

Universidade Estadual Paulista “Júlio de Mesquita Filho”

Instituto de Biociências de Botucatu

Programa de Pós-Graduação em Biotecnologia

Amanda Piveta Schnepper

Caracterização funcional dos *long non-coding* RNAs (lncRNAs) de
Saccharomyces cerevisiae na tolerância ao etanol

Botucatu

2022

Universidade Estadual Paulista “Júlio de Mesquita Filho”

Instituto de Biociências de Botucatu

Programa de Pós-Graduação em Biotecnologia

Mestranda: Amanda Piveta Schnepfer

Orientador: Guilherme Targino Valente

**Caracterização funcional dos *long non-coding* RNAs (lncRNAs) de
Saccharomyces cerevisiae na tolerância ao etanol**

Dissertação apresentada à
Universidade Estadual Paulista,
Instituto de Biociências de Botucatu,
para obtenção da qualificação de
mestrado do Programa de Pós-
Graduação em Biotecnologia, na área
de Biotecnologia

Botucatu

2022

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Schnepper, Amanda Piveta.

Caracterização funcional dos *long non-coding* RNAs (lncRNAs) de *Saccharomyces cerevisiae* na tolerância ao etanol / Amanda Piveta Schnepper. - Botucatu, 2022

Dissertação (mestrado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Guilherme Targino Valente

Capes: 90400003

1. *Saccharomyces cerevisiae*. 2. lncRNAs. 3. Etanol.
4. Leveduras. 5. Engenharia genética.

Palavras-chave: *Long non-coding* RNA; *Saccharomyces cerevisiae*; Tolerância ao etanol.

Agradecimentos

Nas palavras de Agatha Christie:

*“A essência da vida é andar para a frente;
sem possibilidade de fazer ou intentar marcha a trás.”*

Porém, andar para frente nem sempre é fácil, e nesses momentos de dúvidas e dificuldade, temos a sorte de ter pessoas boas ao nosso lado. Mudar de rotina, de casa, de cidade, só foi possível por ter o apoio e a compreensão da minha mãe e do meu pai. Graças a eles, sei que sempre vou ter um “lugar para voltar”, um lugar para descansar, e a certeza de que não estou sozinha. Aqui agradeço também ao meu irmão, minha vó, minhas tias, que mesmo que distantes se fizeram presentes nesta caminhada.

Durante o mestrado, tive a sorte de encontrar bons amigos, que fizeram parte deste trabalho dentro e fora do laboratório. Agradeço ao Luiz, Camila, Farinazzo, Lazari, Flávia, Fabi e Leo, porque sem vocês talvez isso não tivesse sido possível.

Agradeço também ao Prof Guilherme, que como meu orientador me fez superar os desafios, buscar soluções para os problemas (que foram muitos), e fazer este projeto acontecer. Sou grata por ter tido a oportunidade de trabalhar no SBGL, pois aqui encontrei uma realidade muito diferente daquela que eu conhecia, e que com certeza me fez uma pessoa e uma profissional melhor.

Agradeço ao Ivan, que esteve do meu lado nos momentos mais difíceis, que me ajudou e me apoiou mesmo quando o caminho era árduo. Você fez a jornada ser mais colorida.

Agradeço também à CAPES, UNESP e à toda equipe do IBB e do programa de pós-graduação em Biotecnologia.

Juntos passamos por dificuldades, por uma pandemia, mudanças de laboratório, e todos os problemas que são possíveis de imaginar, mas estes momentos difíceis só fazem com que valorizemos mais as boas amizades. Eu agradeço a todos vocês do fundo do meu coração, e dou a certeza de que os levarei comigo para sempre.

*“Respeito, educação e liberdade.
As pessoas têm de ser livres para
serem felizes como elas são.”*

- Rita Lee

Resumo

Saccharomyces cerevisiae é a levedura responsável por grande parte dos processos fermentativos do cotidiano, principalmente a produção de etanol como combustível. Dado o aumento da demanda de uso do etanol, cada vez mais a engenharia genética tem sido usada para entender o metabolismo das leveduras, elucidar as regulações envolvidas nas vias de produção de etanol e melhorar o desempenho das leveduras na fermentação. Neste trabalho, iremos explorar os lncRNA responsivos ao estresse por etanol do ponto de vista funcional. LncRNAs são moléculas de RNA com mais de 200 nt e que normalmente não codificam proteínas. Estes RNAs atuam regulando o metabolismo das células por meio de regulações transcricionais ou traducionais, uma vez que possuem a capacidade de interagir com DNA, RNA e proteínas. Experimentos de modelagem em tripla-hélice mostraram um potencial dos lncRNAs em interagir com a região promotora de diversos genes de *S. cerevisiae*, atuando tanto na sobre-expressão quanto na sub-expressão destes genes. O KO do transcr_9136 reduziu a tolerância ao etanol e alterou a expressão dos genes RRP1 e ILT1, vizinhos a este lncRNA. Na linhagem BY4742, o KO do transcr_10027 resultou em alterações no crescimento, redução da tolerância ao etanol e diferenças significativas nos mecanismos de resposta ao estresse, como a formação de *P-bodies* e grânulos de estresse. Desta forma, reporta-se a importância das regulações mediadas por lncRNAs no metabolismo das leveduras, o que pode contribuir para estudos posteriores e orientar experimentos de engenharia genética.

Palavras-chave: *Saccharomyces cerevisiae*, lncRNAs, tolerância ao etanol

Abstract

Saccharomyces cerevisiae is a yeast responsible for the ordinary fermentation processes, mainly concerning the ethanol-fuel production. The increasing demand for ethanol prompts the use of genetic engineering to understand yeast metabolism, elucidate how regulations are involved in ethanol production, and to improve ethanol yield. Herein, we explore the ethanol stress responsive lncRNAs from a functional view. The lncRNAs are RNAs larger than 200 nt that usually do not code proteins. These RNAs work on regulating the cell metabolism by transcriptional or translational regulation. The ones are also able to interact with DNA, RNA and proteins. Triple-helix interaction experiments revealed that lncRNAs may interact with promoter region of many *S. cerevisiae* genes to either up or down regulate the genes. The transcript_9136 KO albeit the ones had reduced ethanol tolerance and exhibited RRP1 and ILT1 expression changes: these genes are close to that lncRNA. In the BY4742 strain, the transcr_10027 KO lead to growth changes, but reduced the ethanol tolerance and caused differences in stress response mechanisms such as P-bodies and stress granules formation. Thus, the importance of lncRNA-mediated regulations in yeast metabolism is here shown, which may contribute to further studies by guiding genetic engineering experiments.

Keywords: *Saccharomyces cerevisiae*, lncRNAs, ethanol tolerance

Sumário

LISTA DE ABREVIATURAS	9
INTRODUÇÃO	10
Bioetanol como uma alternativa sustentável a matriz energética mundial	10
<i>S. cerevisiae</i> e a tolerância ao etanol	10
lncRNAs e o metabolismo: a chave para a resistência ao etanol pode estar em um RNA não codificante	12
CRISPR-Cas9, uma ferramenta importante na engenharia genética	13
HIPÓTESE	14
Capítulo 1	15
INTRODUÇÃO	15
Estocagem e degradação de moléculas: os lncRNAs como mediadores do ciclo das moléculas dentro da célula	15
Artigo científico a ser submetido	18
Introduction	18
MATERIAL AND METHODS	20
Prediction of lncRNAs-Protein Interaction Network	20
Generating the <i>Saccharomyces cerevisiae</i> mutants	20
Gene expression in mutant strains SEY6210 <i>transcr_9136Δ</i> and BY4742 <i>transcr_10027Δ</i>	24
Protein extraction and western blot analysis for strains BY4742 WT, BY4742 <i>transcr_10027Δ</i> , SEY6210 WT and SEY6210 <i>transcr_9136Δ</i>	25
Analysis of ethanol tolerance of mutant yeasts	26
RESULTS	26
Analysis of the impact of lncRNAs deletion on metabolism	26
DISCUSSION	30
<i>Transcr_10027</i> as one of the stress response regulators through the formation of PBs and SGs	30
The <i>transcr_9136</i> lncRNA as a potential regulator of the translational apparatus	31
CAPÍTULO 2	35
INTRODUÇÃO	35
lncRNAs como reguladores transcricionais, pós-transcricionais e traducionais	35
Bioinformática e o estudo de regulação transcricional pela formação de triplas hélices (RNA:DNA:DNA)	36
MATERIAIS E MÉTODOS	37
Obtenção dos lncRNAs do transcriptoma de <i>Saccharomyces cerevisiae</i>	37
Análise da formação de triplas hélices das linhagens de <i>Saccharomyces cerevisiae</i>	38
RESULTADOS	39
DISCUSSÃO	40

CONSIDERAÇÕES FINAIS.....	43
REFERÊNCIAS	44
APÊNDICE A.....	50
APÊNDICE B.....	51
APÊNDICE C.....	52
Comprovação da presença dos lncRNAs e avaliação da expressão gênica	52
Confirmação da presença dos lncRNAs em <i>Saccharomyces cerevisiae</i>	54

LISTA DE ABREVIATURAS

3'UTR	do inglês 3' untranslated region
cDNA	do inglês Complementary DNA
CT	do inglês <i>Cycle Threshold</i>
DNA	do inglês <i>Desoxyribonucleic acid</i>
dNTP	do inglês <i>deoxyribonucleotide triphosphate</i>
eRNA	do inglês RNA enhancer
gRNA	do inglês <i>guide RNA</i>
KO	do inglês <i>knock out</i>
LB	meio Luria-Bertani
lncRNA	do inglês <i>Long non-coding RNA</i>
miRNA	micro RNA
mRNA	RNA mensageiro
ncRNA	do inglês non-coding RNA
nt	Nucleotídeos
OD ₆₀₀	Densidade óptica em 600nm
pb	Pares de base
PB	do inglês <i>P-body</i>
PCR	do inglês <i>Polymerase chain reaction</i>
PDNR	Desubiquitinação de proteínas e regulação negativa de ubiquitinação
PPPR	Poliubiquitinação de proteínas e regulação positiva de ubiquitinação
PSG	do inglês Proteasome Stress Granules
RNA	do inglês <i>Ribonucleic acid</i>
RNAseq	do inglês RNA next-generation sequencing
RNP	Ribonucleoproteínas
RPM	Rotações por minuto
rRNA	RNA ribossomal
RT-qPCR	do inglês <i>Reverse transcription polymerase chain reaction</i>
SG	Grânulos de estresse
TFBS	Sítio de ligação para fatores de transcrição
TFD	do inglês Transcription Factor Decoy
TFO	do inglês triple-helix forming oligonucleotides
WT	do inglês Wild Type
YPD	do inglês Yeast extract, peptone and dextrose medium

INTRODUÇÃO

Bioetanol como uma alternativa sustentável a matriz energética mundial

Atualmente, os combustíveis fósseis são as principais fontes de energia mundial, mas por diversos motivos (DELLOMONACO; FAVA; GONZALEZ, 2010; MUSSATTO et al., 2010), vários países têm focado no aprimoramento dos biocombustíveis (CHENG; TIMILSINA, 2011). O bioetanol é o biocombustível mais comercializado, sendo o Brasil um pioneiro na produção desse composto com finalidade automobilística e o segundo maior produtor do mundo (LOPES et al., 2016)

As características do etanol possibilitam a combustão mais limpa e o melhor desempenho dos motores, contribuindo para a redução das emissões de gases de efeito estufa (GOLDEMBERG; COELHO; GUARDABASSI, 2008). A produção do bioetanol é feita a partir da fermentação dos açúcares presentes nas matérias primas, sendo a sacarose o açúcar mais comum. Industrialmente, utiliza-se os fungos *Saccharomyces cerevisiae* para tal processo fermentativo (YOU; ROSENFELD; KNIPPLE, 2003). A levedura do gênero *Saccharomyces* é um microrganismo de metabolismo aeróbio facultativo, apresenta alta capacidade fermentativa, alta taxa de crescimento em condições anaeróbias e relativa alta tolerância ao etanol e outros inibidores (MUSSATTO et al., 2010). Enquanto uma porção do açúcar é transformada pela *S. cerevisiae* em biomassa, CO₂ e H₂O, a maior parte é convertida em etanol e CO₂ em anaerobiose, processo denominado de fermentação alcoólica (LIMA; TEIXEIRA, 2001).

S. cerevisiae e a tolerância ao etanol

Para uma boa produção de bioetanol, as cepas devem ser hábeis em sobreviver sob condições industriais (SNOWDON et al., 2009), principalmente em altas concentrações de etanol. O etanol é conhecido como causador de inibição no crescimento de micro-

organismos, implicando em limitação no processo produtivo deste combustível. Porém, existem estudos com linhagens de *S. cerevisiae* que reportam linhagens mais tolerantes ao etanol (YOU; ROSENFELD; KNIPPLE, 2003).

A tolerância ao etanol é o potencial de sobrevivência das células durante uma exposição crônica à essa substância (STANLEY et al., 2010a); foi sugerido que a *S. cerevisiae* tolera altas concentrações de etanol porque evoluiu a capacidade de produzir este composto a partir de açúcar. Na última fase de fermentação industrial, o etanol atinge os níveis tóxicos para a levedura (SNOEK; VERSTREPEN; VOORDECKERS, 2016), causando uma diminuição da viabilidade celular e, conseqüentemente, uma queda de produtividade. Assim, para fins industriais, são utilizadas linhagens mais resistentes, as quais vêm sendo obtidas através de seleção e mutações dirigidas (STANLEY et al., 2010a; ZHANG et al., 2012).

Os efeitos da toxicidade do etanol sobre a *S. cerevisiae* são a inibição do crescimento celular, redução de viabilidade, altas taxas de morte celular, diminuição das taxas de acumulação de RNA e proteínas, aumento de mutações, mudanças no metabolismo e desnaturação de proteínas e enzimas. Os principais alvos do etanol nas células fúngicas são a membrana celular, proteínas e o retículo endoplasmático (STANLEY et al., 2010b).

Apesar dos bons resultados gerados por meio de trabalhos de engenharia genética visando aumentar a tolerância ao etanol e do volume de conhecimento acumulado acerca dos genes e proteínas envolvidos nesse processo, a relação entre esses fatores é ainda pouco esclarecida. Por fim, as escolhas dos alvos para os trabalhos de aumento de tolerância ao etanol enfrentam um problema na seleção de genes relacionados a esse estresse.

LncRNAs e o metabolismo: a chave para a resistência ao etanol pode estar em um RNA não codificante

Considera-se atualmente que pelo menos 75% do genoma de *S. cerevisiae* é transcrito e que muitos dos transcritos não são codificantes (YAMASHITA; SHICHINO; YAMAMOTO, 2016). Os RNAs não codificantes (ncRNAs) são entendidos como sequências responsáveis por agir em mecanismos funcionais, regulatórios e estruturais dos organismos, atuando principalmente na regulação da expressão gênica (STATELLO et al., 2021). Os miRNAs, possuem cerca de 22 nucleotídeos e estão relacionados com a regulação pós-transcricional de genes (SAINI; ENRIGHT; GRIFFITHS-JONES, 2008). Os piRNAs possuem 24-30 nucleotídeos e ligam-se às proteínas responsáveis por manter a estabilidade do genoma em células germinativas. Os snoRNAs possuem de 60 a 300 pares de base e são componentes das ribonucleoproteínas nucleolar pequenas (snoRNPs) (GOMES et al., 2017).

Uma outra classe de ncRNAs são os *long non-coding RNAs* (lncRNAs), um grupo de transcritos não codificantes maiores do que 200 nucleotídeos e conhecidos por atuarem na dinâmica cromossomal, na biologia dos telômeros e na organização estrutural sub-celular. Eles também podem atuar como co-fatores para modular a atividade de fatores de transcrição, regular a RNA polimerase, promover a regulação ao acesso à cromatina, a regulação epigenética, a expressão de mRNAs, atuar como capturadores de microRNAs e também podem codificar micro-peptídeos (ANDERSON et al., 2015). Os lncRNAs podem ser classificados de acordo com as suas posições no genoma, os quais podem ser intrônicos, intergênicos ou até mesmo sobrepor regiões codificantes (CHEKANOVA, 2015).

No caso da *S. cerevisiae*, os lncRNAs estão envolvidos em mudanças metabólicas, iniciação da diferenciação sexual e outros processos ainda pouco esclarecidos (YAMASHITA; SHICHINO; YAMAMOTO, 2016). Além disso, estudos em *Saccharomyces*

relacionam lncRNA com a interação proteína-RNA, atuando como mediador e suporte na ligação de proteínas específicas no RNA, podendo ser um mecanismo de resposta ao estresse (GALIPON et al., 2013). Dessa forma, o estudo dos lncRNA em cepas com alta tolerância ao etanol pode ser uma boa estratégia para encontrar mecanismos responsáveis por essa resistência.

CRISPR-Cas9, uma ferramenta importante na engenharia genética

A biologia sintética atua como um complemento importante da biologia de sistemas concentrando-se na edição do genoma com a finalidade de redirecionar o metabolismo, favorecendo a origem de novos fenótipos.

Atualmente, busca-se reduzir o número de genes-alvo para experimentos laboratoriais devido às dificuldades em trabalhar experimentalmente com um alto número de genes simultaneamente. Porém, isso limita os estudos pois nem sempre produz-se resultados satisfatórios, uma vez que a maioria das características fenotípicas é determinada por herança poligênica. Dessa forma, a tecnologia CRISPR-Cas9 fornece uma vantagem clara quanto à multiplexação, uma vez que interferências múltiplas podem ser introduzidas no genoma simultaneamente (RYAN; PODDAR; CATE, 2016).

CRISPR-Cas9 foi originalmente identificada como parte do sistema imune de bactérias que utiliza nucleases guiadas por RNA para clivar elementos genéticos “estranhos” àqueles organismos. Atualmente os princípios gerais de seu mecanismo de funcionamento são utilizados para a edição de genes no genoma de diversos tipos de organismos. Essa tecnologia consiste no uso de um RNA guia (gRNA) ligado a uma endonuclease Cas9, o qual direciona a clivagem do DNA alvo por meio do pareamento de

20 nucleotídeos (JINEK et al., 2012). Qualquer locus genômico que esteja seguido pela sequência 5'-NGG-3' pode ser alvo de integração ou “gene disrupting” pela complementariedade entre ~20nt de sequência de gRNA complementar e o alvo. Esse sistema de edição gênica é uma ferramenta específica, simples de construir, e vem sendo utilizada com sucesso em *S. cerevisiae*, incluindo em reengenharia-multiplex de rotas metabólicas (JAKOČINAS et al., 2015)

HIPÓTESE

A avaliação da expressão dos diversos lncRNAs e de seus alvos em 6 linhagens de *S. cerevisiae* permitiu criarmos a hipótese de que a ligação de um lncRNA a uma determinada proteína atua como um tipo de “buffer”, gerando um *feedback* negativo na expressão desses alvos por um novo mecanismo proposto (WOLF et al., 2021). Nesse contexto, a hipótese da presente proposta é que alguns lncRNAs estejam negativamente afetando a tolerância ao etanol em *S. cerevisiae* pela sua atuação como “buffer”.

Capítulo 1

INTRODUÇÃO

Estocagem e degradação de moléculas: os lncRNAs como mediadores do ciclo das moléculas dentro da célula

Reportou-se que os lncRNAs têm a capacidade de interagir com diversas proteínas reguladoras do metabolismo de *S. cerevisiae* (WOLF et al., 2021). Na Figura 1, destaca-se várias interações entre lncRNAs e proteínas relacionadas às vias de degradação de moléculas, como formação de *P-bodies* (PBs), proteassomos e poliubiquitinação de proteínas e regulação positiva de ubiquitinação (PPPR). Além das vias de degradação, observa-se também interações com proteínas presentes nas vias de estocagem de moléculas, como a formação de grânulos de estresse (SGs), grânulos de *stress* do proteassomo (PSGs) e desubiquitinação de proteínas e regulação negativa de ubiquitinação (PDNR).

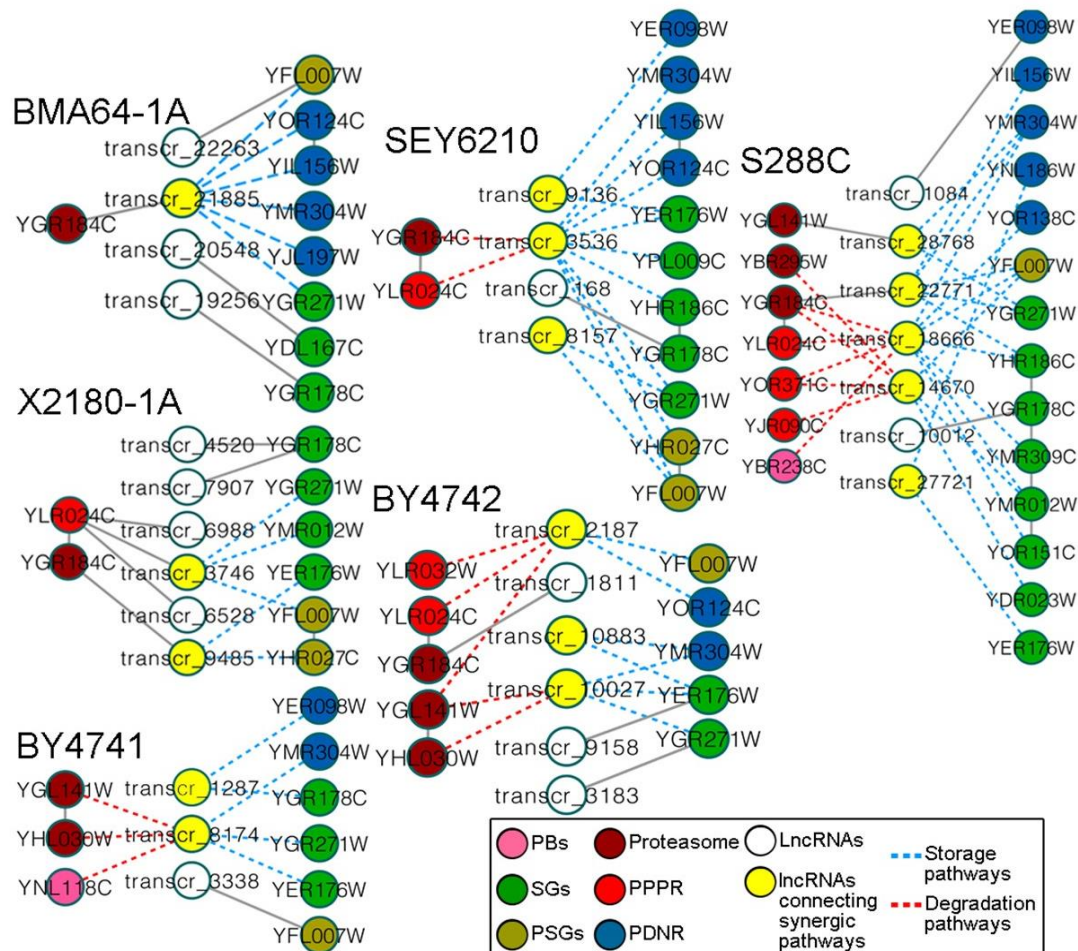


Figura 1: Subredes de interação lncRNA-proteína com foco nas vias de degradação e estocagem, separados por linhagem. Adaptado de (WOLF et al., 2021)

Células eucarióticas possuem complexos de ribonucleoproteínas (RNP), que atuam regulando a tradução, degradação e localização de mRNAs (SHAH et al., 2013). Duas importantes variações de RNPs são os PBs e os SGs, sendo que, nas linhagens estudadas estes mecanismos parecem sofrer um efeito regulatório pelos lncRNAs

Os PBs são estruturas conservadas de resposta ao estresse presentes desde plantas, leveduras, até mamíferos (BUCHAN; MUHLRAD; PARKER, 2008). Os PBs possuem em sua composição um aglomerado de proteínas de ligação ao RNA (AGO1, AGO2, ELAVL1, FRMR1, HNRNPM, HNRNPU, IGF2BP1, IGF2BP2, IGF2BP3, MOV10, PUM2, e ZC3H7B) o que favorece o endereçamento dos mRNAs para dentro dos PBs durante o estresse, além de uma importante exonuclease, a IXR1, sendo que estes mRNAs parecem não entrar nas vias de degradação por possuírem extremidades 5' mais protegidas

em comparação com os mRNAs livres (D et al., 2007; HUBSTENBERGER et al., 2017). Estudos demonstram uma correlação direta entre o número de PBs e o tempo de vida da célula durante a fase estacionária (SHAH et al., 2013) .

SGs são estruturas formadas como resposta a vários fatores estressantes, tais como o estresse oxidativo, privação de nutrientes, estresse osmótico, entre outros (WANG et al., 2012). Os PBs atuam na formação dos SGs, que comumente se formam próximos aos PBs, e juntos formam um complexo que protege os mRNAs e favorece a retomada da tradução após o estresse (BUCHAN; MUHLRAD; PARKER, 2008). Os mRNAs que saem da tradução durante o estresse, primeiramente entram nos PBs e depois transitam para os SGs, onde são preparados para a reentrada na tradução (BUCHAN; MUHLRAD; PARKER, 2008). Sabendo da importância da regulação dos PBs e SGs para a sobrevivência da célula durante períodos de estresse, o lncRNA transcr_10027 da linhagem BY4742 foi selecionado para estudo, uma vez que existem evidências deste lncRNA estar interagindo com proteínas de vias de estocagem e degradação de moléculas (Figura 1).

Estudos anteriores conduzidos pelo grupo de pesquisa, mostraram que o lncRNA transcr_9136 da linhagem SEY6210, além de atuar nas vias de estocagem de proteínas, é um importante regulador do ciclo celular, atuando na modulação de Clb1/2p, um complexo necessário para a entrada da levedura na fase M (LÁZARI et al., 2021; TZENG et al., 2011). Desta forma, selecionamos este lncRNA para estudos mais aprofundados.

**Long non-coding RNAs affects mRNA translation and processing in ethanol stress
in *Saccharomyces cerevisiae***

Amanda Piveta Schnepper¹, Luiz Henrique Cardoso¹, Camila Moreira Pinto¹, Matheus Naia Fioretto², Lucas Cardoso Lazari^{1,3}, Luís Antonio Justulin², Guilherme Targino Valente^{1,4}

¹Department of Bioprocess and Biotechnology. School of Agriculture. São Paulo State University (UNESP). Botucatu, SP. Brazil.

²Department of Structural and Functional Biology, Institute of Biosciences. São Paulo State University (UNESP). Botucatu, SP. Brazil

³Biomedical Sciences Institute, Department of Parasitology. University of São Paulo (USP). São Paulo, SP, Brazil.

⁴Max Planck Institute for Heart and Lung Research. Bad Nauheim, Hessen. Germany

Introduction

Ethanol tolerance is the potential for cell survival during chronic exposure to this substance (STANLEY et al., 2010a). *Saccharomyces cerevisiae* tolerates high concentrations of ethanol because it evolved the ability to produce this compound from sugar. The last phase of industrial fermentation reaches toxic ethanol levels for yeast (SNOEK; VERSTREPEN; VOORDECKERS, 2016), decreasing cell viability and hampering productivity. Resistant strains must be used for industrial purposes, which have been obtained through selection and targeted mutations (STANLEY et al., 2010a; ZHANG et al., 2012).

At least 75% of *S. cerevisiae* genome is transcribed even including many of non-coding RNAs (ncRNAs) (YAMASHITA; SHICHINO; YAMAMOTO, 2016). NcRNAs are transcripts responsible for functional, regulatory and structural mechanisms acting mainly in the gene expression regulation (STATELLO et al., 2021).

LncRNAs in *S. cerevisiae* are involved in metabolic changes, initiation of sexual differentiation and other processes still poorly understood (YAMASHITA; SHICHINO; YAMAMOTO, 2016). Furthermore, studies in *Saccharomyces* evidenced that related lncRNAs interact with proteins and RNAs, acting as a mediator and binding support of specific proteins in RNA, which may be a stress response mechanism (GALIPON et al., 2013).

In animals, both ribosomal proteins and RNAs are processed in lysosomes during autophagy-mediated degradation providing a valuable source of amino acids and

nucleotides under conditions of nutrient deprivation (HUANG et al., 2015; LAFONTAINE, 2010). However, it remains unclear whether mature ribosomes are degraded by autophagy under normal growth conditions and whether ribosome degradation contributes to animal physiology.

LncRNAs may interact with several proteins to regulate the metabolism in *S. cerevisiae*. Many lncRNAs interact with degradation pathways related proteins such as P-bodies (PBs), proteasomes and protein polyubiquitination and up-regulation of ubiquitination (PPPR). Moreover, lncRNAs may also interact with molecule storage pathways related proteins, such as stress granules (SGs), proteasome stress granules (PSGs) and protein desubiquitination and down-regulation of ubiquitination (PDNR) (WOLF et al., 2021).

Eukaryotic ribonucleoprotein complexes (RNP) regulate the translation, degradation and localization of mRNAs (SHAH et al., 2013), such as PBs and the SGs. PBs are conserved stress response structures (BUCHAN; MUHLRAD; PARKER, 2008). In human, PBs present a cluster of RNA binding proteins (AGO1, AGO2, ELAVL1, FRMR1, HNRNPM, HNRNPU, IFGF2BP1, IGF2BP2, IGF2BP3, MOV10, PUM2, and ZC3H7B) which favors the addressing of mRNAs inside the PBs during stress (HUBSTENBERGER et al., 2017). There is a direct correlation between the number of PBs and cell lifetime during the stationary phase (SHAH et al., 2013).

Several stressors such as oxidative stress, nutrient deprivation, osmotic stress, heat shock, among others trigger SGs formations in yeast (BUCHAN; YOON; PARKER, 2011; WANG et al., 2018). PBs lie close to SGs and act in its formation. The SGs and PBs complex protect mRNAs and favors translation resumption after stress (BUCHAN; MUHLRAD; PARKER, 2008). The translation stalling during stress relies on mRNAs entering in PBs and then transiting to SGs, where mRNAs are prepared for re-entry into translation after stress (BUCHAN; MUHLRAD; PARKER, 2008).

The lncRNA lncRNA Inc9136 (hereafter referred to as transcript_9136) of the SEY6210 strain acts on cell cycle arrest release during ethanol stress in yeast by interacting with proteins of this pathway (LÁZARI et al., 2021). However, lncRNA-protein interactions suggested that lncRNA may work in the protein storage pathways (WOLF et al., 2021). Despite the extensive knowledge about stressors in *Saccharomyces*, the regulatory mechanisms impacted by lncRNAs is still unknown. Here, we address these issues through the inactivation of the lncRNAs transcr_10027 and transcr_9136, as well as changes in the stress response by ethanol. The analysis of PBs and SGs allowed us to explore the relationship between transcr_10027 and the formation of PBs and SGs in the BY4742

lineage, as well as the translational changes caused by the deletion of transcr_9136 in the SEY6210 lineage.

MATERIAL AND METHODS

Prediction of lncRNAs-Protein Interaction Network

The transcr_10027-protein and transcr_9136-protein interactions were predicted with the lncPRO program (LU et al., 2013) for all proteins greater than 32 amino acids from each strain of *Saccharomyces cerevisiae* (MARQUES et al., 2021). Interactions with probability ≥ 0.95 were selected to compose the network of lncRNA-protein interactions (MARQUES et al., 2021).

Generating the Saccharomyces cerevisiae mutants

Forward oligos were constructed for lncRNAs transcr_9136 (gi|696450445|gb|JRIW01000133.1|:86676-84821) of SEY6210 and transcr_10027 (gi|696446837|gb|JRIR01000161.1|:54479-51747) for BY4742 specifying the insertion of the target in the gRNA expression plasmid pMEL16 (Addgene assess number 107922). The oligos to target the genome loci of those RNAs were constructed using the CRISPRdirect program (available at <https://crispr.dbcls.jp/>) (Naito et al 2015). The reverse primer “OliB” was used in all reactions (**Table 1**). PCR was performed as follow: 30 ng of pMEL16, 1 μ l of forward primer (10 μ M), 1 μ l of OliB primer (10 μ M), 1 μ l of dNTP (10 mM) (Invitrogen), 5 μ l HF buffer (5x), 1 μ l Phusion™ High-Fidelity DNA Polymerase (ThermoFisher) for a final volume of 25 μ l. Cycling was performed as follow: 98 ° C for 2 min, followed by 10 cycles at 98° C for 30 s, X° C for 30 s and 72° C for 6 min, followed by 10 cycles at 98° C for 30 s, Y° C for 30 s and 72° C for 6 min, followed by 20 cycles at 98° C for 30 s, Z° C for 30 s and 72° C for 6 min, with final extension of 72° C for 10 min. The X, Y and Z are specific for each set of primers they are available in Table 1.

Table 1: Oligos used in the CRISPR-Cas9 system and to analyze the differential expression of lncRNAs neighboring genes.

Strain	Description	Sequence (5' - 3')	Temp. X	Temp. Y	Temp. Z
SEY6210	gRNA transcr_9136	P-GTTTTAGAGCTAGAAATAGCTCATGTACTCCATAGAGTGA	65,4° C	61,4° C	58,4° C
BY4742	gRNA transcr_10027	P-ACAAATCCGTGATAGTTTAAGTTTTAGAGCTAGAAATAGC	65,4° C	61,5° C	58,4° C
-	OliB	P-GATCATTTATCTTTCCTGC	-	-	-
SEY6210	Donor DNA transcr_9136 Fw	TCCACACACCCATTTTCGCTAGCGTAGAACAAGGGGAGACACA AACTTTCTTTTCCTTGCATAATTATTTCCCTCGTTGCTACTCATT GAGGCCGCTCCATATGGAGATTTGAAAAAGGTT	-	-	-
SEY6210	Donor DNA transcr_9136 Rv	AACCTTTTTCAAATCTCCATATGGAGCGGCCTCAATGAGTAGC AACGAGGGAAATAATTATGCAAGGAAAAGAAAGTTTGTGTCTC CCCTTGTTCTACGCTAGCGAAATGGGTGTGTGGA	-	-	-
BY4742	Donor DNA transcr_10027 Fw	GAATTGAACCCCGATCTGGCAGCGACAAGCGCCCATTCTGA CCATTAACTATCACGGAGGGTTCTGCTCAGATTTTCGTAAATA CCCGGACCATCGTAGATGATTGTTTTTACTCTCC	-	-	-
BY4742	Donor DNA transcr_10027 Rv	GGAGAGTAAAAACAATCATCTACGATGGTCCGGGTATTTAACG AAATCTGAGCAGAACCCTCCGTGATAGTTTAATGGTCAGAATG GGCGCTTGTCGCGTGCCAGATCGGGGTTCAATTC	-	-	-
SEY6210	ILT1 Fw	TTATTGCGGCTGATGTTGGC	60°C	-	-
SEY6210	ILT1 Rv	GCCAAGCACCTAATGAATCG	60°C	-	-
SEY6210	RRP1 Fw	CGTCCTAGACCTCAGCAACG	60°C	-	-

SEY6210	RRP1 Rv	GGTCAACTCATCTGCAGTGCTA	60°C	-	-
BY4742	FAA3 Fw	AAACGGCGGTCTCTTTCACT	60°C	-	-
BY4742	FAA3 Rv	GTTGTCCCTTGGGAGTCCAT	60°C	-	-
BY4742	TIR3 Fw	CTCCTCCTCTGCTACCTCCA	60°C	-	-
BY4742	TIR3 Rv	ACCAACACCAGCGGAGAAG	60°C	-	-

PCR products were treated with DpnI restriction enzyme (New England BioLabs), following the manufacturer's recommendations. The reactions were purified using Wizard® SV Gel and PCR Clean-Up System (Promega) and quantified in NanoDrop. Then, 200 ng of products were treated with the enzyme T4 DNA Ligase (Invitrogen), following the manufacturer's recommendations, and stored at -20° C.

A total of 20 ng of plasmid was added to 100 µL of competent bacteria TOP10 and incubated on ice for 30 min. Heat shock was performed at 42 ° C for 90s and incubating on ice for 2 min. Cell wall recovery was performed with incubation in LB liquid medium (KASVI) for 2 h at 37° C followed by inoculation (at 37° C for 16 h) in LB agar medium (KASVI) with the appropriate antibiotic and incubated. Colonies were inoculated in liquid medium and incubated at 37° C for 16 h at 200RPM. Then, plasmid extraction was performed using QuickLyse® Miniprep Kit (Qiagen) following the manufacturer's recommendations.

PCR was performed using the M13 primers to check the target insertion. The PCR reaction followed the following conditions: 1 µl of buffer, 0.2 µl of dNTP (10 mM) (Invitrogen), 0.3 µl of MgCl₂, 0.5 µl of Fw primer (5'-GTAAAACGACGGCCAGT-3') (10 µM), 0.5 µL Rv primer (5'-CAGGAAACAGCTATGAC-3') (10 µM), 0.05 µL Platinum® Taq DNA Polymerase (ThermoFisher) adjusting final volume to 10 µL using water. The cycling used was 94° C for 2 min, followed by 40 cycles of 94° C for 15 s, 50° C for 30 s and 68° C for 1 min, with final extension at 68° C for 5 min, in a GeneAmp thermocycler ® PCR System 9700 (Applied Biosystems). The PCR product was analyzed on a 2.0% agarose gel, sequenced on an ABI3730XL platform and analyzed in Geneious® 7.1 software.

The repair DNA was complementary to the lncRNAs' adjacent regions: 60bp from each side were selected generating a 120bp repair oligo (Table 1). A forward and reverse oligos were synthesized, equimolar amounts of each repair were mixed, heated at 94° C for 10 min, incubated at room temperature for 3 h, and used for transformation in yeast cells.

The lithium acetate was used for chemocompency of selected strains. Cells were stored in 50% glycerol at -80° C until use. 10 µL of Salmon Sperm Deoxyribonucleic acid – SSDNA (Sigma-Aldrich), 1 µg plasmid p414 (Addgene 43802), 1 µg modified plasmid pMEL16, 20 µl double-strand repair DNA, and 100 µl of chemocompetent cells were used in the transformation step. 600 µL of a PLATE solution (Ryan, 2016) was added in the mixture with subsequent incubation at 30° C for 30 min at 200RPM. Then, 10% (v/v) of DMSO was added, followed by an incubation at 42° C for 15 min. Cells were placed on ice for 2 min and incubated at 30° C for 2 h using Yeast Synthetic Drop-out Medium Supplements HIS- (Sigma-Aldrich). After two hours, the sample was plated in Yeast

Synthetic Drop-out Medium Supplements without histidine (Sigma-Aldrich) with Bacto-Agar (Dynamic) and 200 mg/L of G418 (Sigma-Aldrich) and incubated at 30° C for 3-4 days.

The negative control was performed by submitting the competent cells to the same transformation protocol, but without the insertion of plasmids. The positive control followed the same methodology, but with the insertion of plasmids and not the insertion of repair DNA. In addition, viability testing was performed by plating an aliquot of cells onto plates containing YPD-agar. All plates were incubated under the same conditions.

Gene expression in mutant strains SEY6210 transcr_9136 Δ and BY4742 transcr_10027 Δ

To assess the possible variation in gene expression of lncRNAs' neighbor genes, mutants (SEY6210 transcr_9136 Δ and BY4742 transcr_10027 Δ) and corresponding wild-type (WT) strains were grown overnight at 30° C and shaking at 200RPM. OD₆₀₀ was determined in NanoDrop and corrected to 0.2. Yeasts allocated in 6-well plates were treated with ethanol at the maximum tolerance of each strain (20% for SEY6210 and 26% for BY4742) for 1 h, at 30° C and 200RPM, with a final volume of 2 mL. The experiment was conducted in triplicates. After treatment, cells were transferred to 1.5 mL microtubes, centrifuged for 2 min at 2500 x g, the supernatant was discarded, and the pellets stored at -80° C.

Total RNA extraction was performed using SV Total RNA Isolation System (Promega), according to the manufacturer's specifications, after a treatment with 100 μ L of lyticase for 30 min (Sigma-Aldrich) and 0.1 μ L of β -mercaptoethanol at room temperature. RNA quality was assessed on a 1.5% agarose gel stained with GelRed (Sigma) and bromophenol blue and quantified using NanoDrop.

Then, 1 μ g of RNA was used for cDNA synthesis using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystem) following the manufacturer's specifications. The qPCR was performed using the following reaction: 3 μ L cDNA (1:10 dilution), 1.5 μ L forward primer (10 μ M), 1.5 μ L reverse primer (10 μ M), 7.5 μ L GoTaq $\text{\textcircled{R}}$ qPCR MasterMix (Promega) for a final volume of 15 μ L. The qPCR cycling was 95° C for 2 min, 45 cycles of 95° C for 3 s, and 60° C for 30 s, with a melting curve, using a 7500 Real-Time PCR Systems thermocycler (Applied Biosystems). The TDH2 gene (YJR009C) was used as endogenous (VAUDANO et al., 2011). CT data were analyzed using $\Delta\Delta$ Ct methodology and statistical

analysis using GraphPad Prism 7.00 software. The primers used to assess gene expression are described in Table 1.

Protein extraction and western blot analysis for strains BY4742 WT, BY4742 transcr_10027 Δ , SEY6210 WT and SEY6210 transcr_9136 Δ

To assess possible changes caused in the formation of PBs and SGs during ethanol stress, total protein extraction was performed, followed by western blot analysis. For this analysis, yeast strains were grown overnight in YPD medium, at 30° C and 200RPM shaking. OD₆₀₀ was quantified in NanoDrop and diluted to 0.3 in 20 mL of YPD or YPD + ethanol medium. The experiment was carried out with 6 biological replicas.

The cells were centrifuged at 1,000RPM for 1 min, the supernatant was discarded and 2.0 M LiAc and 0.4 M NaOH was added. The contents were again centrifuged, the supernatant was discarded and the cells were stored at -80° C. Protein concentration was measured using the Bradford method (Bradford, 1976). Aliquots (30 to 70 μ g of proteins) were subjected to 10% polyacrylamide gel electrophoresis. After electrophoresis, the proteins were transferred to a Hybond ECL nitrocellulose membrane (Amersham, Little Chalfont, UK). For 1 h, membranes were blocked with 5% skim milk diluted in TBS, followed by an incubation (16 h at 4° C) with primary antibody that binds to the desired target: anti-PABP (sc-166027, Santa Cruz, 1:1000), anti-Elf4e (sc-9976, Santa Cruz, 1:1000) and anti-DCP1a (sc-100706; Santa Cruz, 1:1000) (NILSSON; SUNNERHAGEN, 2011; PARK et al., 2018; RZECZKOWSKI et al., 2011).

After washing in TBS for 2 h, the membranes were incubated with secondary antibody specific for each primary antibody, followed by further washing using TBS. The relative molecular weight of the bands was determined according to the molecular weight standard, Kaleidoscope (Bio-Rad, Hercules, CA, USA), which was added during electrophoresis. The chemiluminescence reaction was developed in an ECL kit and the images were captured using Image Quant 350 (GE - Healthcare). Semiquantitative analyzes by band densitometry were performed and analyzed using the GraphPad Prism 7.00 software.

Analysis of ethanol tolerance of mutant yeasts

Yeast strains were grown overnight in 20 ml of YPD, 30 °C and shaking at 200RPM. After growth, the strains were diluted to an OD₆₀₀ of 0.4 in 2 mL of YPD with different concentrations of ethanol ranging from 6% to 40%. After 1h of treatment at 30°C and 120RPM, the serial dilution of the cultures in double-distilled water was carried out 1:10. Then, 5 µL of each dilution was inoculated into a YPD plate and incubated at 30 °C for 48 h.

RESULTS

Analysis of the impact of lncRNAs deletion on metabolism

Genes flanking the transcr_9136 of SEY6210 strain exhibited an alteration in expression dependent on presence of this lncRNA (Figure 2 – B). In the SEY6210 transcr_9136 Δ under ethanol stress, it is observed that the RRP1 (YDR087C) is down-regulated. Interestingly the ILT1 (YDR090C) gene did not show significant difference in expression under control and treatment condition for that mutant, showing an alteration in relation to the SEY6210 strain, ILT1 presents an up-regulation in the treatment compared to the control (Figure 2 – D). Conversely the transcr_10027 neighbor genes in the BY4742 strain did not present alterations in the gene expression in the absence of this lncRNA (Figure 2 – A and C).

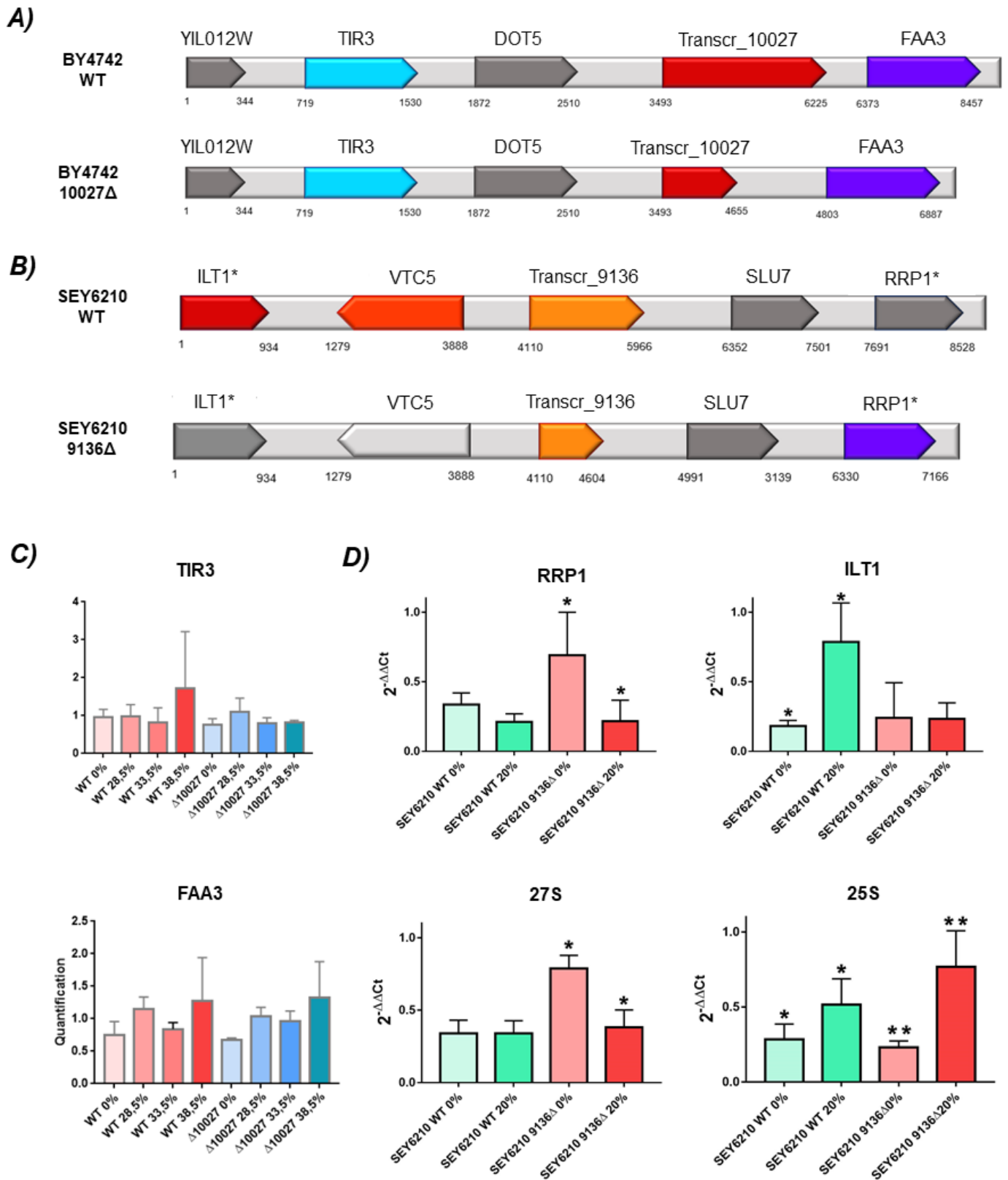


Figura 2. Evaluation of the expression of genes neighboring the lncRNAs *transcr_10027Δ* and *transcr_9136Δ*. A) Expression variation of genes neighboring *transcr_10027Δ* in WT and mutant strains. B) Expression variation of genes neighboring *transcr_9136Δ* in WT and mutant strains. The *VTC5* (YDR089W) and *SLU7* (YDR088C) genes were not evaluated by RT-qPCR, and the

differential expression data came from the foldchange of RNAseq data (WOLF et al., 2021). C) expression data of genes neighboring transcript_10027 of the BY4742 lineage. D) Expression data of genes neighboring transcr_9136 of the SEY6210 lineage, as well as 27S and 25S rRNAs. Red indicates up-regulated, blue indicates down-regulated and gray has no differential expression/not evaluated

The RT-qPCR data showed that 27S pre-rRNA is constant in the WT regardless of stress whereas the 25S was increased under stress. Conversely to the 27S, the 25S was reduced and increased in the SEY6210 transcr_9136 Δ without and under stress, respectively (Figure 2 - D).

The spot test was performed to assess the impact of the transcr_9136 deletion on ethanol response (Figure 3). The ethanol tolerance test showed that the mutants reduced the ability to grow after stress compared to their WT versin (Figure 3 – A and B).

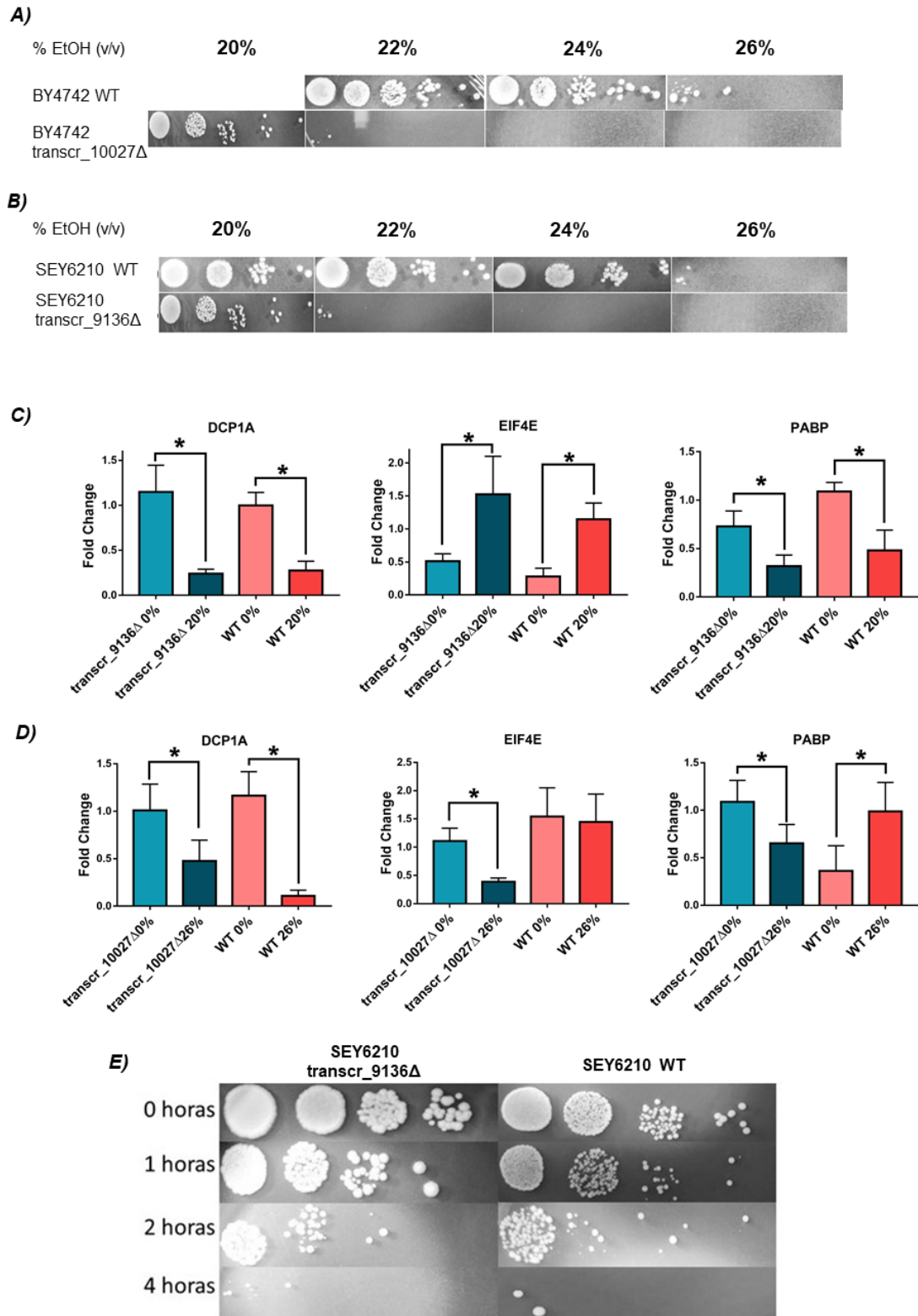


Figura 3: Ethanol tolerance tests and protein quantification A) spot test of strain BY4742 WT and transcr_10027Δ. B) Spot test of the SEY6210 WT and transcr_9136Δ strains. C) Western blot for DCP1A, EIF4E and PABP protein for SEY6210 WT and transcr_9136Δ strains in 0% and 20% ethanol. D) Western blot for DCP1A, EIF4E and PABP protein for BY4742 WT and transcr_10027Δ strains in 0 and 26% ethanol. E) Spot test for SEY6210 WT and transcr_9136Δ strains with 20% ethanol treatment from 0 to 4 hours of exposure

The spot test data showed that the WT tolerates a higher concentration of ethanol when compared to the mutant strain (Figure 3 – B). The spot-test carried out after chronic ethanol exposure for 1 – 4h showed that the SEY6210 *transcr_9136Δ* strain had a higher surveillance during stress compared to the WT (Figure 3 – E).

In order to better understand the role of *transcr_10027* in ethanol stress western blot experiments were made with strain BY4742 *transcr_10027Δ*. Unlike the observed in BY4742 WT, there are no significant differences in EIF4E and PABP (Figure 3 – D). Furthermore, both WT and mutant strains have the same profile in control and treatment for the DCPA1.

The western blot results of SEY6210 WT and *transcr_9136Δ* mutant showed no differences in the patterns between them for all proteins analyzed, although the differences within strains are significant (Figure 3 – C).

DISCUSSION

Transcr_10027 as one of the stress response regulators through the formation of PBs and SGs

It is suggested that the lncRNAs *transcr_10027* of the BY4742 strain may act mainly in signaling, cell division and degradation by interacting with the Pik1p (YNL267W) (WOLF et al., 2021), a kinase acting on cells in nutrient-poor media (DEMMELE et al., 2008). *Transcr_10027* also interacts with Tel1p (YBL088C) (WOLF et al., 2021), an inhibitor of PBs formation (TKACH et al., 2012). The data obtained through western blot demonstrate that one of the markers for PBs, EIF4E, shows a lower yield of this protein in the mutant strain, than the WT strain. The marker for SG (PABP) also shows a variation, demonstrating a reduction in the mutant strain, when relating control and treatment, and an increase in protein concentration in the WT strain. Therefore, we hypothesize that the *transcr_10027* acts as an activator of Tel1p, in which this protein would not be active in the lack of this lncRNA (Figure 4). Hence, the BY4742 *transcr_10027Δ* mutant would present a faster recovery after ethanol stress relief, as observed (WOLF et al., 2021), by producing a greater number of PBs and SGs.

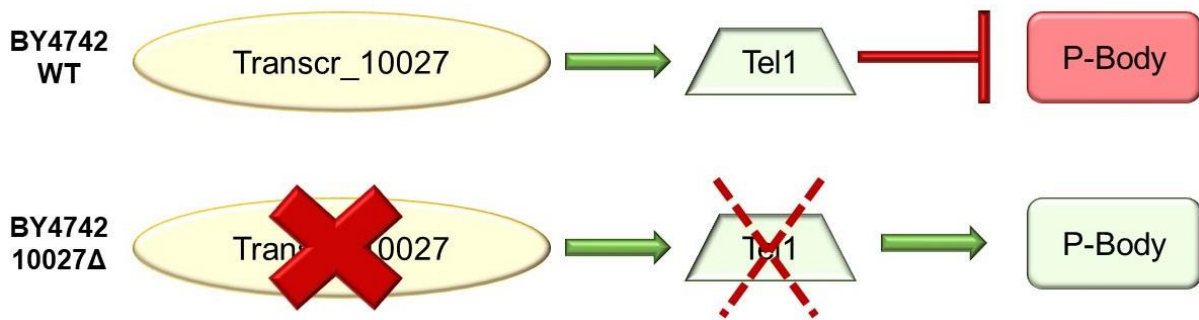


Figura 4: regulation of the formation of PBs mediated by transcr_10027 in the BY4742 lineage.

The eIF4E protein acts favoring the binding between Pab1p and mRNAs for translation initiation (DEVER; KINZY; PAVITT, 2016). The PBs and SGs complex regulates the translation during stress, through the storage and processing of mRNAs, preventing or facilitating entry into translation (JAIN; PARKER, 2013); this mechanism can contribute to cell tolerance to the stressful environment (IVANOV; KEDERSHA; ANDERSON, 2019). The significant reduction in eIF4E and Pab1p in transcr_10027Δ mutant during ethanol stress compared to WT indicates that PBs and SGs may be not working properly during stress in the absence of transcr_10027. Although mRNAs may be protected from degradation within the complexes they may be not underwent to the translation (HUBSTENBERGER et al., 2017). This may explain why, despite the mutant transcr_10027Δ having a faster recovery after stress and better tolerate environments with low amount of nutrients (WOLF et al., 2021), this strain tolerates lower amounts of ethanol when compared to the wild, as observed here.

The transcr_9136 lncRNA as a potential regulator of the translational apparatus

It has been suggested that the lncRNA transcr_9136 of SEY6210 strain is directly related to cell cycle regulation, in which the transcr_9136 indirectly activates the Swe1po (YJL187C) inhibiting the Clb1/2 (YGR108W/YPR119W) during ethanol stress, releasing the cell cycle arrest (LÁZARI et al., 2021). Moreover, this lncRNA also has interactions with proteins Sla1p (YBL007C), Bni1p (YNL271C) and Hbt1p (YDL223C) (WOLF et al., 2021), which are proteins related to cytoskeleton structure associated with endocytosis and cell polarization, respectively (PRUYNE et al., 2004; TCHEPEREGINE; GAO; BI, 2005).

Some genes flanking the *transcr_9136* were differentially expressed comparing the SEY6210 wild-type and SEY6210 *transcr_9136Δ* strains. For instance, *ILT1* (YDR090C) was reduced in the mutant under stress. This gene encodes a plasma membrane protein with unknown function. However, it seems to be related to response to ionic toxicity, including growth reduction in *ILT1Δ* strain (HIGGINS et al., 2018), which may be contributing to a lower ethanol tolerance of the mutant strain when compared to wild type. The *RRP1* gene (YDR087C) is also lower expressed in the mutant lacking the *transcr_9136*. By checking the S288C Hi-C data (SCHALBETTER et al., 2019), we observed that the homologous loci in which *transcr_9136* lie on interacts with the *RRP1* locus during prophase followed by relaxing during mitotic division. Thus, we suggest that the region surrounding this lncRNA may had undergone chromatin compaction changings in the *transcr_9136Δ* mutant, which may be related to the variation observed in differential expression.

We hypothesize that the accumulation of pre-rRNA during ethanol stress is related to rebound capacity after chronic exposure by a reduction in the *RRP1* gene expression, in which the *transcr_9136* plays a negative role in this process. To test this hypothesis, we assessed the processing of 27S pre-rRNA into 25S, and the *RRP1* expression: *RRP1* is essential for the maturation of pre-rRNA into rRNA: *RRP1* mutant presented an accumulation of 66S ribosomal (HORSEY et al., 2004).

The expression of *RRP1* is constant in the WT regardless of stress, whereas this gene is increased in the SEY6210 *transcr_9136Δ* and reduced under stress, suggesting that the *transcr_9136* seems to act as a modulator of *RRP1* stable expression. The SEY6210 *transcr_9136Δ* has a lower population rebound than the WT from 18-22%, higher rebound at 24% and similar rebound at 26% (LÁZARI et al., 2021), a lower ethanol tolerance, and a better growth after chronic long-time exposure to ethanol. Thus, although the WT strain tolerates a higher concentration of ethanol, it has a harmed surveillance under chronic stress. We observed that the lack of *transcr_9136Δ* allowed an 27S pre-rRNA accumulation in strains out of ethanol stress, which is not observed in the WT. Thus, we suggest that the SEY6210 *transcr_9136Δ* under stress recovers the ribosomal synthesis promptly converting the bulky of 27S pre-rRNA, which may be not occurring in the WT (Figure 5B). This mechanism may be related to the observed better surveillance of SEY6210 *transcr_9136Δ* mutant under chronic ethanol exposure, boosted by the observed higher Rrp1p activity under stress.

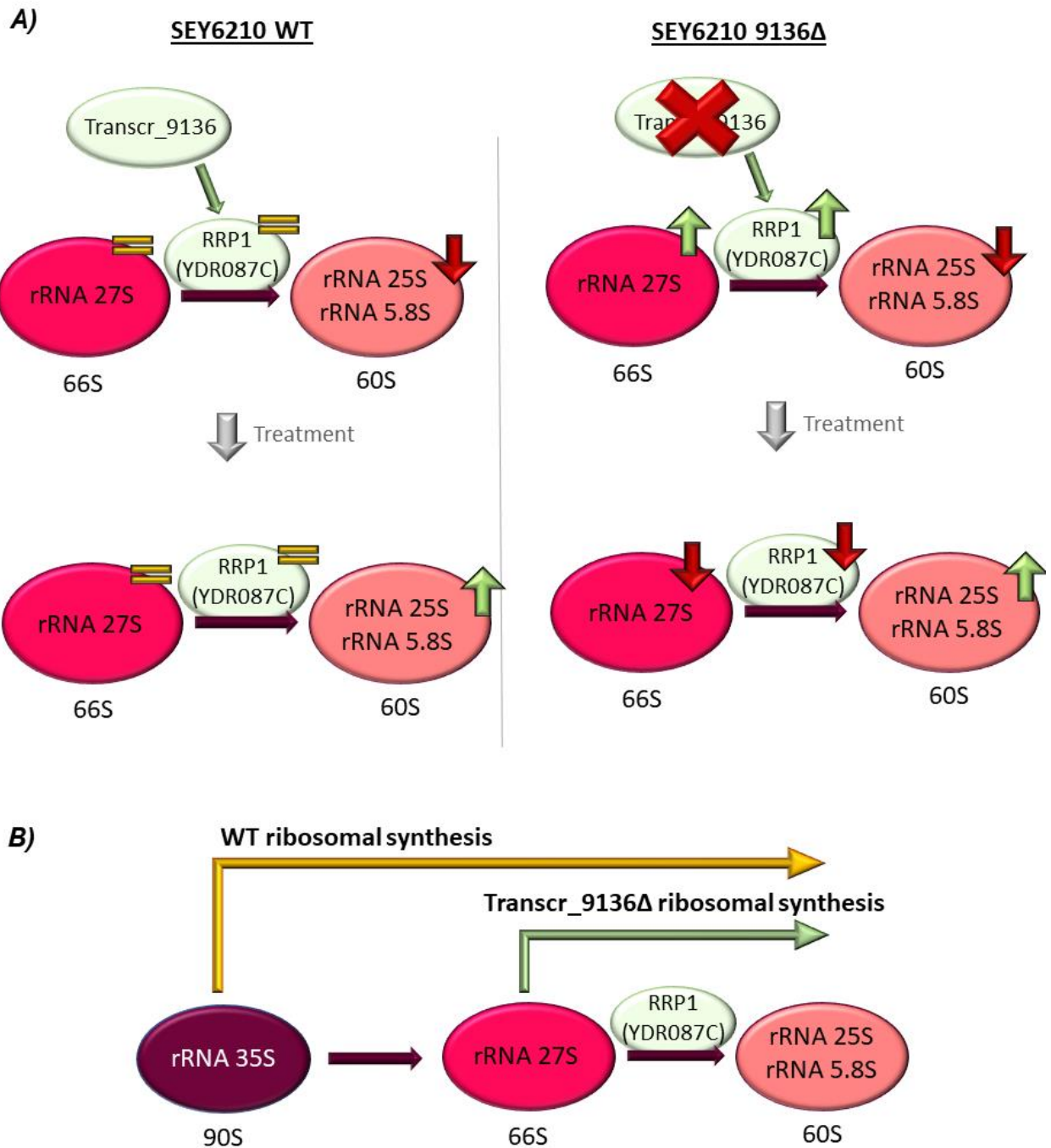


Figura 5: A) Overview of expression data of ribosomal related genes in SEY6210 WT and *transcr_9136Δ* strains. Green arrows indicate up-regulate, red arrows indicate down-regulate and sign equal indicate no significant difference in expression between control and treatment. B) Proposed mechanism of translational initiation of ribosomal syntheses during ethanol stress.

CAPÍTULO 2

INTRODUÇÃO

LncRNAs como reguladores transcricionais, pós-transcricionais e traducionais

Os lncRNAs são moléculas de RNA com mais de 200nt e que geralmente não codificam proteínas funcionais e tem capacidade de atuar em diferentes etapas do ciclo celular, desde a regulação transcricional até a traducional (JARROUX; MORILLON; PINSKAYA, 2017; STATELLO et al., 2021).

Como reguladores transcricionais, os lncRNAs podem atuar diretamente na região promotora dos genes, agindo como silenciadores impedindo a ligação de fatores de transcrição, ou como *enhancers* de lncRNAs (eRNA) acentuando a expressão, tal como diversos lncRNAs diferencialmente expressos em células cancerosas. O lncRNA *HOTTIP* utiliza este mecanismo para regulação dos genes HOX, o qual atua como uma âncora para a ligação de um complexo de fatores de transcrição necessários para a ligação da RNA Pol II (FANG; FULLWOOD, 2016).

As estruturas em tripla hélice, RNA:DNA:DNA são descritas em diversos organismos, possuindo um importante papel regulatório. Por exemplo, uma tripla hélice formada na região telomérica, conservada em leveduras e humanos, é essencial para a atividade catalítica da telomerase (QIAO; CECH, 2008). Em humanos, o lncRNA MALAT1 é conhecido por ser abundante em células cancerosas. Este lncRNA forma estruturas em tripla no genoma afetando processos como degradação de RNAs, transcrição e ação da telomerase (ABULWERDI et al., 2019; BROWN et al., 2014).

As moléculas de RNA também são capazes de formar estruturas chamadas de duplex de RNA, a qual duas fitas de RNA interagem entre si. Este tipo de estrutura pode atuar na regulação pós-transcricional e pode ter um impacto direto na tradução de mRNAs, como o RNA DrsA, o qual atua como um regulador pós-traducional em *Escherichia coli* através da formação de estruturas de RNA:RNA (LEASE; CUSICK; BELFORT, 1998). Em ratos, o lncRNA Sirt1 atua na regulação do ciclo celular de células musculares (mioblastos)

através da formação de duplex de RNA com o microRNA miR-34a (WANG et al., 2016). Em humanos, foram identificadas mais de 5 milhões de interações lncRNA:mRNA, sendo que a formação de duplex de RNA pode ter diferentes efeitos, como regular o *splicing*, competir com miRNAs por sítios de ligação em transcritos, promover eventos de edição de mRNA e hibridizar com 3'UTRs de transcritos contendo Alu e direcionar tais moléculas de mRNA para degradação. (SZCZEŚNIAK; MAKAŁOWSKA, 2016)

As interações lncRNA-proteínas são reportadas como importantes reguladores traducionais. Um dos primeiros lncRNAs conhecidos, HOTAIR, atua interagindo com proteínas relacionadas a metilação de histonas em células cancerosas (BHAN; MANDAL, 2015). Estudos demonstram que várias proteínas podem ligar-se ao mesmo lncRNA de forma colaborativa ou competitiva (FERRÈ; COLANTONI; HELMER-CITTERICH, 2016).

Devido a esta grande flexibilidade de atuação dos lncRNAs, os mesmos são bons candidatos na busca pelo entendimento das regulações metabólicas de *S. cerevisiae*, principalmente por estas vias regulatórias ainda serem pouco exploradas na literatura.

Bioinformática e o estudo de regulação transcricional pela formação de triplas hélices (RNA:DNA:DNA)

Os lncRNAs são capazes de atuarem em diferentes formas de regulação gênica. Esta regulação pode ser mediada por interações RNA-proteína, RNA:RNA ou RNA:DNA; a última permite a formação de uma estrutura em tripla-hélice RNA:DNA:DNA, como já mencionado

O domínio da fita simples de RNA capaz de formar uma tripla hélice recebe o nome de TFO (do inglês *triplex-forming oligonucleotide*), e o domínio de polipurina/polipirimidina (*polypurine/polypyrimidine*) de uma dupla fita de DNA capaz de acomodar um TFO é chamado de TFD (do inglês *triplex target site*) (ZHOU; GILES; FELSENFELD, 2019).

As ligações entre nucleotídeos que são capazes de formar um domínio tripla hélice foi descrito pela primeira vez por Morgan e Wells em 1968. Porém este tipo de regulação é

ainda pouco estudado, implicando em uma escassez de dados e softwares de predição destas estruturas.

A estabilidade da estrutura em tripla hélice se dá pelas ligações de hidrogênio de Hoogsteen, que consistem em um grupo doador e um aceptor de hidrogênio (ESCUDEÉ et al., 1993). Desta forma, três principais domínios são descritos: 1- domínio de pirimidinas, com interações entre timinas e citosinas; 2- domínio de purinas, com interações entre guaninas e adeninas; 3- domínio de purinas e pirimidinas, com interações entre guaninas e timinas (SUN; GARESTIER; HÉLÈNE, 1996).

A configuração de Hoogsteen é formada por T:AT e C⁺:GC no domínio de pirimidinas, sendo que a citosina protonada só é encontrada em pH ácido, não sendo viável em pH biológico. No domínio de purinas, temos a formação de estruturas caracterizadas por G:GC e A:AT, e no domínio de purinas e pirimidinas a configuração de Hoogsteen se dá por G:GC e T:AT (sendo T substituído por uracila nos RNAs).

Neste contexto, os lncRNAs são candidatos a formarem estruturas de tripla hélice em regiões de regulação transcricional, podendo atuar na expressão, silenciamento ou ativação de genes importantes para o metabolismo de *Saccharomyces cerevisiae*.

MATERIAIS E MÉTODOS

Obtenção dos lncRNAs do transcriptoma de *Saccharomyces cerevisiae*

A montagem dos *long non-coding RNAs* (lncRNAs) foi obtida de Wolf *et al.* e Marques *et al.* (MARQUES et al., 2021; WOLF et al., 2021). Brevemente, as bibliotecas de RNA-Seq foram avaliadas utilizando os dados obtidos pelo programa FASTQC (["https://www.bioinformatics.babraham.ac.uk/projects/fastqc/"](https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)), leituras biblioteca de sequências (*reads*) de baixa qualidade e com possíveis resquícios de adaptadores foram removidos utilizando o programa Trimmomatic v.0.36 (BOLGER; LOHSE; USADEL, 2014). Então, os *reads* foram mapeados em um banco de dados de genes codificantes de proteínas de múltiplas espécies com o programa HISAT2 (KIM; LANGMEAD; SALZBERG, 2015) e possíveis sequências codificantes foram removidas. Os *reads* restantes foram

montados utilizando múltiplos algoritmos e um *score* foi calculado para determinar a melhor montagem (MARQUES et al., 2021). Redundâncias entre os transcritos montados foram removidas com o programa CD-HIT (FU et al., 2012), montagens espúrias foram removidas por meio do mapeando de transcritos nos genomas de referência das linhagens de *Saccharomyces cerevisiae* com o programa GMAP (WU; WATANABE, 2005). Finalmente uma última checagem sobre o potencial codificante dos transcritos montados foi efetuada, transcritos sem potencial codificantes foram selecionados para o compor o conjunto final de lncRNAs (MARQUES et al., 2021).

Análise da formação de triplas hélices das linhagens de Saccharomyces cerevisiae

Foi selecionada uma região de até 1500pb *upstream* aos genes mapeados das linhagens analisadas de *S. cerevisiae*, descartando as regiões codificantes, para obtenção das prováveis regiões promotoras dos genes. Posteriormente, estas sequências foram analisadas pelo software Triplexator (BUSKE et al., 2012), juntamente com as sequências dos transcritos dos lncRNAs, com o intuito de determinar possíveis sítios de formação de triplas hélices. Esse processo foi realizado para as linhagens BMA64-1A, BY4742, SEY6210, BY4741, X2180-1A e S288C, independentemente.

O resultado do software Triplexator foi filtrado seguindo os parâmetros: %GC \geq 20, domínios permitidos = 'R' e 'M', taxa de erro = 0, orientação = 'P' e 'A', fita = '+' e '-'. Ao final, os pares de gene:lncRNA foram analisados quanto à sua possível correlação da expressão no controle e tratamento com etanol, *time-course* e expressão diferencial.

Os resultados para as seis linhagens foram classificados de acordo com o *fold change* do RNAseq (WOLF et al., 2021) dos lncRNAs e genes alvo. Essa classificação abrange 9 classes.

A linhagem BMA64-1A foi selecionada para um estudo mais aprofundado da estrutura das interações tripla hélice por ser a linhagem de maior tolerância ao etanol (WOLF et al., 2021) e por possuir dados de *time-course*. A região promotora dos genes selecionados da BMA64-1A, após um teste de correlação utilizando o *time-course*, foi analisada no software Conreal (BEREZIKOV; GURYEV; CUPPEN, 2005), para determinar

os sítios de ligação para fatores de transcrição (TFBS), TATAbox, CCAATbox e elementos transponíveis.

RESULTADOS

O resultado da análise de formação de tripla hélice da região promotora dos genes das seis linhagens de *S. cerevisiae* foram separadas em nove classes, considerando a expressão diferencial dos genes e dos lncRNAs (Figura 6).

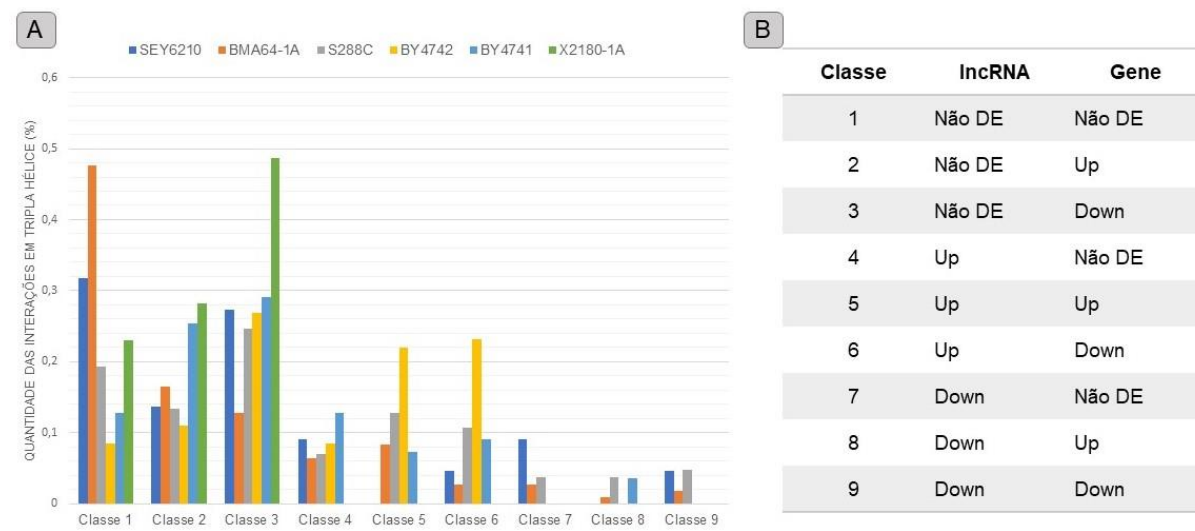


Figura 5: classificação das interações em tripla hélice. **A**: gráfico de porcentagem de cada classe por linhagem; **B**: classes utilizadas para classificação das interações tripla hélice quanto a expressão diferencial

Nota-se a presença de um padrão em todas as linhagens, a qual há uma maior quantidade de interações quando o lncRNA não tem expressão diferencial (classes 1, 2 e 3), uma quantidade moderada de interações quando o lncRNA está *sobre-expresso* (classes 4, 5 e 6) e uma quantidade muito pequena de interações quando o lncRNA está *down-regulado* (classes 7, 8 e 9).

A linhagem BMA64-1A possui um *time-course* de 1h, 2h e 4h para a expressão dos genes e lncRNAs (WOLF et al., 2021). Isso permitiu uma análise mais robusta de algumas interações, tais como as transcr_19008:YLR019W, transcr_3150:YBR154C e transcr_21069: YNL088W (Figura 7).

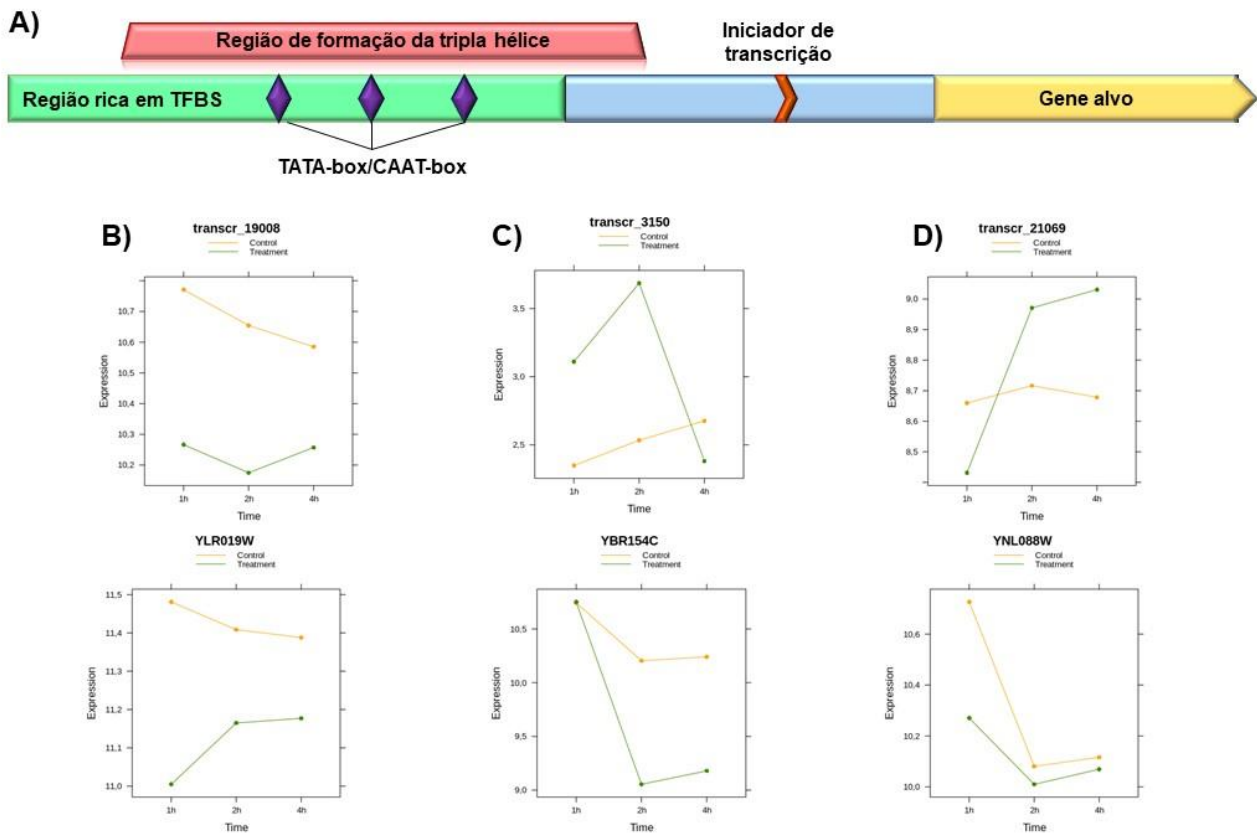


Figura 6: A) Representação da estrutura dos genes analisados, bem como a região da formação das triplas hélices. B) Time-course da expressão diferencial das interações transcr_19008:YLR019W, transcr_3150:YBR154C e transcr_21069: YNL088W.

A análise da região promotora destes genes demonstrou que os sítios de formação da tripla hélice estão localizados muito próximos de regiões ricas em TFBS e TATAbox, como é o caso do transcr_19008:YLR019W, ou diretamente sobre estas regiões (ex.: transcr_3150:YBR154C e o transcr_21069:YNL088W) (Figura 7). Além disso, análise do *time-course* dos genes e lncRNAs da BMA64-1A mostrou uma correlação entre o aumento da expressão lncRNA e a redução da expressão do gene alvo, principalmente durante o tratamento com etanol.

DISCUSSÃO

Baseado nas classes propostas para classificação das interações entre lncRNAs e regiões promotoras do gene, é possível propor alguns mecanismos de ação para explicar estas regulações. Para as classes 1, 2 e 3 não foi possível propor um mecanismo, pois,

como o lncRNA não possui expressão diferencial, não identificamos um comportamento que explicasse essas regulações. Para as classes 4 a 9, os mecanismos propostos estão ilustrados na Figura 8

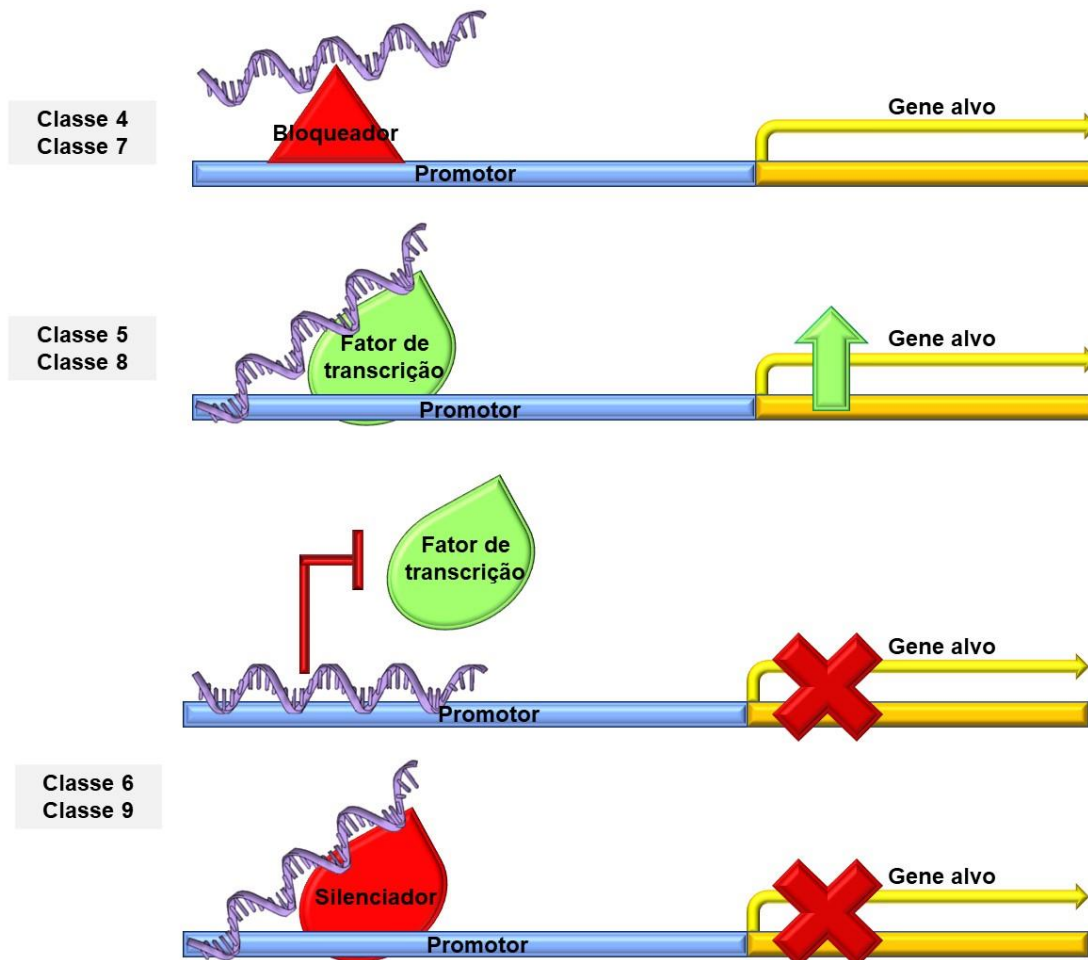


Figura 7: representação da regulação transcricional mediada por interação em tripla hélice na região promotora (azul) do gene alvo (amarelo)

As classes 4 e 7 têm em comum o fato de o gene alvo não possuir expressão diferencial significativa quando em estresse por etanol, sugerindo a ação de um bloqueador impedindo que a interação em tripla hélice ocorra. Nas classes 5 e 8 o gene alvo está sobre-expresso, o que indica que o lncRNA está agindo sobre a região promotora e favorecendo a transcrição, provavelmente agindo como uma âncora para fatores de transcrição que orientam a RNA Pol II. Nas classes 6 e 9, o gene alvo está sub-expresso, indicando que a formação em tripla hélice age como um bloqueador para fatores de transcrição, reduzindo

a afinidade da RNA Pol II sobre o promotor destes genes. Este fenômeno deve ocorrer principalmente por impedimento alostérico, já que os lncRNAs são moléculas grandes que podem impedir que fatores de transcrição ou a própria RNA Pol II consiga chegar ao DNA de forma eficiente para que a transcrição ocorra.

Já foi reportada a atuação dos lncRNAs na ativação ou orientação da transcrição de forma significativa (classes 5 e 8). Em células tronco embrionárias humanas, o lncRNA *Haunt* tem papel significativo na ativação do gene HOXA, um gene essencial para diferenciação celular durante o período embrionário. Este lncRNA possivelmente atua como um enhancer na região promotora de HOXA, favorecendo a atuação dos fatores de transcrição (YIN et al., 2015). Outro exemplo interessante, é o lncRNA *Chaserr*, células tronco embrionárias de ratos atua sobre o gene Chd2 direcionando o aparato de transcrição sobre um start códon diferente do que seria direcionado sem a presença do lncRNA (ALI; GROTE, 2020; ROM et al., 2019).

Além de favorecer a transcrição, os lncRNAs podem atuar de diferentes formas silenciando genes (classe 6 e 9), como é o caso do lncRNA COLDAIR. Em *Arabidopsis thaliana* que durante o frio, esse lncRNA atua impedindo a floração através do recrutamento da proteína PCR2 na região promotora do gene FLC. O complexo formado por COLDAIR:PCR2 silencia o gene FLC e modula a floração (HEO; SUNG, 2011).

Ressalta-se aqui o transcr_10944 da linhagem S288C. Este lncRNA parece estar formando estruturas em triplas hélices com vários lncRNAs das diferentes classes. Da classe 5, as interações são com os genes ATP1 (YBL099W), RME1 (YGR044C), TAM41 (YGR046W), COQ2 (YNR041C), SPO24 (YPR036W-A) e IRC16 (YPR038W). Estes genes fazem parte de estruturas ou vias similares ou conectadas, sendo os genes ATP1 (ATP-sintase) e TAM41 genes importantes para funcionamento da mitocôndria, enquanto que os genes RME1 e SPO24 são essenciais para a regulação da meiose (DEVENISH et al., 2000; RINE; GF; HERSKOWITZ, 1981; TKACH et al., 2012). Pertencentes à classe 7, observa-se o transcr_10944 interagindo com os genes SAC3 (YDR159W), TPA1 (YER049W) e YGR045C, sendo SAC3 um gene de exportação de mRNA essencial para a biogênese dos ribossomos (LEI et al., 2003), o TPA1 atua na síntese de ribossomos e estabilidade de mRNAs (KEELING et al., 2006; SHIVANGE et al., 2014), e o YGR045C não possui função

conhecida. O lncRNA transcr_10944 é o único da linhagem S288C que possui interações pertencentes à classe 9, sendo alvos os genes: ANP1 (YEL036C), HYP2 (YEL034W), MTC7 (YEL033W), RPL2A (YFR031C-A), ERG11 (YHR007C), YNL089C, TOP2 (YNL088W) e AGA1 (YNR044W). Entre estas 8 interações, temos o gene HYP2, um importante fator de alongamento da tradução (SAINI et al., 2009), o RPL2A, uma proteína que faz parte da subunidade ribossomal 60S (MESKAUSKAS; RUSS; DINMAN, 2008) e TOP2, uma topoisomerase essencial para duplicação do DNA (GITTENS et al., 2019). Este caso reporta como um lncRNA pode interferir em diferentes vias dentro da célula, além de indicar que as interações em tripla hélice podem ter um papel essencial para determinadas etapas do metabolismo e do ciclo celular.

CONSIDERAÇÕES FINAIS

O trabalho apresentado tem como objeto contribuir para a compreensão função e o impacto dos lncRNAs no metabolismo de *S. cerevisiae*, determinar se os lncRNAs são determinantes para a tolerância ao etanol e para a recuperação da célula após o estresse.

No primeiro capítulo, reportamos que os lncRNAs possuem um papel importante para a tolerância ao etanol. A deleção dos lncRNAs transcr_9136 e transcr_10027 causou alterações no crescimento e recuperação das linhagens pós o estresse, mudanças nas vias de resposta ao estresse como PBs e SGs, e também modificações na expressão de genes próximos.

No segundo capítulo, por meio de análises de tripla hélice, foi possível compreender melhor um dos mecanismos de regulação que envolve os lncRNAs, os quais vários transcritos atuam diretamente na região promotora dos genes.

Dito isto, de diferentes formas, os lncRNAs mostram-se moléculas importantes dentro da célula, atuando direta e indiretamente nas vias de transcrição e tradução, o que os torna candidatos para estudos posteriores para o entendimento do metabolismo de leveduras.

REFERÊNCIAS

ABULWERDI, F. A.; XU, W.; AGEELI, A. A.; YONKUNAS, M. J.; ARUN, G.; NAM, H.; SCHNEEKLOTH, J. S.; DAYIE, T. K.; SPECTOR, D.; BAIRD, N.; LE GRICE, S. F. J. Selective Small-Molecule Targeting of a Triple Helix Encoded by the Long Noncoding RNA, MALAT1. **ACS Chemical Biology**, [s. l.], 2019.

ALI, T.; GROTE, P. Beyond the RNA-dependent function of LncRNA genes. **eLife**, [s. l.], v. 9, n. 3, p. 1–14, 2020.

ANDERSON, D. M.; ANDERSON, K. M.; CHANG, C. L.; MAKAREWICH, C. A.; NELSON, B. R.; MCANALLY, J. R.; KASARAGOD, P.; SHELTON, J. M.; LIOU, J.; BASSEL-DUBY, R.; OLSON, E. N. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. **Cell**, [s. l.], v. 160, n. 4, p. 595–606, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.cell.2015.01.009>>

BEREZIKOV, E.; GURYEV, V.; CUPPEN, E. CONREAL web server: Identification and visualization of conserved transcription factor binding sites. **Nucleic Acids Research**, [s. l.], v. 33, n. SUPPL. 2, p. 447–450, 2005.

BHAN, A.; MANDAL, S. S. LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. **Biochimica et Biophysica Acta - Reviews on Cancer**, [s. l.], v. 1856, n. 1, p. 151–164, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.bbcan.2015.07.001>>

BROWN, J. A.; BULKLEY, D.; WANG, J.; VALENSTEIN, M. L.; YARIO, T. A.; STEITZ, T. A.; STEITZ, J. A. Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. **Nature Structural and Molecular Biology**, [s. l.], v. 21, n. 7, p. 633–640, 2014.

BUCHAN, J. R.; MUHLRAD, D.; PARKER, R. P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. **Journal of Cell Biology**, [s. l.], v. 183, n. 3, p. 441–455, 2008.

BUCHAN, J. R.; YOON, J. H.; PARKER, R. Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. **Journal of Cell Science**, [s. l.], v. 124, n. 2, p. 228–239, 2011.

BUSKE, F. A.; BAUER, D. C.; MATTICK, J. S.; BAILEY, T. L. Triplexator: Detecting nucleic acid triple helices in genomic and transcriptomic data. **Genome Research**, [s. l.], v. 22, n. 7, p. 1372–1381, 2012.

CHEKANOVA, J. A. Long non-coding RNAs and their functions in plants. **Current Opinion in Plant Biology**, [s. l.], v. 27, p. 207–216, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.pbi.2015.08.003>>

CHENG, J. J.; TIMILSINA, G. R. Status and barriers of advanced biofuel technologies: A review. **Renewable Energy**, [s. l.], v. 36, n. 12, p. 3541–3549, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.renene.2011.04.031>>

D, □; INÉS FERNÁNDEZ-ULIBARRI, □ V; VILELLA, M.; LÁZARO-DIÉGUEZ, F.; SARRI, E.; MARTÍNEZ, S. E.; JIMÉNEZ, N.; CLARO, E.; MÉRIDA, I.; BURGER, K. N. J.; EGEA,

G. Diacylglycerol Is Required for the Formation of COPI Vesicles in the Golgi-to-ER Transport Pathway. **Molecular Biology of the Cell**, [s. l.], v. 18, n. June, p. 3250–3263, 2007.

DELLOMONACO, C.; FAVA, F.; GONZALEZ, R. *Path Of Next Generation Biofuels* fulltext.Pdf. [s. l.], p. 1–15, 2010.

DEMMELE, L.; BECK, M.; KLOSE, C.; SCHLAITZ, A.-L.; GLOOR, Y.; HSU, P. P.; HAVLIS, J.; SHEVCHENKO, A.; KRAUSE, E.; KALAIIDZIDIS, Y.; WALCH-SOLIMENA, C. Nucleocytoplasmic shuttling of the Golgi phosphatidylinositol 4-kinase Pik1 is regulated by 14-3-3 proteins and coordinates Golgi function with cell growth. **Molecular biology of the cell**, [s. l.], v. 19, n. 3, p. 1046–1061, 2008.

DEVENISH, R. J.; PRESCOTT, M.; ROUCOU, X.; NAGLEY, P. Insights into ATP synthase assembly and function through the molecular genetic manipulation of subunits of the yeast mitochondrial enzyme complex. **Biochimica et Biophysica Acta - Bioenergetics**, [s. l.], v. 1458, n. 2–3, p. 428–442, 2000.

ESCUDEÉ, C.; FRANÇOIS, J. christophe; SUN, J. sheng; OTT, G.; SPRINZL, M.; GARESTIER, T.; HEÉLENE, J. christophe. Stability of triple helices containing RNA and DNA strands: Experimental and molecular modeling studies. **Nucleic Acids Research**, [s. l.], v. 21, n. 24, p. 5547–5553, 1993.

FANG, Y.; FULLWOOD, M. J. **Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer**. [s.l.] : Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China, 2016. v. 14 Disponível em: <<http://dx.doi.org/10.1016/j.gpb.2015.09.006>>

FERRÉ, F.; COLANTONI, A.; HELMER-CITTERICH, M. Revealing protein-lncRNA interaction. **Briefings in Bioinformatics**, [s. l.], v. 17, n. 1, p. 106–116, 2016.

GALIPON, J.; MIKI, A.; ODA, A.; INADA, T.; OHTA, K. Stress-induced lncRNAs evade nuclear degradation and enter the translational machinery. **Genes to Cells**, [s. l.], v. 18, n. 5, p. 353–368, 2013.

GITTENS, W. H.; JOHNSON, D. J.; ALLISON, R. M.; COOPER, T. J.; THOMAS, H.; NEALE, M. J. A nucleotide resolution map of Top2-linked DNA breaks in the yeast and human genome. **Nature Communications**, [s. l.], v. 10, n. 1, p. 1–16, 2019. Disponível em: <<http://dx.doi.org/10.1038/s41467-019-12802-5>>

GOLDEMBERG, J.; COELHO, S. T.; GUARDABASSI, P. The sustainability of ethanol production from sugarcane. **Energy Policy**, [s. l.], v. 36, n. 6, p. 2086–2097, 2008.

GOMES, A. Q.; REAL, C.; ANTUNES, F.; MARINHO, H. S.; NOLASCO, S.; SOARES, H. **Noncoding RNAs as Critical Players in Regulatory Accuracy, Redox Signaling, and Immune Cell Functions**. [s.l.] : Elsevier B.V., 2017. Disponível em: <<http://dx.doi.org/10.1016/B978-0-444-63660-7.00010-3>>

HEO, J. B.; SUNG, S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. **Science**, [s. l.], v. 331, n. 6013, p. 76–79, 2011.

HUBSTENBERGER, A.; COUREL, M.; BÉNARD, M.; SOUQUERE, S.; ERNOULT-

LANGE, M.; CHOUAIB, R.; YI, Z.; MORLOT, J. B.; MUNIER, A.; FRADET, M.; DAUNESSE, M.; BERTRAND, E.; PIERRON, G.; MOZZICONACCI, J.; KRESS, M.; WEIL, D. P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. **Molecular Cell**, [s. l.], v. 68, n. 1, p. 144- 157.e5, 2017.

IVANOV, P.; KEDERSHA, N.; ANDERSON, P. Stress granules and processing bodies in translational control. **Cold Spring Harbor Perspectives in Biology**, [s. l.], v. 11, n. 5, 2019.

JAKOČINAS, T.; BONDE, I.; HERRGÅRD, M.; HARRISON, S. J.; KRISTENSEN, M.; PEDERSEN, L. E.; JENSEN, M. K.; KEASLING, J. D. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. **Metabolic Engineering**, [s. l.], v. 28, p. 213–222, 2015.

JARROUX, J.; MORILLON, A.; PINSKAYA, M. **History, discovery, and classification of lncRNAs**. [s.l: s.n.]. v. 1008

JINEK, M.; CHYLINSKI, K.; FONFARA, I.; HAUER, M.; DOUDNA, J. A.; CHARPENTIER, E. A Programmable Dual-RNA – Guided. [s. l.], v. 337, n. August, p. 816–822, 2012.

KEELING, K. M.; SALAS-MARCO, J.; OSHEROVICH, L. Z.; BEDWELL, D. M. Tpa1p Is Part of an mRNP Complex That Influences Translation Termination, mRNA Deadenylation, and mRNA Turnover in *Saccharomyces cerevisiae* . **Molecular and Cellular Biology**, [s. l.], v. 26, n. 14, p. 5237–5248, 2006.

LÁZARI, L. C.; WOLF, I. R.; SCHNEPPER, A. P.; VALENTE, G. T. LncRNAs of *Saccharomyces cerevisiae* dodge the cell cycle arrest imposed by the ethanol stress. [s. l.], v. 6210, p. 1–29, 2021.

LEASE, R. A.; CUSICK, M. E.; BELFORT, M. Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. **Proceedings of the National Academy of Sciences of the United States of America**, [s. l.], v. 95, n. 21, p. 12456–12461, 1998.

LEI, E. P.; STERN, C. A.; FAHRENKROG, B.; KREBBER, H.; MOY, T. I.; AEBI, U.; SILVER, P. A. Sac3 is an mRNA export factor that localizes to cytoplasmic fibrils of nuclear pore complex. **Molecular Biology of the Cell**, [s. l.], v. 14, n. 3, p. 836–847, 2003.

LIMA, N.; TEIXEIRA, A. Alcohol Production from Cheese Whey. **Biotechnology and Bioengineering**, [s. l.], v. 72, n. 5, p. 507–514, 2001.

LOPES, M. L.; PAULILLO, S. C. de L.; GODOY, A.; CHERUBIN, R. A.; LORENZI, M. S.; GIOMETTI, F. H. C.; BERNARDINO, C. D.; DE AMORIM NETO, H. B.; DE AMORIM, H. V. Ethanol production in Brazil: a bridge between science and industry. **Brazilian Journal of Microbiology**, [s. l.], v. 47, p. 64–76, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.bjm.2016.10.003>>

LUCAS FARINAZZO, M.; IVAN RODRIGO, W.; LAZARI, L. C.; ALMEIDA, L. F. De; SCHNEPPER, A. P.; CARDOSO, L. H.; MORAES, L. N. De; GROTTTO, R. M. T.; SIMÕES, R. P.; RAMOS, É.; VALENTE, G. T. Long non-coding RNAs bind to proteins relevant to the ethanol tolerance in yeast: a systems biology view. [s. l.], p. 1–27, 2021.

MESKAUSKAS, A.; RUSS, J. R.; DINMAN, J. D. Structure/function analysis of yeast ribosomal protein L2. **Nucleic Acids Research**, [s. l.], v. 36, n. 6, p. 1826–1835, 2008.

MUSSATTO, S. I.; DRAGONE, G.; GUIMARÃES, P. M. R.; SILVA, J. P. A.; CARNEIRO, L. M.; ROBERTO, I. C.; VICENTE, A.; DOMINGUES, L.; TEIXEIRA, J. A. Technological trends, global market, and challenges of bio-ethanol production. **Biotechnology Advances**, [s. l.], v. 28, n. 6, p. 817–830, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.biotechadv.2010.07.001>>

QIAO, F.; CECH, T. R. Triple-helix structure in telomerase RNA contributes to catalysis. **Nature Structural and Molecular Biology**, [s. l.], v. 15, n. 6, p. 634–640, 2008.

RINE, J.; GF, S.; HERSKOWITZ, I. rme1 Mutation of *Saccharomyces cerevisiae*: map position and bypass of mating type locus control of sporulation. **Molecular and cellular biology**, [s. l.], v. 1, n. 10, p. 958–960, 1981.

ROM, A.; MELAMED, L.; GIL, N.; GOLDRICH, M. J.; KADIR, R.; GOLAN, M.; BITON, I.; PERRY, R. B. T.; ULITSKY, I. Regulation of CHD2 expression by the Chaserr long noncoding RNA gene is essential for viability. **Nature Communications**, [s. l.], v. 10, n. 1, 2019. Disponível em: <<http://dx.doi.org/10.1038/s41467-019-13075-8>>

RYAN, O. W.; PODDAR, S.; CATE, J. H. D. Crispr–cas9 genome engineering in *Saccharomyces cerevisiae* cells. **Cold Spring Harbor Protocols**, [s. l.], v. 2016, n. 6, p. 525–533, 2016.

SAINI, H. K.; ENRIGHT, A. J.; GRIFFITHS-JONES, S. Annotation of mammalian primary microRNAs. **BMC Genomics**, [s. l.], v. 9, p. 1–19, 2008.

SAINI, P.; EYLER, D. E.; GREEN, R.; DEVER, T. E. Hypusine-containing protein eIF5A promotes translation elongation. **Nature**, [s. l.], v. 459, n. 7243, p. 118–121, 2009.

SHAH, K. H.; ZHANG, B.; RAMACHANDRAN, V.; HERMAN, P. K. Processing body and stress granule assembly occur by independent and Differentially regulated pathways in *Saccharomyces cerevisiae*. **Genetics**, [s. l.], v. 193, n. 1, p. 109–123, 2013.

SHIVANGE, G.; KODIPELLI, N.; MONISHA, M.; ANINDYA, R. A role for *saccharomyces cerevisiae* Tpa1 protein in direct alkylation repair. **Journal of Biological Chemistry**, [s. l.], v. 289, n. 52, p. 35939–35952, 2014.

SNOEK, T.; VERSTREPEN, K. J.; VOORDECKERS, K. How do yeast cells become tolerant to high ethanol concentrations? **Current Genetics**, [s. l.], v. 62, n. 3, p. 475–480, 2016.

SNOWDON, C.; SCHIERHOLTZ, R.; POLISZCZUK, P.; HUGHES, S.; VAN DER MERWE, G. ETP1/YHL010c is a novel gene needed for the adaptation of *Saccharomyces cerevisiae* to ethanol. **FEMS Yeast Research**, [s. l.], v. 9, n. 3, p. 372–380, 2009.

STANLEY, D.; BANDARA, A.; FRASER, S.; CHAMBERS, P. J.; STANLEY, G. A. The ethanol stress response and ethanol tolerance of *Saccharomyces cerevisiae*. **Journal of Applied Microbiology**, [s. l.], v. 109, n. 1, p. 13–24, 2010. a.

STANLEY, D.; CHAMBERS, P. J.; STANLEY, G. A.; BORNEMAN, A.; FRASER, S.

Transcriptional changes associated with ethanol tolerance in *Saccharomyces cerevisiae*. **Applied Microbiology and Biotechnology**, [s. l.], v. 88, n. 1, p. 231–239, 2010. b.

STATELLO, L.; GUO, C. J.; CHEN, L. L.; HUARTE, M. Gene regulation by long non-coding RNAs and its biological functions. **Nature Reviews Molecular Cell Biology**, [s. l.], v. 22, n. 2, p. 96–118, 2021. Disponível em: <<http://dx.doi.org/10.1038/s41580-020-00315-9>>

SUN, J. S.; GARESTIER, T.; HÉLÈNE, C. Oligonucleotide directed triple helix formation. **Current Opinion in Structural Biology**, [s. l.], v. 6, n. 3, p. 327–333, 1996.

SZCZEŚNIAK, M. W.; MAKALOWSKA, I. LncRNA-RNA interactions across the human transcriptome. **PLoS ONE**, [s. l.], v. 11, n. 3, p. 1–11, 2016.

TKACH, J. M.; YIMIT, A.; LEE, A. Y.; RIFFLE, M.; COSTANZO, M.; JASCHOB, D.; HENDRY, J. A.; OU, J.; MOFFAT, J.; BOONE, C.; DAVIS, T. N.; NISLOW, C.; BROWN, G. W. Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. **Nature Cell Biology**, [s. l.], v. 14, n. 9, p. 966–976, 2012. Disponível em: <<http://dx.doi.org/10.1038/ncb2549>>

TZENG, Y. W.; HUANG, J. N.; SCHUYLER, S. C.; WU, C. H.; JUANG, Y. L. Functions of the mitotic B-type cyclins CLB1, CLB2, and CLB3 at mitotic exit antagonized by the CDC14 phosphatase. **Fungal Genetics and Biology**, [s. l.], v. 48, n. 10, p. 966–978, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.fgb.2011.07.001>>

VAUDANO, E.; COSTANTINI, A.; NOTI, O.; GARCIA-MORUNO, E. An RT-qPCR approach to study the expression of genes responsible for sugar assimilation during rehydration of active dry yeast. **Food Microbiology**, [s. l.], v. 27, n. 6, p. 802–808, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.fm.2010.04.010>>

WANG, C. Y.; WEN, W. L.; NILSSON, D.; SUNNERHAGEN, P.; CHANG, T. H.; WANG, S. W. Analysis of stress granule assembly in *Schizosaccharomyces pombe*. **Rna**, [s. l.], v. 18, n. 4, p. 694–703, 2012.

WANG, G. Q.; WANG, Y.; XIONG, Y.; CHEN, X. C.; MA, M. L.; CAI, R.; GAO, Y.; SUN, Y. M.; YANG, G. S.; PANG, W. J. Sirt1 AS lncRNA interacts with its mRNA to inhibit muscle formation by attenuating function of miR-34a. **Scientific Reports**, [s. l.], v. 6, n. October 2015, p. 1–13, 2016.

WANG, J.; TAN, J.; QI, Q.; YANG, L.; WANG, Y.; ZHANG, C.; HU, L.; CHEN, H.; FANG, X. MiR-487b-3p suppresses the proliferation and differentiation of myoblasts by targeting IRS1 in skeletal muscle myogenesis. **International Journal of Biological Sciences**, [s. l.], v. 14, n. 7, p. 760–774, 2018.

WOLF, I. R.; MARQUES, L. F.; ALMEIDA, L. F. De; LÁZARI, L. C.; MORAES, L. N. De; CARDOSO, L. H.; OLIVEIRA, C. C. De; NAKAJIMA, R. T.; SCHNEPPER, A. P.; GOLIM, M. de A.; CATALDI, T. R.; NIJLAND, J. G.; PINTO, C. M.; FIORETTO, M. N.; ALMEIDA, R. O.; DRIESSEN, A. J. M.; SIMÕES, R. P.; LABATE, M. V.; TOMMASINI, REJANE M GROTTO, C. A. L.; JUNIOR, A. F.; JUSTULIN, L. A.; COAN, R. L. B.; RAMOS, É.; FURTADO, F. B.; MARTINS, C.; VALENTE, G. T. The ethanol tolerance in *Saccharomyces cerevisiae* under a phenomics perspective. **bioRxiv**, [s. l.], p. 1–41, 2021.

YAMASHITA, A.; SHICHINO, Y.; YAMAMOTO, M. The long non-coding RNA world in yeasts. **Biochimica et Biophysica Acta - Gene Regulatory Mechanisms**, [s. l.], v. 1859, n. 1, p. 147–154, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.bbagrm.2015.08.003>>

YIN, Y.; YAN, P.; LU, J.; SONG, G.; ZHU, Y.; LI, Z.; ZHAO, Y.; SHEN, B.; HUANG, X.; ZHU, H.; ORKIN, S. H.; SHEN, X. Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. **Cell Stem Cell**, [s. l.], v. 16, n. 5, p. 504–516, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.stem.2015.03.007>>

YOU, K. M.; ROSENFELD, C. L.; KNIPPLE, D. C. Ethanol tolerance in the yeast *Saccharomyces cerevisiae* is dependent on cellular oleic acid content. **Applied and Environmental Microbiology**, [s. l.], v. 69, n. 3, p. 1499–1503, 2003.

ZHANG, M.; XIAO, Y.; ZHU, R.; ZHANG, Q.; WANG, S. L. Enhanced thermotolerance and ethanol tolerance in *Saccharomyces cerevisiae* mutated by high-energy pulse electron beam and protoplast fusion. **Bioprocess and Biosystems Engineering**, [s. l.], v. 35, n. 9, p. 1455–1465, 2012.

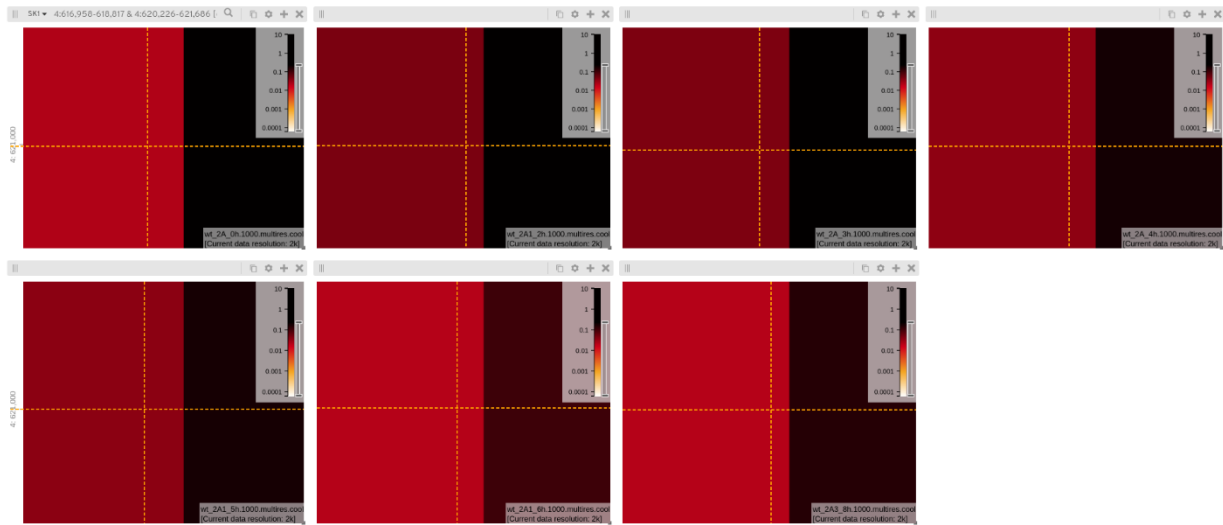
ZHOU, Z.; GILES, K. E.; FELSENFELD, G. DNA-RNA triple helix formation can function as a cis-acting regulatory mechanism at the human β -globin locus. **Proceedings of the National Academy of Sciences of the United States of America**, [s. l.], v. 116, n. 13, p. 6130–6139, 2019.

APÊNDICE A

Apêndice A: Descrição e tolerância máxima ao etanol por linhagem

Strain	EtOH tol. (%)	Aces. num.	MAT locus	Genotype	Phenotype
BMA64-1A	30	Euroscarf/20000A	a	MATa; his3-11_15; leu2-3_112; ura3-1; trp1D2; ade2-1; can1-100	HT
BY4741	22	Euroscarf/Y00000 SGD/BY4741	a	S288C isogenic yeast strain: MATa; his3D1; leu2D0; met15D0; ura3D0	LT
BY4742	26	Euroscarf/Y10000 SGD/BY4742	alpha	S288C isogenic yeast strain: MATalpha; his3D1; leu2D0; lys2D0; ura3D0	HT
S288C	20	NBRP/BY20118 SGD/S288c	alpha	MAT α SUC2 mal mel gal2CUP1	LT
SEY6210	20	NBRP/BY3553 SGD/SEY6210	alpha	MAT α suc2- Δ 9 ura3-52 leu2-3,112 his3- Δ 200 trp1- Δ 901 lys2-801	LT
X2180-1A	24	NBRP/BY21559 SGD/X2180-1A	a	MATa SUC2 mel gal2 CUP1	HT

APÊNDICE B



Análise de dados de HI-C da linhagem S288C no decorrer do ciclo celular. Em destaque, a região de interação entre o lncRNA *transcr_9136* e o gene *RRP1*.

APÊNDICE C

Atualmente, o grupo vem estudando os lncRNAs de *S. cerevisiae* responsivos em etanol em várias vertentes. Além dos novos resultados aqui apresentados, reportamos que alguns lncRNAs são determinantes na tolerância ao etanol, uma vez que suas deleções reduzem essa tolerância. Além disso, evidenciou-se que esses lncRNAs atuam no ciclo celular causando uma parada do ciclo celular, resultado dos danos ao DNA oriundos do estresse por etanol. Também, reporta-se a participação dos lncRNAs em muitos outros processos e estruturas celulares, principalmente àqueles relacionados aos processos ao acúmulo e degradação de RNAs e proteínas (LÁZARI et al., 2021; LUCAS FARINAZZO et al., 2021; WOLF et al., 2021).

Ao longo desse processo, várias técnicas foram empregadas para comprovar as evidências desses lncRNAs, resultados esses reportados nesse apêndice.

Comprovação da presença dos lncRNAs e avaliação da expressão gênica

Para avaliar a expressão gênica dos lncRNAs, todas as linhagens de estudadas pelo grupo foram crescidas e tratadas da mesma forma como reportado no capítulo 2. Os controles experimentais foram realizados utilizando solução fisiológica correspondente a porcentagem de etanol utilizada no tratamento. O experimento foi conduzido em triplicata (Figura 1). Os RNAs, cDNAs e qPCR foram conduzidos tal como descrito no capítulo 2. Porém, o gene endógeno aqui utilizado foi o ACT1 (YFL039C) (VAUDANO et al., 2010). Os primers utilizados para a avaliação da expressão gênica estão descritos na Tabela 1.

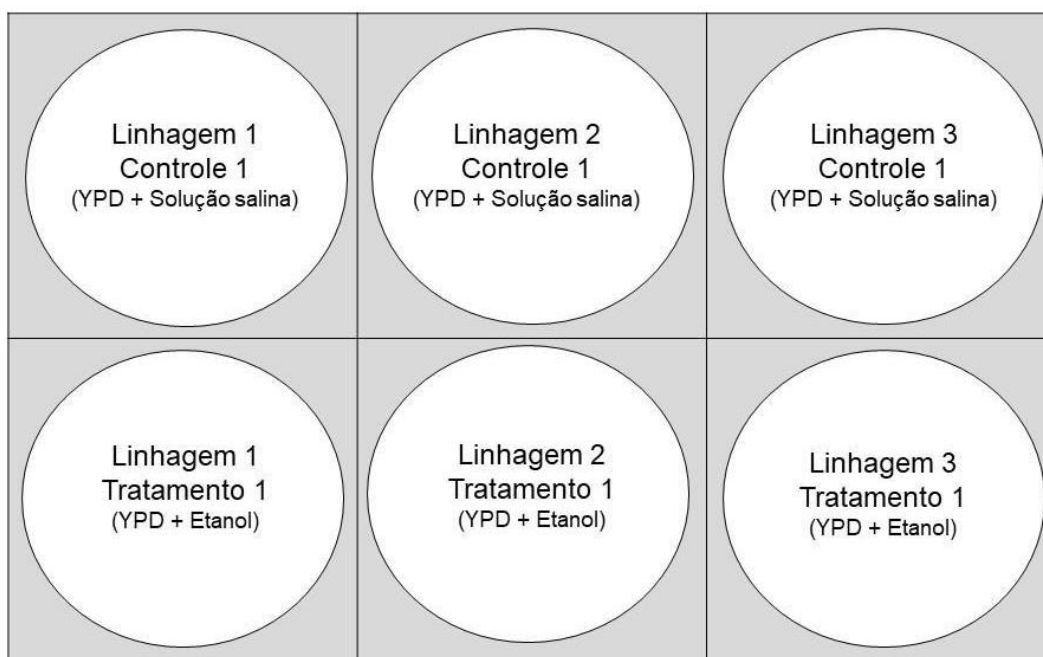


Figura 1: Representação da montagem das placas de 6 poços para controle e tratamento das linhagens de *Saccharomyces cerevisiae*.

Tabela 1: oligos utilizados para a validação da expressão diferencial nas linhagens de *S. cerevisiae*

Linhagem	IncRNA/Gene		Primer (5' → 3')
S288c	transc_14670	Fw	CGGGCCCCGGAAC TTTATAA
		Rv	TACTCCTTTGGGGTGGGTCC
BY4742	transc_10883	Fw	TCGCATCCACGTTAGCATGA
		Rv	GTTTGAACGTCGCCTTGTGT
X2180-1A	transc_9485	Fw	TGGCCTCTCTCCATTCGGTA
		Rv	TCGTACGTGGGCACAATGAA
BY4741	transc_8174	Fw	AGATTCTCGCACGACAGCAT
		Rv	TGGAGTTCGGGGAAAGATCG
SEY6210	transc_3536	Fw	CAATACGGCCATGTTGAGCG
		Rv	GTGCTGGCCTCGATAGGTAC
BMA64-1A	transc_20548	Fw	TGTAGAGCTTTTCCCGGCAT
		Rv	CGGTGATGAATATGAGCCGC
Endógeno	ACT1	Fw	AATCACCGCTTTGGCTCCAT
		Rv	AGAACCACCAATCCAGACGG

Em paralelo, o cDNA da linhagem S288C foi submetido a PCR para avaliar a veracidade do comprimento do lncRNA transcr_28768 (2739 nts), utilizando a reação: 2,5 µL de tampão, 0,75 µL de MgCl₂, 0,6µL de dNTP (10 mM) (Invitrogen), 0,6 µL de primer Fw (10 µM), 0,6 µL de primer Rv (10 µM), 2 µL de cDNA, 0,5 µL de Platinum® Taq DNA Polymerase (ThermoFisher) e 17,45 uL de água, com volume final de 25 µL. A reação de PCR utilizada foi 95 ° C por 1 min, seguido de 40 ciclos de 95 ° C por 30 s, 59 ° C por 30 s e 72 ° C por 2 min 30 s, finalizando com 72 ° C por 7 min, utilizando o termociclador GeneAmp® PCR System 9700 (Applied Biosystems). O produto de PCR foi analisado em gel de agarose 1,5% corado com GelRed (Sigma) e azul de bromofenol.

A seguir, o produto de PCR foi purificado utilizando Wizard® SV Gel and PCR Clean-Up System (Promega) e a sequência foi obtida através de sequenciamento automático na plataforma ABI3730XL. A sequência foi analisada no software Geneious 7.1.

Confirmação da presença dos lncRNAs em Saccharomyces cerevisiae

Anteriormente, os lncRNAs foram preditos pelo grupo de pesquisa utilizando dados de RNAseq das 6 linhagens de *Saccharomyces cerevisiae* (LUCAS FARINAZZO et al., 2021; WOLF et al., 2021). Este trabalho consistiu em realizar a comprovação da existência e da expressão diferencial dos lncRNAs preditos por bioinformática, através da extração de RNA e análise por RTqPCR. Os resultados obtidos para um lncRNA de cada linhagem, bem como sua comparação com o foldchange obtido via análise de RNAseq estão descritos na Figura 2.

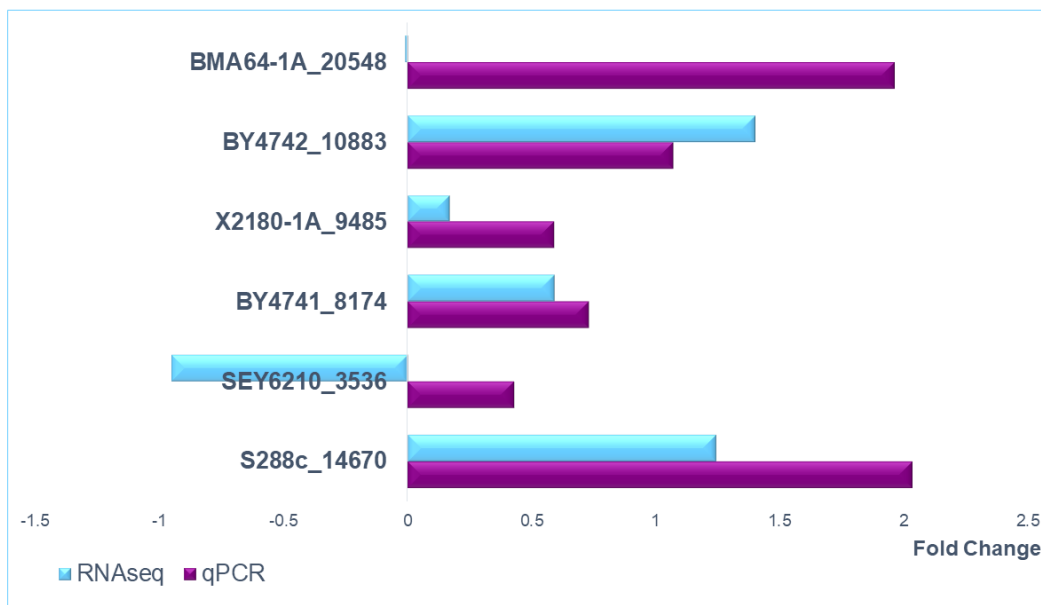


Figura 2: Dados de foldchange obtidos através de qPCR (roxo) e RNAseq (azul) para um lncRNA de cada linhagem de *Saccharomyces cerevisiae*

Em paralelo, foi realizada amplificação do transcr_28278 da S288c a partir do cDNA, sendo este o maior lncRNA descrito para esta linhagem. Este lncRNA se mostrou presente tanto nas condições controle e tratamento (Figura 3), e sua sequência foi confirmada via sequenciamento SANGER.

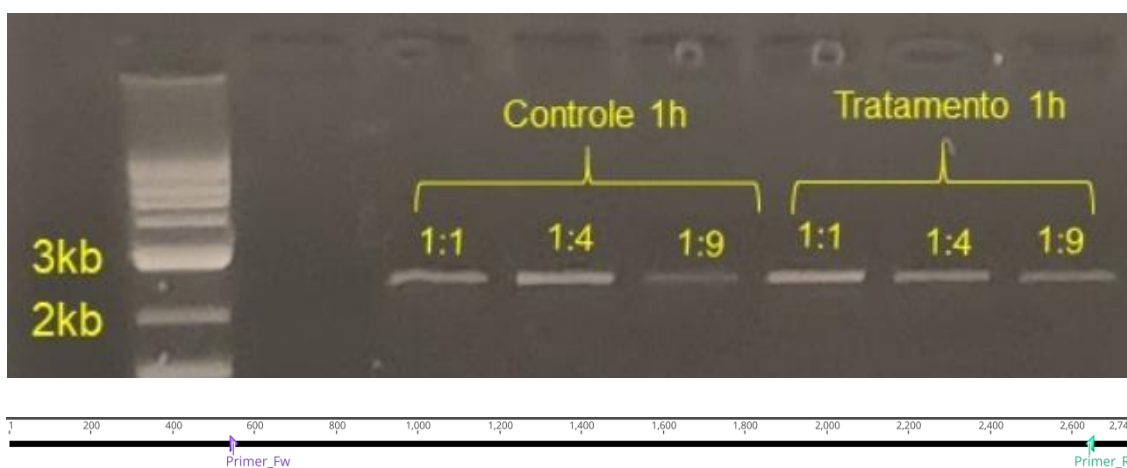


Figura 3: Bandas de identificação do lncRNA transcr_28278 da S288c, do tamanho esperado de 2744pb. Em roxo a localização do primer Fw e em verde a localização do primer Rv que foram utilizados amplificação e sequenciamento.

Os dados de RTqPCR dos lncRNAs das seis linhagens de *S. cerevisiae*, juntamente com os dados de sequenciamento do transcr_28278, corroboram as previsões realizadas pelo grupo (LUCAS FARINAZZO et al., 2021) quanto a presença dos lncRNAs, bem como os níveis de expressão destes transcritos.