalmente descritas como chaperonas protegendo e mantendo funções de proteínas em múltiplos locais. A função dos produtos gênicos é classificada com base na ontologia gênica. Fonte: (MA; LIU, 2010).

Os fatos supracitados deixam claro que a relação entre os elementos moleculares-chave que sustentam a aclimatação da levedura ao etanol e àqueles associados a determinar sua tolerância, são ainda pouco esclarecidos (STANLEY et al., 2010b; SWINNEN et al., 2012). Isso provavelmente está relacionado ao fato de os estudos tradicionais avaliarem as funções gênicas abordando poucos genes por vez, desconsiderando que os produtos gênicos atuam como componentes de uma rede integrada (GE; WALHOUT; VIDAL, 2003). O mecanismo de tolerância ao etanol está associado a muitos genes e possui múltiplas camadas de funções e interações complexas (MA; LIU, 2010) (**Figura 5**).

1.3. Biologia de sistemas, transcriptomas e seus empregos no estudo de tolerância ao etanol

Tradicionalmente os estudos de biologia molecular utilizam abordagens reducionistas, as quais objetivam elucidar os princípios biológicos a partir de uma ou poucas moléculas (como genes, proteínas e metabólitos) por vez (GE; WALHOUT; VIDAL, 2003). No entanto, o estudo de moléculas isoladas não permite a compreensão de características multifatoriais, sendo necessário uma abordagem sistêmica de análises (BRUGGEMAN; WESTERHOFF, 2007).

Assim, para aprimorar o conhecimento sobre a função dos módulos que regem os sistemas biológicos de maneira integrada, é imprescindível combinar múltiplos conjuntos de dados (GE; WALHOUT; VIDAL, 2003), comparar os resultados obtidos com outros já existentes e verificar se os resultados se repetem em conjuntos de dados auxiliares, atingindo assim um “padrão ouro” para estas análises (RITCHIE et al., 2015). Por

preencher os requisitos supracitados, a metodologia de análise de redes vem sendo cada vez mais utilizada. Nesta abordagem os vértices (também chamados de nós) geralmente constituem os componentes biológicos e as arestas (também chamadas de conexões) representam as suas interações.

Análises de redes são amplamente utilizadas para descrever diversos sistemas (LUSCOMBE et al., 2004). Dentre suas vantagens, elas independem do tipo de dados e permitem sua aplicação em análises funcionais integradas (PIRUZIAN et al., 2010). Além disso, testes estatísticos sobre essas redes permitem que os resultados geralmente sejam interpretados e comparados, assim os mecanismos capazes de perturbá-los podem ser identificados (CREIXELL et al., 2015).

Trabalhos utilizando metodologias de análise de redes em *S. cerevisiae* foram aplicados em estudos de redes regulatórias (ver (LEE et al., 2002; LUSCOMBE et al., 2004; HERRGARD, 2006)), redes metabólicas (ver (FAMILI et al., 2003)), redes globais (ver (SAID et al., 2004; REGULY et al., 2006; YU et al., 2008)), redes do ciclo celular (LI et al., 2004) e em apenas dois trabalhos abordou-se a tolerância ao etanol (KASAVI et al., 2014, 2016).

Os estudos previamente citados (KASAVI et al., 2014, 2016) exploram as propriedades das redes de maneira estática. No entanto, existe também a necessidade de se compreender a conectividade das redes de maneira dinâmica (LEE et al., 2012). Como exemplo, já foi observado que diferentes tratamentos podem, sinergicamente, alterar o estado intrínseco das células por meio de reconexões dinâmicas nas redes regulatórias (LUSCOMBE et al., 2004; LEE et al., 2012). Desta forma, não é aconselhável desconsiderar que as interações em uma rede biológica variem de acordo com os diferentes estímulos (LUSCOMBE et al., 2004; CREIXELL et al., 2015), o que pode alterar suas medidas topológicas.

Neste contexto, somente uma análise de redes dinâmicas (considerando aqui a dinâmica entre diferentes condições) foi efetuada em *S. cerevisiae* (LUSCOMBE et al., 2004). Neste trabalho os autores verificaram que as respostas ambientais em medições topológicas globais bem conhecidas em redes, tais como os *degrees* de um vértice, indicam alterações dinâmicas transitórias nas interações entre as proteínas nas diferentes condições avaliadas. Estes conceitos de alterações topológicas e estados de transição podem ser aplicados a outras redes biológicas, incluindo sub-sistemas complexos (LUSCOMBE et al., 2004) como é o caso da característica de tolerância ao etanol.

No contexto do uso de análises em *high-throughput* de RNAs visando a tolerância ao etanol, há relatos de estudos realizados com as bactérias *Zymomonas mobilis* ZM4 (YANG et al., 2013), *Clostridium thermocellum* (YANG et al., 2012) e cianobactéria *Synechocystis* sp. (WANG et al., 2012). Dos poucos trabalhos envolvendo transcriptomas com foco na tolerância ao etanol em *S. cerevisiae*, a metodologia empregada foi a de *microarray* (ALEXANDRE et al., 2001; CHANDLER et al., 2004; STANLEY et al., 2010b) e nenhum deles realizou análises sistêmicas a fim de avaliar eventuais diferenças entre o controle e o tratamento com etanol (o mesmo vale para os trabalhos com bactérias e cianobactérias supracitados). Adicionalmente a tecnologia de *microarray* é conhecida por requerer uma alta quantidade de RNA, gerar mais ruídos que o RNA-Seq e não permite o estudo de isoformas e de diferentes alelos (WANG; GERSTEIN; SNYDER, 2009). Por outro lado, análises de RNA-Seq evitam alguns dos problemas citados e apresentam um alcance dinâmico muito maior no que diz respeito aos níveis de expressão (FU et al., 2009; WANG; GERSTEIN; SNYDER, 2009; XU et al., 2013).

No que se refere a redes, Kasavi *et al.* (KASAVI et al., 2014, 2016) utilizando redes de interações entre proteínas (PPIs), redes regulatórias e dados de *microarray*, identificaram 17 novos genes candidatos relacionados a tolerância ao etanol. Análises de

microarray foram efetuadas para verificar a influência da deleção de dois destes genes. No entanto, as linhagens avaliadas possuíam apenas 1% de divergência quanto à tolerância ao etanol, não houve tratamento com etanol antecedendo a extração de RNA no primeiro experimento e, além disso, as três linhagens testadas tinham o mesmo *background* genético da linhagem BY4743. Logo, apesar de obterem resultados interessantes, não houve nenhuma tentativa de compreender a importância do comportamento dinâmico das redes geradas junto aos perfis de expressão obtidos.

Assim, ao observar os estudos anteriores, é notável a abundância de informação disponível a cerca do fenômeno de tolerância ao etanol. Infelizmente, devido a quantidade e complexidade, a análise integrada destes dados torna-se inviável por meio de métodos convencionais. Neste contexto, ferramentas e técnicas são necessárias para lidar com as grandes quantidades de informações biológicas sendo geradas e o *Machine learning* tem se destacado nesse campo por permitir a análise destes conjuntos de dados complexos (CAMACHO et al., 2018).

O *Machine learning* é uma área derivada da inteligência artificial, no qual um programa (algoritmo) aprende a respeito de uma tarefa e otimiza seu aprendizado através de uma medida de performance (DAS et al., 2015) (**Figura 6**). Diferentes algoritmos podem ser usados dependendo do conjunto de dados e da tarefa a ser executada. No caso de uma tarefa de predição na qual previamente sabe-se os rótulos/classes (*labels*) dos dados, pode-se trabalhar com o aprendizado supervisionado (*supervised learning*) para criar modelos de classificação ou regressão. Caso o objetivo seja encontrar estrutura nos dados sem nenhuma informação prévia, pode-se utilizar o aprendizado não-supervisionado (*unsupervised learning*) para executar uma tarefa de agrupamento (*clustering*) (DAS et al., 2015; CAMACHO et al., 2018). Diversos algoritmos de *Machine learning* estão disponíveis, sendo que alguns deles permitem de forma

simples a interpretação dos modelos gerados, tal como observa-se no uso das árvores de decisão. Por outro lado, o uso de redes neurais artificiais geram dados de difícil interpretação (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2006). Finalmente, apesar de técnicas de *Machine learning* estarem sendo aplicadas a diversos tipos de dados biológicos (DAS et al., 2015; CAMACHO et al., 2018), não há nenhum estudo onde estas técnicas tenham sido aplicadas à tolerância ao etanol em *S. cerevisiae*.

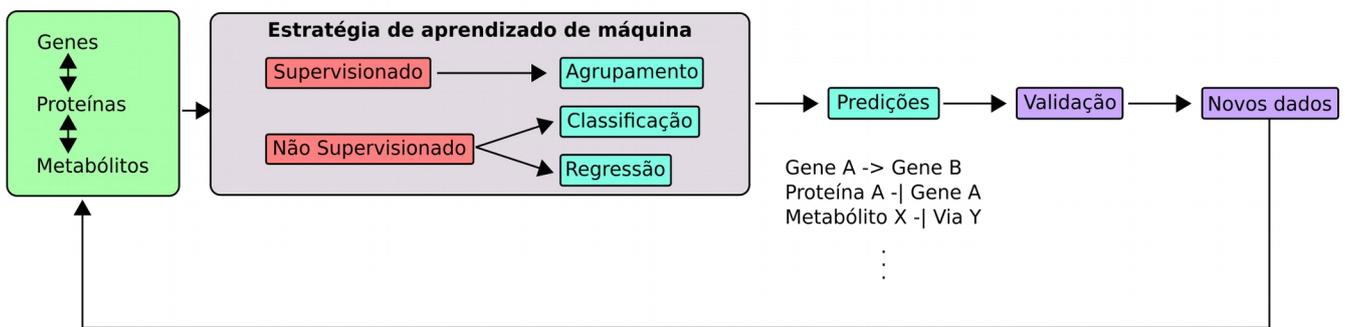


Figura 6: Tipos de dados e técnicas de *Machine learning*. De acordo com os objetivos e os dados fornecidos uma tarefa preditiva (*Supervised learning*) ou uma tarefa de agrupamento (*Unsupervised learning*) pode ser utilizada. Fonte: (CAMACHO et al., 2018).

Deste modo, essa tese objetiva identificar assinaturas sistêmicas associadas à tolerância ao etanol por meio do uso de transcriptoma e metaboloma, técnicas de bioinformática, inteligência artificial, biologia celular e redes dinâmicas, ou seja, análises de alterações dinâmicas nas propriedades dos nós e das conexões das redes em diferentes condições (presença ou ausência de etanol). Experimentalmente, diferentes linhagens de *S. cerevisiae* com diferentes tolerâncias ao etanol foram submetidas à esse estressor. Os transcriptomas e metabolomas dessas linhagens foram obtidos durante o estresse máximo e tais dados foram explorados também em nível sistêmico, integrando dados de diferentes fontes.

7. Referências

- ALEXANDRE, H. et al. Global gene expression during short-term ethanol stress in *Saccharomyces cerevisiae*. **FEBS Letters**, v. 498, n. 1, p. 98–103, jun. 2001.
- ALMEIDA, L. F. de. **Análise das linhagens de *Saccharomyces cerevisiae* expostas ao estresse por etanol**. 2017. Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências de Botucatu, 2017. Disponível em: <<http://hdl.handle.net/11449/152537>>.
- ALPER, H. et al. Engineering yeast transcription machinery for improved ethanol tolerance and production. **Science (New York, N.Y.)**, v. 314, n. 5805, p. 1565–1568, 8 dez. 2006.
- ASLANKOOHI, E. et al. Dynamics of the *Saccharomyces cerevisiae* Transcriptome during Bread Dough Fermentation. **Applied and Environmental Microbiology**, v. 79, n. 23, p. 7325–7333, 1 dez. 2013. Disponível em: <<http://aem.asm.org/cgi/doi/10.1128/AEM.02649-13>>.
- AUESUKAREE, C. Molecular mechanisms of the yeast adaptive response and tolerance to stresses encountered during ethanol fermentation. **Journal of Bioscience and Bioengineering**, v. 124, n. 2, p. 133–142, ago. 2017. Disponível em: <<http://dx.doi.org/10.1016/j.jbiosc.2017.03.009>>.
- AUGUSTYNIAK, A.; SKRZYDLEWSKA, E. I-Carnitine in the lipid and protein protection against ethanol-induced oxidative stress. **Alcohol**, v. 43, n. 3, p. 217–223, maio 2009. Disponível em: <<http://dx.doi.org/10.1016/j.alcohol.2008.12.005>>.
- BALASUNDARAM, D.; TABOR, C. W.; TABOR, H. Spermidine or spermine is essential for the aerobic growth of *Saccharomyces cerevisiae*. **Proceedings of the National Academy of Sciences**, v. 88, n. 13, p. 5872–5876, 1 jul. 1991. Disponível em: <<http://www.pnas.org/cgi/doi/10.1073/pnas.88.13.5872>>.
- BANKEVICH, A. et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455–477, maio 2012. Disponível em: <<http://online.liebertpub.com/doi/abs/10.1089/cmb.2012.0021>>.
- BARRAT, A. et al. The architecture of complex weighted networks. **Proceedings of the National Academy of Sciences**, v. 101, n. 11, p. 3747–3752, 16 mar. 2004. Disponível em: <<http://www.pnas.org/cgi/doi/10.1073/pnas.0400087101>>.
- BEGLEY, U. et al. Trm9-Catalyzed tRNA Modifications Link Translation to the DNA Damage Response. **Molecular Cell**, v. 28, n. 5, p. 860–870, dez. 2007. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1097276507006338>>.
- BLOHM, P. et al. Negatome 2.0: a database of non-interacting proteins derived by literature mining, manual annotation and protein structure analysis. **Nucleic Acids**

- Research**, v. 42, n. D1, p. D396–D400, jan. 2014. Disponível em: <<https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1079>>.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics (Oxford, England)**, v. 30, n. 15, p. 2114–20, 1 ago. 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24695404>>.
- BRUGGEMAN, F. J.; WESTERHOFF, H. V. The nature of systems biology. **Trends in Microbiology**, v. 15, n. 1, p. 45–50, 2007.
- BURT, R. S. Structural Holes and Good Ideas. **American Journal of Sociology**, v. 110, n. 2, p. 349–399, set. 2004. Disponível em: <<http://www.journals.uchicago.edu/doi/10.1086/421787>>.
- BUSKENS, V.; VAN DE RIJT, A. Dynamics of Networks if Everyone Strives for Structural Holes. **American Journal of Sociology**, v. 114, n. 2, p. 371–407, set. 2008. Disponível em: <<http://www.journals.uchicago.edu/doi/10.1086/590674>>.
- BUTLER, G. Fungal Sex and Pathogenesis. **Clinical Microbiology Reviews**, v. 23, n. 1, p. 140–159, 1 jan. 2010. Disponível em: <<http://cmr.asm.org/cgi/doi/10.1128/CMR.00053-09>>.
- CAMACHO, D. M. et al. Next-Generation Machine Learning for Biological Networks. **Cell**, v. 173, n. 7, p. 1581–1592, jun. 2018. Disponível em: <<https://doi.org/10.1016/j.cell.2018.05.015>>.
- CANTAREL, B. L. et al. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. **Genome Research**, v. 18, n. 1, p. 188–196, 2008.
- CARR, M.; BENSASSON, D.; BERGMAN, C. M. Evolutionary Genomics of Transposable Elements in *Saccharomyces cerevisiae*. **PLoS ONE**, v. 7, n. 11, p. e50978, 30 nov. 2012. Disponível em: <<https://dx.plos.org/10.1371/journal.pone.0050978>>.
- CHAKRABORTY, S. et al. Biomass to biofuel: a review on production technology. **Asia-Pacific Journal of Chemical Engineering**, v. 7, n. SUPPL. 3, p. S254–S262, ago. 2012. Disponível em: <<http://doi.wiley.com/10.1002/apj.1642>>.
- CHANDLER, M. et al. A genomic approach to defining the ethanol stress response in the yeast *Saccharomyces cerevisiae*. **Annals of Microbiology**, v. 54, n. 4, p. 427–454, 2004.
- CHAROENBHAKDI, S. et al. Vacuolar H⁺-ATPase Protects *Saccharomyces cerevisiae* Cells against Ethanol-Induced Oxidative and Cell Wall Stresses. **Applied and Environmental Microbiology**, v. 82, n. 10, p. 3121–3130, 15 maio 2016. Disponível em: <<http://aem.asm.org/lookup/doi/10.1128/AEM.00376-16>>.
- CHATR-ARYAMONTRI, A. et al. The BioGRID interaction database: 2015 update. **Nucleic acids research**, v. 43, n. Database issue, p. D470–8, jan. 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/25428363>>.

- CHERRY, J. M. et al. Saccharomyces Genome Database: the genomics resource of budding yeast. **Nucleic acids research**, v. 40, n. Database issue, p. D700-5, 1 jan. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22110037>>.
- CHONG, J.; XIA, J. MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. **Bioinformatics**, v. 34, n. 24, p. 4313–4314, 15 dez. 2018. Disponível em: <<https://academic.oup.com/bioinformatics/article/34/24/4313/5046255>>.
- COSTA, V. et al. Mitochondrial superoxide dismutase is essential for ethanol tolerance of Saccharomyces cerevisiae in the post-diauxic phase. **Microbiology**, v. 143, n. 5, p. 1649–1656, 1 maio 1997. Disponível em: <<http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-143-5-1649>>.
- CREIXELL, P. et al. Pathway and network analysis of cancer genomes. **Nature Methods**, v. 12, n. 7, p. 615–621, 30 jun. 2015. Disponível em: <<http://www.nature.com/doi/10.1038/nmeth.3440>>.
- CSARDI, G.; NEPUSZ, T. The igraph software package for complex network research. **InterJournal**, v. Complex Sy, p. 1695, 2006. Disponível em: <<http://igraph.org>>.
- CUADROS-INOSTROZA, Á. et al. TargetSearch - a Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. **BMC Bioinformatics**, v. 10, n. 1, p. 428, 16 dez. 2009. Disponível em: <<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-10-428>>.
- DAS, S. et al. Applications of Artificial Intelligence in Machine Learning: Review and Prospect. **International Journal of Computer Applications**, v. 115, n. 9, p. 31–41, 22 abr. 2015. Disponível em: <<http://research.ijcaonline.org/volume115/number9/pxc3902402.pdf>>.
- DEMEKE, M. M. et al. Development of a D-xylose fermenting and inhibitor tolerant industrial Saccharomyces cerevisiae strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. **Biotechnology for biofuels**, v. 6, n. 1, p. 89, jan. 2013. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3698012&tool=pmcentrez&rendertype=abstract>>. Acesso em: 27 ago. 2014.
- DEMIRBAS, A. Tomorrow's biofuels: Goals and hopes. **Energy Sources, Part A: Recovery, Utilization, and Environmental Effects**, v. 39, n. 7, p. 673–679, 3 abr. 2017. Disponível em: <<http://dx.doi.org/10.1080/15567036.2016.1252815>>.
- DING, J. et al. **Tolerance and stress response to ethanol in the yeast Saccharomyces cerevisiae** *Applied Microbiology and Biotechnology*, nov. 2009. .
- DONCHEVA, N. T. et al. Topological analysis and interactive visualization of biological networks and protein structures. **Nature Protocols**, v. 7, n. 4, p. 670–685, 2012. Disponível em: <<http://dx.doi.org/10.1038/nprot.2012.004>>.

DOUGLAS, L. M. et al. Membrane Compartment Occupied by Can1 (MCC) and Eisosome Subdomains of the Fungal Plasma Membrane. **Membranes**, v. 1, n. 4, p. 394–411, 13 dez. 2011. Disponível em: <<http://www.mdpi.com/2077-0375/1/4/394>>.

EISENBERG, T. et al. Induction of autophagy by spermidine promotes longevity. **Nature Cell Biology**, v. 11, n. 11, p. 1305–1314, 4 nov. 2009. Disponível em: <<http://dx.doi.org/10.1038/ncb1975>>.

FABREGAT, A. et al. The Reactome Pathway Knowledgebase. **Nucleic Acids Research**, v. 44, n. D1, p. D481–D487, 14 jan. 2014. Disponível em: <<https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1102>>.

FAMILI, I. et al. Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. **Proceedings of the National Academy of Sciences**, v. 100, n. 23, p. 13134–13139, 11 nov. 2003. Disponível em: <<http://www.pnas.org/cgi/doi/10.1073/pnas.2235812100>>.

FERNANDES, L. P. et al. Protein Networks Reveal Detection Bias and Species Consistency When Analysed by Information-Theoretic Methods. **PLoS ONE**, v. 5, n. 8, p. e12083, 18 ago. 2010. Disponível em: <<https://dx.plos.org/10.1371/journal.pone.0012083>>.

FRANKEN, J. et al. Carnitine and carnitine acetyltransferases in the yeast Saccharomyces cerevisiae: a role for carnitine in stress protection. **Current Genetics**, v. 53, n. 6, p. 347–360, 22 jun. 2008. Disponível em: <<http://link.springer.com/10.1007/s00294-008-0191-0>>.

FU, X. et al. Estimating accuracy of RNA-Seq and microarrays with proteomics. **BMC genomics**, v. 10, p. 161, 2009.

GALLONE, B. et al. Domestication and Divergence of Saccharomyces cerevisiae Beer Yeasts. **Cell**, v. 166, n. 6, p. 1397–1410.e16, set. 2016. Disponível em: <<http://dx.doi.org/10.1016/j.cell.2016.08.020>>.

GE, H.; WALHOUT, a. J. M.; VIDAL, M. Integrating “omic” information: A bridge between genomics and systems biology. **Trends in Genetics**, v. 19, n. 10, p. 551–560, 2003.

GULLBERG, J. et al. Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of Arabidopsis thaliana samples in metabolomic studies with gas chromatography/mass spectrometry. **Analytical Biochemistry**, v. 331, n. 2, p. 283–295, ago. 2004. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0003269704003811>>.

GUPTA, A.; VERMA, J. P. Sustainable bio-ethanol production from agro-residues: A review. **Renewable and Sustainable Energy Reviews**, v. 41, p. 550–567, jan. 2015. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1364032114007084>>.

GUREVICH, A. et al. QCAST: quality assessment tool for genome assemblies. **Bioinformatics**, v. 29, n. 8, p. 1072–1075, 15 abr. 2013. Disponível em:

<<https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btt086>>.

HALL, M. et al. The WEKA Data Mining Software : An Update. **SIGKDD Explorations**, v. 11, n. 1, p. 10–18, 2009.

HERRGARD, M. J. Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*. **Genome Research**, v. 16, n. 5, p. 627–635, 1 maio 2006. Disponível em: <<http://www.genome.org/cgi/doi/10.1101/gr.4083206>>.

HOFFMAN, D. E. et al. Changes in diurnal patterns within the *Populus* transcriptome and metabolome in response to photoperiod variation. **Plant, Cell & Environment**, v. 33, n. 8, p. no-no, 18 mar. 2010. Disponível em: <<http://doi.wiley.com/10.1111/j.1365-3040.2010.02148.x>>.

IGNATIUS PANG, C. N.; GOEL, A.; WILKINS, M. R. Investigating the Network Basis of Negative Genetic Interactions in *Saccharomyces cerevisiae* with Integrated Biological Networks and Triplet Motif Analysis. **Journal of Proteome Research**, v. 17, n. 3, p. 1014–1030, 2 mar. 2018. Disponível em: <<http://pubs.acs.org/doi/10.1021/acs.jproteome.7b00649>>.

ISHMAYANA, S.; KENNEDY, U. J.; LEARMONTH, R. P. Further investigation of relationships between membrane fluidity and ethanol tolerance in *Saccharomyces cerevisiae*. **World Journal of Microbiology and Biotechnology**, v. 33, n. 12, p. 218, 27 dez. 2017. Disponível em: <<http://dx.doi.org/10.1007/s11274-017-2380-9>>.

JACKMAN, S. D. et al. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. **Genome Research**, v. 27, n. 5, p. 768–777, maio 2017. Disponível em: <<http://genome.cshlp.org/lookup/doi/10.1101/gr.214346.116>>.

JIA, K.; ZHANG, Y.; LI, Y. Systematic engineering of microorganisms to improve alcohol tolerance. **Engineering in Life Sciences**, v. 10, n. 5, p. 422–429, out. 2010. Disponível em: <<http://doi.wiley.com/10.1002/elsc.201000076>>.

KABECHE, R.; HOWARD, L.; MOSELEY, J. B. Eisosomes provide membrane reservoirs for rapid expansion of the yeast plasma membrane. **Journal of Cell Science**, v. 128, n. 22, p. 4057–4062, 15 nov. 2015. Disponível em: <<http://jcs.biologists.org/cgi/doi/10.1242/jcs.176867>>.

KASAVI, C. et al. A system based network approach to ethanol tolerance in *Saccharomyces cerevisiae*. **BMC Systems Biology**, v. 8, n. 1, p. 90, 2014. Disponível em: <<http://www.biomedcentral.com/1752-0509/8/90>>.

KASAVI, C. et al. An integrative analysis of transcriptomic response of ethanol tolerant strains to ethanol in *Saccharomyces cerevisiae*. **Mol. BioSyst.**, v. 12, n. 2, p. 464–476, 2016. Disponível em: <<http://xlink.rsc.org/?DOI=C5MB00622H>>.

KIM, D.; LANGMEAD, B.; SALZBERG, S. L. HISAT: a fast spliced aligner with low memory requirements. **Nature Methods**, v. 12, n. 4, p. 357–360, 9 mar. 2015. Disponível em: <<http://www.nature.com/doi/10.1038/nmeth.3317>>.

KIM, S.-K. et al. Enhanced ethanol fermentation by engineered *Saccharomyces cerevisiae* strains with high spermidine contents. **Bioprocess and Biosystems Engineering**, v. 40, n. 5, p. 683–691, 24 maio 2017. Disponível em: <<http://link.springer.com/10.1007/s00449-016-1733-3>>.

KIM, S. et al. Evaluation and optimization of metabolome sample preparation methods for *Saccharomyces cerevisiae*. **Analytical Chemistry**, v. 85, n. 4, p. 2169–2176, 2013.

KIM, S. Y.; IMOTO, S.; MIYANO, S. Inferring gene networks from time series microarray data using dynamic Bayesian networks. **Briefings in bioinformatics**, v. 4, n. 3, p. 228–35, 1 set. 2003. Disponível em: <<https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/4.3.228>>.

KLEIN, C. et al. Structural and dynamical analysis of biological networks. **Briefings in Functional Genomics**, v. 11, n. 6, p. 420–433, nov. 2012. Disponível em: <<https://academic.oup.com/bfg/article-lookup/doi/10.1093/bfgp/els030>>.

KOHANY, O. et al. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. **BMC bioinformatics**, v. 7, p. 474, jan. 2006. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1634758&tool=pmcentrez&rendertype=abstract>>. Acesso em: 2 jun. 2014.

KOPKA, J. et al. GMD@CSB.DB: the Golm Metabolome Database. **Bioinformatics**, v. 21, n. 8, p. 1635–1638, 15 abr. 2005. Disponível em: <<https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bti236>>.

KOTSIANTIS, S. B.; ZAHARAKIS, I. D.; PINTELAS, P. E. Machine learning: a review of classification and combining techniques. **Artificial Intelligence Review**, v. 26, n. 3, p. 159–190, 10 nov. 2006. Disponível em: <<http://link.springer.com/10.1007/s10462-007-9052-3>>.

LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature methods**, v. 9, n. 4, p. 357–9, abr. 2012. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3322381&tool=pmcentrez&rendertype=abstract>>. Acesso em: 10 jul. 2014.

LEE, M. J. et al. Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. **Cell**, v. 149, n. 4, p. 780–794, maio 2012. Disponível em: <<http://dx.doi.org/10.1016/j.cell.2012.03.031>>.

LEE, T. I. et al. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. **Science (New York, N.Y.)**, v. 298, n. 5594, p. 799–804, 25 out. 2002. Disponível em: <<http://www.sciencemag.org/cgi/doi/10.1126/science.1075090>>.

- LEWIS, J. a. et al. Exploiting natural variation in *Saccharomyces cerevisiae* to identify genes for increased ethanol resistance. **Genetics**, v. 186, n. 4, p. 1197–1205, 1 dez. 2010.
- LI, B.; DEWEY, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics**, v. 12, n. 1, p. 323, 2011. Disponível em: <<http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-323>>.
- LI, F. et al. The yeast cell-cycle network is robustly designed. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 14, p. 4781–4786, 2004.
- LICATA, L. et al. MINT, the molecular interaction database: 2012 update. **Nucleic Acids Research**, v. 40, n. D1, p. D857–D861, jan. 2012. Disponível em: <<https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr930>>.
- LOHMANN, G. et al. Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain. **PLoS ONE**, v. 5, n. 4, 2010.
- LOVE, M. I.; HUBER, W.; ANDERS, S. **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**bioRxiv. [s.l: s.n.]. Disponível em: <<http://biorxiv.org/content/early/2014/05/27/002832.abstract>>.
- LUO, W. et al. GAGE: generally applicable gene set enrichment for pathway analysis. **BMC bioinformatics**, v. 10, n. 1, p. 161, 27 maio 2009. Disponível em: <<http://www.biomedcentral.com/1471-2105/10/161>>.
- LUO, W.; BROUWER, C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. **Bioinformatics (Oxford, England)**, v. 29, n. 14, p. 1830–1, 15 jul. 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/23740750>>.
- LUSCOMBE, N. M. et al. Genomic analysis of regulatory network dynamics reveals large topological changes. **Nature**, v. 431, n. 7006, p. 308–312, 16 set. 2004.
- MA, M.; LIU, Z. L. **Mechanisms of ethanol tolerance in saccharomyces cerevisiae****Applied Microbiology and Biotechnology**, jul. 2010. . Disponível em: <<http://link.springer.com/10.1007/s00253-010-2594-3>>.
- MANGAN, S.; ZASLAVER, A.; ALON, U. The Coherent Feedforward Loop Serves as a Sign-sensitive Delay Element in Transcription Networks. **Journal of Molecular Biology**, v. 334, n. 2, p. 197–204, nov. 2003. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/14607112>>.
- MANI, R.-S.; CHINNAIYAN, A. M. Triggers for genomic rearrangements: insights into genomic, cellular and environmental influences. **Nature reviews. Genetics**, v. 11, n. 12, p. 819–29, dez. 2010. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21045868>>. Acesso em: 29 out. 2014.
- MARIS, A. F. et al. Diauxic shift-induced stress resistance against hydroperoxides in *Saccharomyces cerevisiae* is not an adaptive stress response and does not depend on

functional mitochondria. **Current Genetics**, v. 39, n. 3, p. 137–149, 16 maio 2001. Disponível em: <<http://link.springer.com/10.1007/s002940100194>>.

MUSSATTO, S. I. et al. **Technological trends, global market, and challenges of bio-ethanol production** *Biotechnology Advances* Elsevier Inc., , nov. 2010. . Disponível em: <<http://dx.doi.org/10.1016/j.biotechadv.2010.07.001>>.

NARAYANAN, A. et al. Defects in Protein Folding Machinery Affect Cell Wall Integrity and Reduce Ethanol Tolerance in *S. cerevisiae*. **Current Microbiology**, v. 73, n. 1, p. 38–45, 18 jul. 2016. Disponível em: <<http://link.springer.com/10.1007/s00284-016-1024-x>>.

NAVARRO-TAPIA, E. et al. Ethanol Cellular Defense Induce Unfolded Protein Response in Yeast. **Frontiers in Microbiology**, v. 7, n. FEB, p. 1–12, 18 fev. 2016. Disponível em: <<http://journal.frontiersin.org/Article/10.3389/fmicb.2016.00189/abstract>>.

NIGAM, P. S.; SINGH, A. Production of liquid biofuels from renewable resources. **Progress in Energy and Combustion Science**, v. 37, n. 1, p. 52–68, fev. 2011. Disponível em: <<http://dx.doi.org/10.1016/j.pecs.2010.01.003>>.

OGATA, H. et al. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, v. 27, n. 1, p. 29–34, 1 jan. 1999. Disponível em: <<https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/27.1.29>>.

PENG, Y. et al. IDBA – A Practical Iterative de Bruijn Graph De Novo Assembler. In: **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**. [s.l.: s.n.]6044 LNBIp. 426–440.

PEREIRA, F. B. et al. Genome-wide screening of *Saccharomyces cerevisiae* genes required to foster tolerance towards industrial wheat straw hydrolysates. **Journal of Industrial Microbiology and Biotechnology**, v. 41, n. 12, p. 1753–1761, 2014.

PÉREZ-GALLARDO, R. V. et al. Reactive oxygen species production induced by ethanol in *Saccharomyces cerevisiae* increases because of a dysfunctional mitochondrial iron-sulfur cluster assembly system. **FEMS Yeast Research**, v. 13, n. 8, p. 804–819, dez. 2013. Disponível em: <<https://academic.oup.com/femsyr/article-lookup/doi/10.1111/1567-1364.12090>>.

PIRUZIAN, E. et al. Integrated network analysis of transcriptomic and proteomic data in psoriasis. **BMC systems biology**, v. 4, p. 41, 2010.

PISKUR, J. et al. How did *Saccharomyces* evolve to become a good brewer? **Trends in Genetics**, v. 22, n. 4, p. 183–186, abr. 2006. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0168952506000333>>.

PRONK, J. T.; YDE STEENSMA, H.; VAN DIJKEN, J. P. Pyruvate Metabolism in *Saccharomyces cerevisiae*. **Yeast**, v. 12, n. 16, p. 1607–1633, dez. 1996. Disponível em: <<http://doi.wiley.com/10.1002/%28SICI%291097-0061%28199612%2912%3A16%3C1607%3A%3AAID-YEA70%3E3.0.CO%3B2-4>>.

QUINLAN, A. R.; HALL, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. **Bioinformatics**, v. 26, n. 6, p. 841–842, 2010.

R CORE TEAM. **R: A Language and Environment for Statistical Computing** Vienna, Austria, 2015. Disponível em: <<http://www.r-project.org>>.

REGULY, T. et al. Comprehensive curation and analysis of global interaction networks in *Saccharomyces cerevisiae*. **Journal of biology**, v. 5, n. 4, p. 11, 2006. Disponível em: <<http://jbiol.com/content/5/4/11>>.

REIMAND, J. et al. g:Profiler—a web server for functional interpretation of gene lists (2016 update). **Nucleic Acids Research**, v. 44, n. W1, p. W83–W89, 8 jul. 2016. Disponível em: <<https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw199>>.

RICHARDS, V. P. et al. Transcriptomic and genomic evidence for *Streptococcus agalactiae* adaptation to the bovine environment. **BMC genomics**, v. 14, p. 920, jan. 2013. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3890567&tool=pmcentrez&rendertype=abstract>>.

RITCHIE, M. D. et al. Methods of integrating data to uncover genotype – phenotype interactions. **Nature Publishing Group**, n. January, 2015. Disponível em: <<http://dx.doi.org/10.1038/nrg3868>>.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139–140, 1 jan. 2010. Disponível em: <<https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp616>>.

ROLLAND, T. et al. A Proteome-Scale Map of the Human Interactome Network. **Cell**, v. 159, n. 5, p. 1212–1226, nov. 2014. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0092867414014226>>.

SAID, M. R. et al. Global network analysis of phenotypic effects: protein networks and toxicity modulation in *Saccharomyces cerevisiae*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 52, p. 18006–18011, 28 dez. 2004.

SCHAFFTER, T.; MARBACH, D.; FLOREANO, D. GeneNetWeaver: in silico benchmark generation and performance profiling of network inference methods. **Bioinformatics**, v. 27, n. 16, p. 2263–2270, 15 ago. 2011. Disponível em: <<https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btr373>>.

SCHAUER, N. et al. GC-MS libraries for the rapid identification of metabolites in complex biological samples. **FEBS Letters**, v. 579, n. 6, p. 1332–1337, 28 fev. 2005. Disponível em: <<http://doi.wiley.com/10.1016/j.febslet.2005.01.029>>.

SONG, G. et al. AGAPE (Automated Genome Analysis PipelinE) for Pan-Genome Analysis of *Saccharomyces cerevisiae*. **PLOS ONE**, v. 10, n. 3, p. e0120671, 17 mar. 2015. Disponível em: <<http://dx.plos.org/10.1371/journal.pone.0120671>>.

STANLEY, D. et al. The ethanol stress response and ethanol tolerance of *Saccharomyces cerevisiae*. **Journal of applied microbiology**, v. 109, n. 1, p. 13–24, jul. 2010a. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/20070446>>.

STANLEY, D. et al. Transcriptional changes associated with ethanol tolerance in *Saccharomyces cerevisiae*. **Applied Microbiology and Biotechnology**, v. 88, n. 1, p. 231–239, set. 2010b.

STEFANINI, I. et al. A Systems Biology Approach to Dissection of the Effects of Small Bicyclic Peptidomimetics on a Panel of *Saccharomyces cerevisiae* Mutants. **Journal of Biological Chemistry**, v. 285, n. 30, p. 23477–23485, 23 jul. 2010. Disponível em: <<http://www.jbc.org/lookup/doi/10.1074/jbc.M110.125153>>.

SUPEK, F. et al. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. **PLoS ONE**, v. 6, n. 7, p. e21800, 18 jul. 2011. Disponível em: <<https://dx.plos.org/10.1371/journal.pone.0021800>>.

SWINNEN, S. et al. Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. **Genome Research**, v. 22, n. 5, p. 975–984, 1 maio 2012.

TAKAGI, H.; TAGUCHI, J.; KAINO, T. Proline accumulation protects *Saccharomyces cerevisiae* cells in stationary phase from ethanol stress by reducing reactive oxygen species levels. **Yeast**, v. 33, n. 8, p. 355–363, ago. 2016. Disponível em: <<http://doi.wiley.com/10.1002/yea.3154>>.

TARAILO-GRAOVAC, M.; CHEN, N. Using RepeatMasker to identify repetitive elements in genomic sequences. **Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.]**, v. Chapter 4, n. March, p. Unit 4.10, mar. 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19274634>>. Acesso em: 28 maio. 2014.

TEIXEIRA, M. C. et al. Genome-Wide Identification of *Saccharomyces cerevisiae* Genes Required for Maximal Tolerance to Ethanol. **Applied and Environmental Microbiology**, v. 75, n. 18, p. 5761–5772, 15 set. 2009. Disponível em: <<http://aem.asm.org/cgi/doi/10.1128/AEM.00845-09>>.

VALENTE, G. T. et al. The development of a universal in silico predictor of protein-protein interactions. **PloS one**, v. 8, n. 5, p. e65587, jan. 2013. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3669264&tool=pmcentrez&rendertype=abstract>>. Acesso em: 3 nov. 2014.

VERES, D. V. et al. ComPPI: a cellular compartment-specific database for protein-protein interaction network analysis. **Nucleic acids research**, v. 43, n. Database issue, p. D485–93, 28 jan. 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/25348397>>.

WANG, J. et al. RNA-seq based identification and mutant validation of gene targets related to ethanol resistance in cyanobacterial *Synechocystis* sp. PCC 6803. **Biotechnology for Biofuels**, v. 5, n. 1, p. 89, 2012. Disponível em:

<<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3564720&tool=pmcentrez&rendertype=abstract>>.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature reviews. Genetics**, v. 10, n. 1, p. 57–63, jan. 2009. Disponível em: <<http://www.nature.com/doi/10.1038/nrg2484>>.

WASSERMAN, S.; FAUST, K. **Social Network Analysis: Methods and Applications**. Cambridge: Cambridge University Press, 1994.

WU, T. D.; WATANABE, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. **Bioinformatics**, v. 21, n. 9, p. 1859–1875, 1 maio 2005. Disponível em: <<https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bti310>>.

XU, K. et al. Path lengths in protein-protein interaction networks and biological complexity. **PROTEOMICS**, v. 11, n. 10, p. 1857–1867, maio 2011. Disponível em: <<http://doi.wiley.com/10.1002/pmic.201000684>>.

XU, X. et al. Parallel comparison of Illumina RNA-Seq and Affymetrix microarray platforms on transcriptomic profiles generated from 5-aza-deoxy-cytidine treated HT-29 colon cancer cells and simulated datasets. **BMC bioinformatics**, v. 14 Suppl 9, n. Suppl 9, p. S1, 2013. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3697991&tool=pmcentrez&rendertype=abstract>>.

YANG, S. et al. **Clostridium thermocellum ATCC27405 transcriptomic, metabolomic and proteomic profiles after ethanol stress** **BMC Genomics** **BMC Genomics**, , 2012. . Disponível em: <[BMC Genomics](#)>.

YANG, S. et al. Systems Biology Analysis of *Zymomonas mobilis* ZM4 Ethanol Stress Responses. **PLoS ONE**, v. 8, n. 7, p. 89, 2013. Disponível em: <<http://www.biotechnologyforbiofuels.com/content/5/1/89>>.

YANG, T.-H. et al. YTRP: a repository for yeast transcriptional regulatory pathways. **Database : the journal of biological databases and curation**, v. 2014, n. November, p. bau014, 2014. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24608172>>.

YOSHIKAWA, K. et al. Comprehensive phenotypic analysis for identification of genes affecting growth under ethanol stress in *Saccharomyces cerevisiae*. **FEMS Yeast Research**, v. 9, n. 1, p. 32–44, fev. 2009. Disponível em: <<https://academic.oup.com/femsyr/article-lookup/doi/10.1111/j.1567-1364.2008.00456.x>>.

YU, H. et al. High-Quality Binary Protein Interaction Map of the Yeast Interactome Network. **Science**, v. 322, n. 5898, p. 104–110, 3 out. 2008. Disponível em: <<http://www.sciencemag.org/cgi/doi/10.1126/science.1158684>>.

YU, K. O. et al. Increased ethanol production from glycerol by *Saccharomyces cerevisiae* strains with enhanced stress tolerance from the overexpression of SAGA complex components. **Enzyme and Microbial Technology**, v. 51, n. 4, p. 237–243, set. 2012. Disponível em: <<http://dx.doi.org/10.1016/j.enzmictec.2012.07.003>>.

ZERBINO, D. R.; BIRNEY, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. **Genome Research**, v. 18, n. 5, p. 821–829, 21 fev. 2008. Disponível em: <<http://www.genome.org/cgi/doi/10.1101/gr.074492.107>>.

ZHANG, Y. et al. The importance of engineering physiological functionality into microbes. **Trends in Biotechnology**, v. 27, n. 12, p. 664–672, dez. 2009.

ZHAO, F. et al. Enhancing *Saccharomyces cerevisiae* reactive oxygen species and ethanol stress tolerance for high-level production of protopanaxadiol. **Bioresource Technology**, v. 227, p. 308–316, mar. 2017. Disponível em: <<http://dx.doi.org/10.1016/j.biortech.2016.12.061>>.

ZYRINA, A. N. et al. Mitochondrial Superoxide Dismutase and Yap1p Act as a Signaling Module Contributing to Ethanol Tolerance of the Yeast *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 83, n. 3, p. 1–13, 1 fev. 2017. Disponível em: <<http://aem.asm.org/lookup/doi/10.1128/AEM.02759-16>>.