Long non-coding RNAs bind to proteins relevant to the ethanol tolerance in yeast: a systems biology view

Lucas Farinazzo Marques[1], Ivan Rodrigo Wolf[1], Lucas Cardoso Lazari[1], Lauana Fogaça de Almeida[1], Amanda Piveta Schnepper[1], Luiz Henrique Cardoso[1], Leonardo Nazário de Moraes[1], Rejane Maria Tommasini Grotto[1], Rafael Plana Simões[1], Érica Ramos[2], Guilherme Targino Valente[1,3]

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

2- Department of Morphology. São Paulo State University (UNESP). Bioscience Institute, Botucatu, SP. Brazil.

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

Abstract

The ethanol disturbs the cell cycle, transcription, translation, protein folding, cell wall, membranes, and many *Saccharomyces cerevisiae* metabolic processes. Long non-coding RNAs (lncRNAs) are regulatory molecules binding onto the genome or proteins. The number of lncRNAs described for yeast is still scarce, and little is known concerning their roles in the system. There is a lack of knowledge concerning how lncRNAs are responsive to the ethanol tolerance in yeast and whether they act in this tolerance. Hence, by using RNA-Seq data from *S. cerevisiae* strains with different ethanol tolerance phenotypes, we found the severe ethanol responsive lncRNAs. We modeled how they participate in the ethanol tolerance by analyzing lncRNA-protein interactions. The results showed that the EtOH tolerance responsive lncRNAs, in both higher tolerant and lower tolerant phenotypes, work on different pathways: cell wall, cell cycle, growth, longevity, cell surveillance, ribosome biogenesis, intracellular transport, trehalose metabolism, transcription, and nutrient shifts. In summary, lncRNAs seems to interconnect essential systems' modules to overcome the ethanol stress. Finally, here we also found the most extensive catalog of lncRNAs in yeast.

Introduction

Biofuels can be produced from different sources (Demirbas 2017). Bioethanol is the most important biofuel as the most promising gasoline substitute (Chakraborty *et al.* 2012; Gupta and Verma 2015). The yeast *Saccharomyces cerevisiae* is the most used organism for bioethanol production (Mussatto *et al.* 2010; Demeke *et al.* 2013). Thus, understanding relevant factors to improve ethanol yield is essential for an ever-growing environment that increasingly demands more fuel.

However, increasing the ethanol (EtOH) concentration compromises the yeast survival affecting its growth rate and production (Stanley *et al.* 2010; Auesukaree 2017). For instance, EtOH rapidly affects the yeast's plasma membrane integrity (Ding *et al.* 2009; Ma and Liu 2010; Navarro-Tapia *et al.* 2016), leading to protein dysfunction and denaturation, affecting the molecules intake (e.g., glucose and amino acids), and causing an efflux of nucleotides and potassium (Ding *et al.* 2009; Ma and Liu 2010).

Experiments overexpressing or repressing candidate genes change the EtOH tolerance in yeast (Alper *et al.* 2006; Teixeira *et al.* 2009; Mussatto *et al.* 2010; Lewis *et al.* 2010; Swinnen *et al.* 2012). However, the complex molecular mechanisms concerning this pathway are still poorly understood.

The long non-coding RNAs (lncRNAs) promptly respond to external stimuli (Yamashita *et al.* 2016), regulating the gene expression and epigenetic modifications (Anderson *et al.* 2015). Furthermore, the few known cases concerning lncRNA-protein interactions suggest that lncRNAs work as a framework for macromolecular complexes assembly, or bait transcription factors dampening the association of these proteins to the DNA, or guide the chromatin modifiers (Tripathi *et al.* 2010; Geisler and Coller 2013; Ferrè *et al.* 2016; Li *et al.* 2019).

The lncRNA description for *S. cerevisiae* is scarce. Only 18 lncRNAs are appropriately described and annotated for this species (Till *et al.* 2018). These lncRNAs are directly involved in metabolic changes, sexual differentiation initiation, and other unknown processes (Yamashita *et al.* 2016). Recently, experiments showed that four ncRNAs affect the transcriptional systems as a whole mediated by a *trans* effect on transcription factors. These ncRNAs may be associated with the EtOH tolerance and stress response (Balarezo-Cisneros *et al.* 2020). However, little is known concerning the roles of lncRNA-protein interactions for all species.

Here we aimed to raise hypotheses concerning the relationship between lncRNAs and the EtOH tolerance. For this purpose, we focused on analyzing the interactions between the EtOH tolerance responsive lncRNAs and proteins. RNA-Seq data from *S. cerevisiae* strains with different EtOH tolerance phenotypes (higher and lower EtOH tolerant ones) were used to seek the lncRNAs expressed during the severe EtOH stress; in this case, we developed a pipeline to assemble these lncRNAs. Then, we analyzed how the lncRNAs impact the EtOH tolerance based on the prediction of lncRNA-protein interactions, guilt-by-association, and information flow throughout network approaches. The main pathways that EtOH stress-responsive LncRNAs work on are the cell wall, cell cycle and growth, cell longevity, cell surveillance, ribosome biogenesis, intracellular transport, trehalose metabolism, transcription, and nutrients shifts. Our findings indicate that EtOH stress-responsive lncRNAs interconnect essential systems' modules in a sort of strain-specific way to surpass a stressed environment's challenges. Altogether, here we provide insights and hypotheses on how lncRNAs may be working to the EtOH tolerance.

Material and Methods

Strains selection, ethanol tolerance definitions, and sequencing

The haploid strains BMA64-1A (Euroscarf/20000A), BY4742 (SGD/BY4742), X2180-1A (SGD/X2180-1A), BY4741 (SGD/BY4741), SEY6210 (SGD/SEY6210), and S288C (SGD/S288c) were obtained from Euroscarf (European *Saccharomyces cerevisiae* Archive for Functional Analysis) or NRBP (National Bioresources Project).

For the EtOH tolerance experiments, strains were grown overnight in YPD (2% of peptone, 1% of yeast extract, and 2% of glucose) and further diluted to an OD_{600} of 0.2. Then, 100 µL of cells were harvested by centrifugation at 2,000 RPM at 4°C for 5 min. Pellets were resuspended using YPD with different EtOH or physiological solution concentrations (the treatment and control condition, respectively) in plate-wells. Plates were incubated at 30°C for 1h and shaken at 120 RPM; the EtOH or physiological solution ranged from 2% to 32% (v/v). The content of each plate-well was plated on YPD and incubated at 30°C. Visual inspection allowed to set up the highest EtOH tolerance level supported for each strain.

Total RNA of cells under control and treatment conditions (the highest EtOH level supported per strain) was extracted from 1.5 mL of cells using Trizol after a cell wall digestion using Lyticase. The RNA quality was checked by electrophoresis in

agarose gel and Bioanalyzer, and Nanodrop and Qubit estimated the RNA quantification. The samples were treated with TURBO DNase (ThermoFisher), and rRNA depletion was performed. The transcriptome was obtained for 36 samples (6 strains x 2 (treatment and control) x 3 replicates). The genomic DNA of BMA64-1A was extracted using phenol-chloroform, and the quality and quantity were estimated using the equipment mentioned.

The RNA-Seq was performed by the LcScience (Texas, USA) Company using the Illumina HiSeq 4000 (100 nt, paired-end reads, insert size of 24-324 bp, and at least 40 million reads per sample); the company ensured the absence of small RNAs (<200 nts). The GenOne (Rio de Janeiro, Brazil) sequenced the genome using the Illumina HiSeq 2500 (250 nt, paired-end reads, 1 Gb of throughput, and insert size ~500 bp).

Genome assembling, lncRNA identification, and differential expression

The paired-end reads of the BMA64-1A genome were cleaned with Trimmomatic v.0.36 and independently assembled using AbySS v.2.0.2 (Jackman *et al.* 2017), IDBA v.1 (Peng *et al.* 2010), MIRA v.4.0.2, SPAdes v.3.10.1 (Bankevich *et al.* 2012) and Velvet v.1.2.10 (Zerbino and Birney 2008), varying parameters and assembling strategies. For an individual chromosome assembling, the reads were mapped against the reference genome (S288C version R64-2-1) using the HISAT2 v.2.1.0 (Kim *et al.* 2015), and the mapped reads were assembled using IDBA v.1 (Peng *et al.* 2010).

The genomic assembling metrics of each assembly were obtained with QUAST v. 4.5 (Gurevich *et al.* 2013). The QUAST values were normalized, generating a score from 0 to 1. Since QUAST generate many metrics for the assembling and reference, the score mentioned was calculated for each assembling considering all QUAST-normalized metrics (equation 1); the assembling with score value ≈0 was considered the most similar to the reference genome, and we assumed as the final data:

$$d = \sum_{i=1}^{n} \sqrt{\left(metric_j - metric_i\right)^2} \qquad \textbf{eq. 1}$$

where *metric_j* and *metric_i* are the metrics of the reference genome and a given assembling, respectively. To obtain *d*, all *n* metrics were considered.

For the BMA64-1A genome annotation, the transcripts of S288C from SGD were mapped over the best assembling using GMAP (Wu and Watanabe 2005), and

a *de novo* annotation was performed using MAKER v.2.32 (Cantarel *et al.* 2008). A manual curation comparing the two annotations was performed: 1- annotations present in both strategies were considered correct; 2- by visual inspection of the reference genome, the regions with disagreements passed by manual adjustments. Finally, we adjusted the annotations using the results from AGAPE (Song *et al.* 2015). The protein sequences were translated from annotated regions for further analysis.

The paired-end reads for all RNA-Seq libraries were trimmed using Trimmomatic v.0.36 (Bolger *et al.* 2014) and used to identify the lncRNAs.

We developed a pipeline to assembly the lncRNAs. Overall, the filtered reads were mapped over coding sequences of many different species. After, the non-mapped paired-reads were assembled using different algorithms, and a score was calculated to rank the best assembling. The transcript redundancies within strains were excluded, and the remaining transcripts were mapped over the genome of each strain to exclude spurious assembling. The mapped transcripts that may be coding sequences or mobile elements were excluded, and the remaining sequences were rechecked concerning their potential coding (Figure 1). Details of this pipeline are below described.
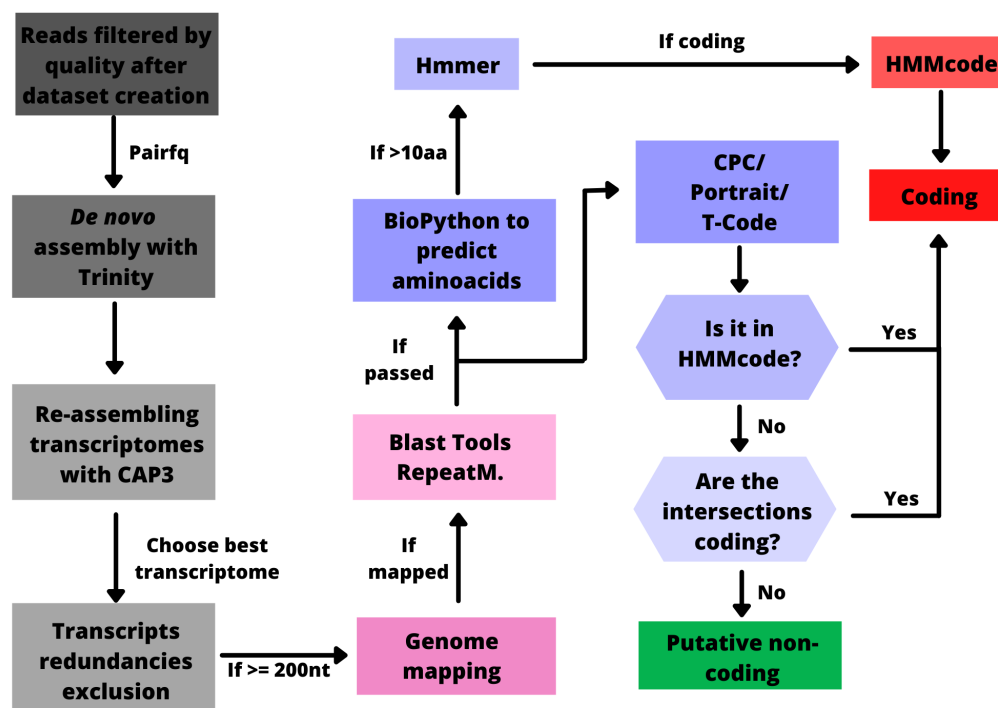

Figure 1: Pipeline to assembly the lncRNAs.

It was created a dataset with millions of sequences including: 1- coding sequences (CDS), and proteomes of eukaryotes and bacteria; 2- genomes; 3- microRNAs precursors; 4- ncRNA families present in the Rfam database (Burge *et al.* 2013), being the lncRNAs (Rfam accession number 01884) the only exception; 5- mobile elements. The RNA-Seq filtered reads of each strain were independently aligned on nucleotide sequences of the database mentioned using HISAT2 v.2.1.0 (Kim *et al.* 2015). Then, we assembled the non-aligned read-pairs selected by Pairfq script (The MIT License) since the ones are reads without similarity with coding, ncRNAs (excepting lncRNAs), mobile elements, and mitochondrial and contaminants genomes; reads that lost a member of the pair were excluded of this assembling (Figure 1).

The "Single Assembler Multiple Parameters" strategy (He *et al.* 2015) was used for the *de novo* assembling mentioned; it was performed parameter adjustments for each step using the S288C reads, and then, we applied the best parameter set to independently assembly the reads of all other strains. First, the Velvet/Oases (Zerbino and Birney 2008; Schulz *et al.* 2012), Trinity (Haas *et al.* 2013), IDBA-tran (Peng *et al.* 2012), and rnaSPAdes (Bankevich *et al.* 2012) were independently tested to determine which one was the best assembler for our dataset. For Velvet/Oases, rnaSPAdes, and IDBA-tran parameters, the *kmers* ranged from 19 to 81, and for Velvet/Oases and rnaSPAdes, we had set-up an automatic coverage cutoff and no scaffolding assembling. For Trinity assembler, we set-up the *kmers* ranging from 19 to 31. These programs were set-up to assemble only transcripts ≥200 nts when this option was available (Figure 1).

All assembled transcriptomes within each assembler were merged in a single file, sequences with ≥10% of "Ns" (non-identified nucleotides) were removed, followed by a generation of 1 transcriptome per assembler; this re-assembling was performed using CAP3 (Huang 1999). The singlets and contigs from CAP3 were merged, and the filtered reads (the ones used in the first assembling) were mapped over this file using Bowtie2 (Langmead and Salzberg 2012). A score was used (equation 2) to evaluate the final assembling, and the highest value indicates the best transcriptome. As mentioned, after these adjustments of the *de novo* assembling pipeline using the reads of S288C, the same settings were applied to independently assemble the lncRNAs of other strains (Figure 1).

$$\begin{cases} Score = (1 - AG)3 + P \times 3 + (\frac{PA \times 4}{100}) \\ \qquad AG = \frac{S+C}{ICAP} \\ \quad P = \frac{S \geq 200nts + C \geq 200nts}{S+C} \end{cases}$$ **eq. 2**

where *AG* (the assembling gain) is a rate of how many sequences the CAP3 used to assemble. The *P* is the percentage of CAP3 assembled sequences ≥200 nts, and the *PA* is the percentage of reads aligned using Bowtie2 over CAP3 assembling. The *S* and *C* are the numbers of sequences in the "singlets" and "contigs" CAP3's output, respectively. The *ICAP* is the number of sequences with ≤10% of "Ns" (assembled by IDBA, rnaSPAdes, Trinity, or Velvet/Oases) used as CAP3 input. The *PA* and *P* range from 0 to 100, while *AG* ranges from 0 to 1. The *AG* rates ≈0 are considered better values since it expresses that a higher number of input sequences was re-assembled by CAP3. The *P* and *PA* ≈1 are better, expressing that CAP3 assembled a higher number of large transcripts, and a higher number of reads was used by the first assemblers (the ones before CAP3), respectively.

The set of non-redundant transcripts found using CD-HIT (Li and Godzik 2006) were merged. The non-redundant sequences of each strain were mapped over its own genomes using GMAP (Wu and Watanabe 2005). The mapped transcripts were then rechecked concerning their potential to be coding, ncRNAs (except lncRNAs), repetitive elements, or contaminant genomes (viruses, bacteria, and mitochondria). For this purpose, the mapped transcripts were: 1- aligned against the proteomes using Blastx; 4- aligned against bacteria, viruses, and mitochondrial genomes using GMAP (Wu and Watanabe 2005). The data distribution of all results was analyzed to establish cutoffs to filter out undesirable sequences (Figure 1).

To verify the lack of coding regions on filtered sequence transcripts were protein translated (≥10 aa) using the Getorf (Rice *et al.* 2000), and transcripts/proteins were evaluated using Hmmer (Mistry *et al.* 2013) with Pfam database v.31.0 (Finn *et al.* 2014), Tcode (Rice *et al.* 2000), Portrait (Arrial *et al.* 2009) and CPC (Kong *et al.* 2007). The transcripts with motifs fitting Hmm models using Hmmer were assumed as potentially coding, whereas the other sequences that fit coding sequences using at least two out of the other three programs (Tcode, Portrait, or CPC) were defined as coding sequences. Hence, sequences not found neither by Hmmer nor by two other programs were assumed as putative lncRNAs (Figure 1).

The RNA-Seq filtered reads of each strain were mapped over its genomes using HISAT2 v.2.10 (Kim *et al.* 2015) ("--no-softclip"), the counting per transcript was obtained using the Bedtools multiBamCov v.2.26.0 (Quinlan and Hall 2010), and the lncRNAs differentially expressed comparing treatment *vs.* control was provided by DESeq2 (Love *et al.* 2014) considering a false discovery rate <0.01.

LncRNA-protein interaction prediction and analysis

It was predicted the lncRNAs-protein interactions (LNCPI), considering proteins ≥32 amino acids of each strain (downloaded from SGD and translated from the BMA64-1A genome assembling), using the lncPRO tool (Lu *et al.* 2013). The probability distributions of interactions follow a normal distribution, then the ones ≥0.95 of probability were selected for further analysis.

We downloaded yeast's protein-protein interactions (PPI) from Biogrid (Chatr-Aryamontri *et al.* 2013) and MINT (Licata *et al.* 2012), gene regulatory networks from the YTRP database (Yang *et al.* 2014), and metabolic networks from REACTOME (Fabregat *et al.* 2014). Only physical interactions among proteins were selected from Biogrid, and redundancies with MINT were excluded; we excluded the interactions present in the NEGATOME (Blohm *et al.* 2014). The networks were combined to create a single integrated network (Uninet). The Uninet was unified with the LNCPI of each strain generating 6 networks (one per strain and each of them including its LNCPI strain-specific) (equation 3). These networks were considered undirected graphs.

$$\begin{cases} G_k = (V_k, E_k) = LNCPI_k \cup Uninet = A_k \\ \qquad G'_k = (V'_k, E'_k) \\ \quad V'_k = \{v_i\} = \{gff_k\} \vee gff_k \subseteq GFF_k \\ \qquad E'_k = \{(v_i, v_j)\} \vee a_{ij} = 1 \end{cases} \qquad \textbf{eq.} \quad \textbf{3}$$

which G_k is an adjacency matrix A_k (where a_{ij} is an element of A_k) of a given strain (k). The G'_k is a subgraph from G_k. V'_k is a subset of nodes v_i from G'_k generated based on GFF annotation (gff_k). E'_k is a subset of edges of V'_k from G'_k. Hence, G'_k are strain-specific networks harboring their respective LNCPI as well.

To assess the influence of EtOH on the lncRNA target-protein and how the lncRNAs could help to the EtOH tolerance, we analyzed the information flow throughout LNCPI networks from each differentially expressed lncRNA (hereafter referred to as lncRNA-propagation analysis). For this analysis, we selected the up-regulated edges (log2 fold-change ≥ 1E-6) of each strain (the network is the G'_k

depicted in equation 3). Unconnected nodes or nodes with self-loops were excluded. The network propagation using the diffusion algorithm (Cowen *et al.* 2017) was assessed starting from all differential expressed lncRNAs. After, the lncRNAs and their first neighbors were selected, and only the top 20 ranked nodes from diffusion information were retrieved to compose new subgraphs (the nodes with self-loops or unconnected nodes were excluded again). Altogether, we could assess the participation of lncRNAs in the EtOH tolerance and whether this compound triggers similar effects on systems harboring different lncRNAs and interactions. The same processes were performed for lncRNAs and genes down-regulated (log2 fold-change ≤ -1E-6).

The gene ontology of lncRNA target-proteins was analyzed using g:Profiler (Reimand *et al.* 2016). The enriched terms were summarized by REVIGO (Supek *et al.* 2011). For this purpose, we considered all target-proteins indiscriminate neither by differential expression nor by strain.

Results

The highest EtOH tolerance level observed for each strain allowed classifying the ones as higher tolerant (HT) or lower tolerant (LT) phenotypes. The HT strains are BMA64-1A (tolerates 30% of EtOH), BY4742 (tolerates 26% of EtOH), and X2180-1A (tolerates 24% of EtOH). The strains BY4741 (tolerates 22% of EtOH), SEY6210 (tolerates 20% of EtOH), and S288C (tolerates 20% of EtOH) were considered LTs.

A total of 6,410,152 filtered reads with ~250 nts length allowed to achieve 196 assemblings of the BMA64-1A genome (Table 1). The best assembling has 11,880,801 bp, with 683 scaffolds, and with 134.88X of coverage. Only 36 out of 6,551 annotated coding genes did not present any similarity with the S288C genome (the reference genome).

Value	N50	Contigs/scaffolds	GC%	Genome size	Genes	AS
Minimum	1,437	17	37.93	126,580	1,744	0.054
Maximum	924,585	27,639	38.34	12,162,499	5,872	0.63
*Selected	202,259	683	38.06	11,880,801	5,668	0.29

Table 1: Metrics of BMA64-1A genome considering all assembling. ***:** metrics of selected assembling; **AS:** assembling score.

The best lncRNA assemblings pipeline uses the Trinity, followed by a re-assembling using CAP3. Although Velvet/Oases outperforms the Trinity (see the score in Table 2), the latter presented the highest computational performance. Hence, all lncRNA assembling used Trinity giving from 87 to 259 lncRNAs identified after the filtering steps (Table 3).

Pipeline	ICAP	AG	S+C	Avg. len.	P (%)	PA (%)	Score
Velvet/Oases	467816	0.18	82010	1792	99.30	99.46	9.43
Trinity	127202	0.26	33581	1161.5	100	99.62	9.19
IDBA	277106	0.07	18999	2144	100	78.65	8.64
rnaSPAdes	973409	0.48	464859	850.5	25.15	99.72	6.31

Table 2: Overview of *de novo* final assemblings of S288C's lncRNAs. The ICAP, AG, S+C, P, PA, and Score are described in equation 2. **Avg. len.:** the average sequence length of CAP3 re-assembling.

Most lncRNAs range from ~200-400 nts (Figure 2), being the lncRNA transcr_28768 of S288c the largest one (2,739 nts). Interestingly, all lncRNAs do not present similarities to those previously published (Parker *et al.* 2017, 2018).
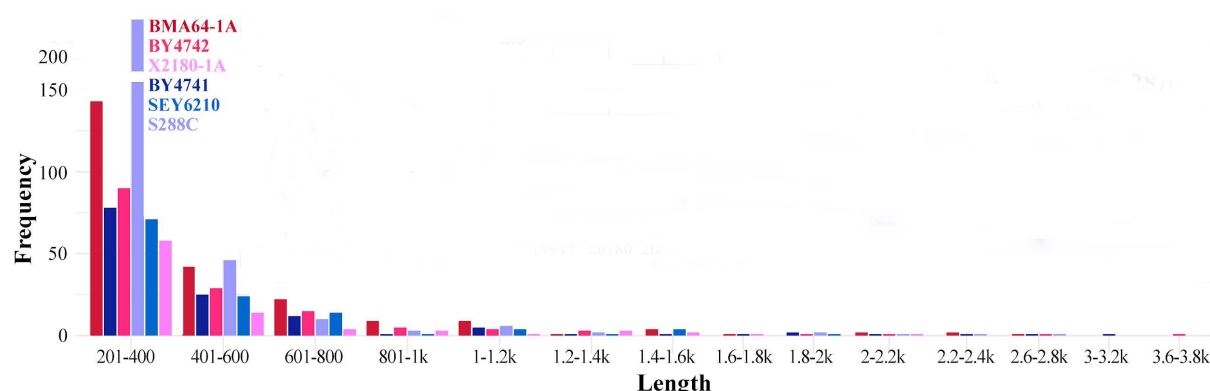
Figure 2: The distribution of lncRNA lengths.

The filtered LNCPI networks (probability ≥95%) harbor most of the identified lncRNAs. The filtered networks presented from 17 to 44 lncRNAs and from 86 to 394 target-proteins. The number of edges found is higher than the number of target-proteins, indicating that some lncRNAs interact with multiple proteins (Table 3). Indeed, the average of target proteins divided by the number of lncRNAs is ~8.1.

Strain	*Num. lncRNAs	**Num. target-proteins	Num. edges
BMA64-1A	36	159	244
BY4742	25	200	369
X2180-1A	17	86	153
BY4741	17	230	286
SEY6210	20	174	233
S288C	44	394	706

Table 3: Summary of LNCPI. ***:** the number of lncRNAs with at least one interaction; ****:** the number of proteins potentially binding to lncRNAs.

The EtOH stress-responsive lncRNAs lie in four functional categories. These categories were identified based on the function of target-proteins selected by lncRNA-propagation analysis from each differentially expressed lncRNAs. The 1st category ("essential process") includes genes related to transcription, replication, and ribosome biogenesis. The 2nd ("membrane dependent process") includes genes related to signaling and cell division, membranes and cell wall, and intracellular transport. The 3rd ("metabolic process") includes genes related to oxidative stress response, diauxic shift, fermentation, and other metabolisms. The 4th includes genes of the "degradation

11

process" (Figure 3A). Moreover, the whole set of target-proteins (without considering differential expression and strain-specificities) are mainly related to "RNA polymerase assembling" and "negative regulation of protein kinase activity by protein phosphorylation" (Figure 3B).
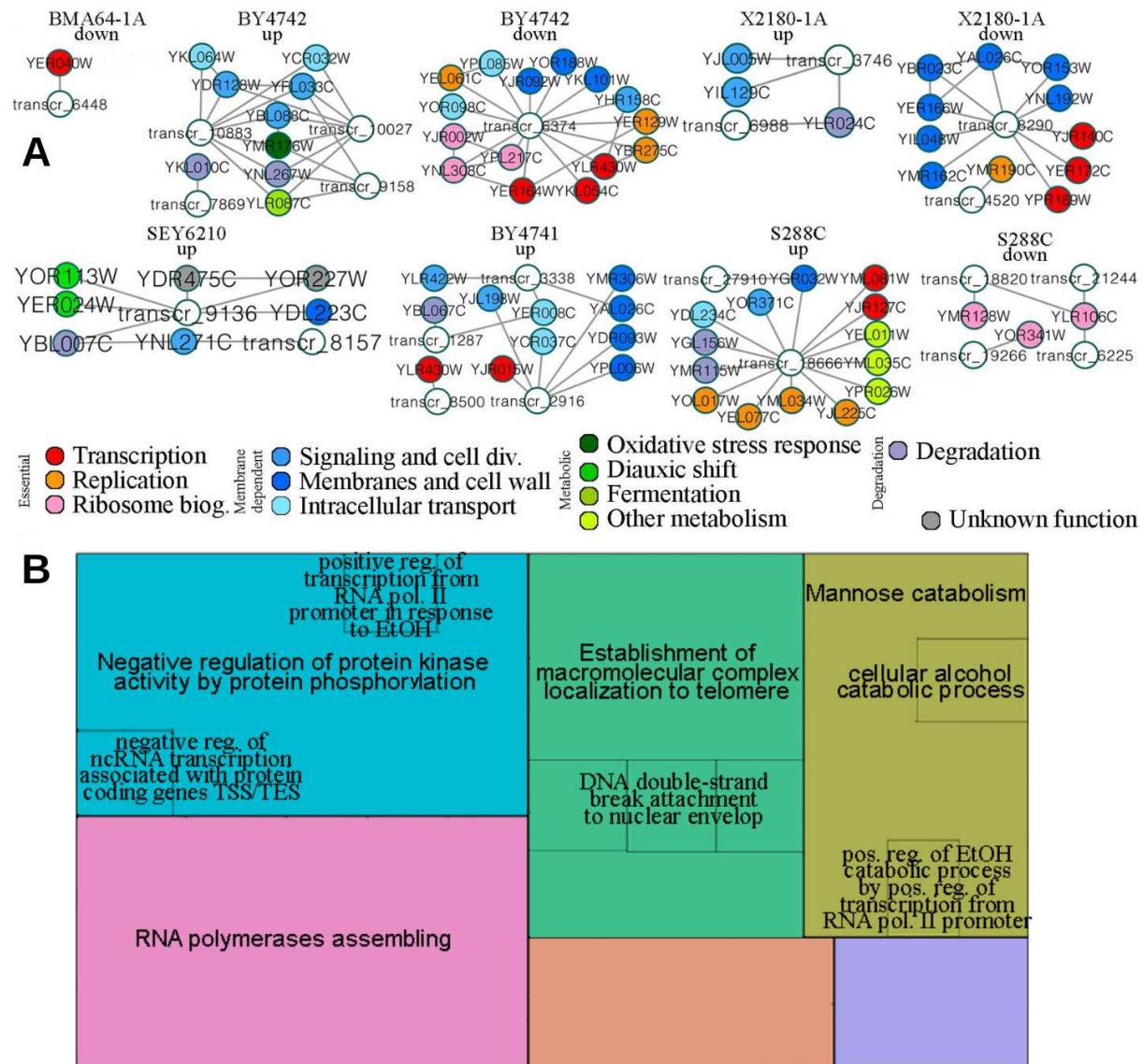

Figure 3: Functional annotation of lncRNAs **A:** selected subsystems using the lncRNA-propagation analysis. The node colors are related to biological functions depicted under the graphs; **B:** enriched terms of all target-proteins without considering neither strains nor differential expression and without excluding redundant proteins.

Discussion

EtOH stress-responsive lncRNAs are diverse and may be directly involved in the EtOH tolerance

Few lncRNAs have been found in yeasts, counting up to 18 lncRNAs for *S. cerevisiae* (Till *et al.* 2018). Here, we present the most extensive catalog of lncRNAs of 6 different *S. cerevisiae* strains; the ones were never identified before since there was no similarity to the previously lncRNAs found in yeast (Parker *et al.* 2017, 2018).

LncRNAs can act as protein-baits providing a negative impact on the target-protein by blocking its function. However, lncRNA-scaffolders help the protein complex organization positively or negatively (Wang and Chang 2011). The LNCPI here predicted fits the number of edges (ncRNA-protein interactions) previously reported (Panni *et al.* 2017), as well as the number of target proteins (~8.1 proteins per transcript); in fact, one small RNA (100 nts) can trap around 5-20 proteins (Chujo *et al.* 2016).

We used the guilt-by-association and the information flow throughout LNCPI (the lncRNA-propagation) to assess the lncRNA functions. Overall, the EtOH tolerance responsive lncRNAs in both higher tolerant (HT) and lower tolerant (LT) phenotypes work on different pathways; we highlight the cell wall, cell cycle and growth, cell longevity, cell surveillance, ribosome biogenesis, intracellular transport, trehalose metabolism, transcription, and nutrients shifts pathways. Below, we discuss some exciting interactions that may be related to EtOH tolerance.

The down-regulated lncRNA transcr_63478 of BY4742 binds to the Cin8p (YEL061C), responsible for the mitotic spindle assembly and chromosome segregation (Roof *et al.* 1992). The lack of CIN8 leads to cell cycle progression delay (Straight *et al.* 1998; Mittal *et al.* 2020). Remarkably, CIN8 is down-regulated in all HT strains and not differentially expressed in most LTs (data not shown). Additionally, this lncRNA also binds to Def1p, an RNAPII degradation factor (Woudstra *et al.* 2002).

The EtOH affects yeast's cell wall components (Aguilar-Uscanga and Francois 2003). The down-regulated lncRNA transcr_8290 of X2180-1A bind to the ATPases Dnf1p (YER166W), Neo1p (YIL048W), Dsr2p (YAL026C), Dnf3 (YMR162C), and chitin synthases Chs3p (YBR023C), and Chs1p (YNL192W). The 4 P-type ATPases mentioned transport phospholipids through a bilayer membrane (Paulusma and Oude Elferink 2005), and then, chitin synthases produce chitins for the cell wall (Ziman *et al.* 1996). Moreover, these proteins lie in the membrane, working as multi-drug transport

in the responsive drug system (Golin *et al.* 2007). Altogether, we suggest that transcr_8290 of X2180-1A may act as an assembler of protein complexes related to membrane and cell wall.

The down-regulated lncRNAs transcr_18820, transcr_21244, and transcr_6225 of S288C bind to the Ecm16p (YMR128W), and Rea1p/Mdn1p (YLR106C), which are small nucleolar ribonucleoprotein (snoRNP) and ribosome biogenesis protein, respectively (Shiratori *et al.* 1999; Colley *et al.* 2000; Nissan 2002). Additionally, transcr_19266 interacts with Rpa190p (YOR341W), which is part of RNA Polymerase I (RNAPI) (Kuhn *et al.* 2007).

The up-regulated lncRNAs transcr_3338 and transcr_2916 of BY4741 bind to the flippases Drs2p (YAL026C) and Dnf2p (YDR093W), respectively. These flipases are P-type ATPases that concentrate phosphatidylserine and phosphatidylethanolamine on the cytosolic leaflet, contributing to endocytosis, intracellular transport, and cell polarity (Chen *et al.* 1999; Hua *et al.* 2002; Pomorski *et al.* 2003; Iwamoto *et al.* 2004). Transcr_2916 also binds to the two low-affinity phosphate transporters Pho90p (YJL198W) and Pho87p (YCR037C); the overexpression of these genes leads to abnormal cell cycle progression and a reduction of vegetative growth rate (Stevenson *et al.* 2001; Sopko *et al.* 2006; Yoshikawa *et al.* 2011). BY4741 is the only strain with up-regulation of these PHO genes (data not shown).

The up-regulated lncRNAs of SEY6210 seem to act mainly on "membrane dependent processes" in the cell modeling processes (such as the ones during the cell cycle). In this case, the interaction between transcr_8157 and transcr_9136 with the proteins Bni1p (YNL271C), Sla1p (YBL007C), and Hbt1p (YDL223C) indicate that these lncRNAs may work on the cell cortex modeling system, albeit the mentioned genes have different roles. Sla1p is a cytoskeletal binding protein associated with endocytosis or binds to proteins to regulate actin dynamics (Pruyne and Bretscher 2000; Howard *et al.* 2002). Bni1p and Hbt1p are responsible for cell polarization (Lee *et al.* 1999; Dittmar 2002; Pruyne *et al.* 2002, 2004; Tcheperegine *et al.* 2005; Guarente 2010).

The up-regulated lncRNAs transcr_18666 of S288C seems to work on trehalose metabolism actively. There is a positive correlation between cell viability and trehalose concentration in cells under EtOH stress, providing a protective effect to the EtOH (Mansure *et al.* 1994). The Ath1p (YPR026W), a trehalose degradation protein (Nwaka *et al.* 1996; Jules *et al.* 2004), interacts with transcr_18666. Interestingly,

S288C has the lowest EtOH tolerance. Altogether, our initial hypothesis is that the transcr_18666 may negatively impact ATH1 under the severe EtOH stress.

The X2180-1A up-regulated lncRNA transcr_3746 may be acting on cellular surveillance mediated by a "membrane dependent process". We found that the Cyr1p (YJL005W) (works on signal transduction, which is required for cAMP production (Kataoka *et al.* 1985)) also binds to transcr_3746. The cAMP is involved in cell cycle progression, sporulation, cell growth, stress response, and longevity (Casperson *et al.* 1985). Hence, we suggest that the transcr_3746 may also be related to cell longevity, growth, and proliferation.

The up-regulated lncRNAs transcr_10883, transcr_10027, and transcr_9158 of BY4742 may be directly involved in the EtOH tolerance controlling nutrient supply. These lncRNAs bind to Pik1p (YNL267W), a kinase that can rapidly restore the nutrients supply in cells under nutrient deprivation (Demmel *et al.* 2008). Interestingly, cells under EtOH stress ongoing by a variation on nutrients depletion (Tesnière *et al.* 2013).

The down-regulated lncRNA transcr_6448 of BMA64-1A had one target-protein coded by the gene YER040W (GLN3). Gln3p is a transcriptional activator of genes subject to nitrogen catabolite repression (Feller *et al.* 2013). The deletion of GLN3 boosts the yeast to the branched-chain alcohols tolerance (Kuroda *et al.* 2019). The finding mentioned fits our hypothesis since BMA64-1A is the only strain with the down-regulation of GLN3 (data not shown) and presents the highest EtOH tolerance here analyzed. In this case, we suggest that the transcr_6448 may act as a sort of Gln3p repressor under the severe EtOH stress by an unknown negative-feedback loop. Altogether, we hypothesize that the highest EtOH tolerance observed for BMA64-1A may be a by-product of the negative GLN3 regulation by the transcr_6448 under stress.

Plenty of lncRNA's target-proteins are related to transcriptional mechanisms ("RNA polymerase assembling" and "negative regulation of protein kinase activity by protein phosphorylation"). The proteins related to the "RNA polymerase assembling" mechanism are related to the "transcription by RNA polymerase II". LncRNAs can contribute to a balance of transcription/degradation rate (Timmers and Tora 2018). Therefore, we hypothesize that lncRNA-RNAPII interactions might be acting as signaling molecules to counterbalance the transcription/degradation rate of mRNAs.

Altogether, network analysis provided hints about the role of many EtOH stress-

responsive lncRNA. These ncRNAs seem to act as baits, backbones, or adapter molecules to maintain protein complexes essentials to improve yeast capacity to endure the severe EtOH stress. Finally, our findings indicate that yeasts' lncRNAs under a severe EtOH stress seem to interconnect a diversity of modules to surpass hurdles imposed by this stressor.

References

Aguilar-Uscanga B., and J. M. Francois, 2003 A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. Lett. Appl. Microbiol. 37: 268–274. https://doi.org/10.1046/j.1472-765X.2003.01394.x

Alper H., J. Moxley, E. Nevoigt, G. R. Fink, and G. Stephanopoulos, 2006 Engineering yeast transcription machinery for improved ethanol tolerance and production. Science 314: 1565–1568. https://doi.org/10.1126/science.1131969

Anderson D. M., K. M. Anderson, C.-L. Chang, C. A. Makarewich, B. R. Nelson, *et al.*, 2015 A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance. Cell 160: 595–606. https://doi.org/10.1016/j.cell.2015.01.009

Arrial R. T., R. C. Togawa, and M. de M. Brigido, 2009 Screening non-coding RNAs in transcriptomes from neglected species using PORTRAIT: case study of the pathogenic fungus Paracoccidioides brasiliensis. BMC Bioinformatics 10: 239. https://doi.org/10.1186/1471-2105-10-239

Auesukaree C., 2017 Molecular mechanisms of the yeast adaptive response and tolerance to stresses encountered during ethanol fermentation. J. Biosci. Bioeng. 124: 133–142. https://doi.org/10.1016/j.jbiosc.2017.03.009

Balarezo-Cisneros L. N., S. Parker, M. G. Fraczek, S. Timouma, P. Wang, *et al.*, 2020 Functional and transcriptional profiling of non-coding RNAs in yeast reveal context-dependent phenotypes and widespread in trans effects on the protein regulatory network. bioRxiv 1–26. https://doi.org/10.1101/2020.04.07.029611

Bankevich A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, *et al.*, 2012 SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J. Comput. Biol. 19: 455–477. https://doi.org/10.1089/cmb.2012.0021

Blohm P., G. Frishman, P. Smialowski, F. Goebels, B. Wachinger, *et al.*, 2014 Negatome 2.0: a database of non-interacting proteins derived by literature mining, manual annotation and protein structure analysis. Nucleic Acids Res. 42: D396–D400. https://doi.org/10.1093/nar/gkt1079

Bolger A. M., M. Lohse, and B. Usadel, 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114–20. https://doi.org/10.1093/bioinformatics/btu170

Burge S. W., J. Daub, R. Eberhardt, J. Tate, L. Barquist, *et al.*, 2013 Rfam 11.0: 10 years of RNA families. Nucleic Acids Res. 41: D226-32. https://doi.org/10.1093/nar/gks1005

Cantarel B. L., I. Korf, S. M. C. Robb, G. Parra, E. Ross, *et al.*, 2008 MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18: 188–196. https://doi.org/10.1101/gr.6743907

Casperson G. F., N. Walker, and H. R. Bourne, 1985 Isolation of the gene encoding adenylate cyclase in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. 82: 5060–5063. https://doi.org/10.1073/pnas.82.15.5060

Chakraborty S., V. Aggarwal, D. Mukherjee, and K. Andras, 2012 Biomass to biofuel: a review on production technology. Asia-Pacific J. Chem. Eng. 7: S254–S262. https://doi.org/10.1002/apj.1642

Chatr-Aryamontri A., B. J. Breitkreutz, S. Heinicke, L. Boucher, A. Winter, *et al.*, 2013 The BioGRID interaction database: 2013 Update. Nucleic Acids Res. 41: 470–478. https://doi.org/10.1093/nar/gks1158

Chen C. Y., M. F. Ingram, P. H. Rosal, and T. R. Graham, 1999 Role for Drs2p, a P-type ATPase and potential aminophospholipid translocase, in yeast late Golgi function. J. Cell Biol. 147: 1223–36. https://doi.org/10.1083/jcb.147.6.1223

Chujo T., T. Yamazaki, and T. Hirose, 2016 Architectural RNAs (arcRNAs): A class of long noncoding RNAs that function as the scaffold of nuclear bodies. Biochim. Biophys. Acta - Gene Regul. Mech. 1859: 139–146. https://doi.org/10.1016/j.bbagrm.2015.05.007

Colley A., J. D. Beggs, D. Tollervey, and D. L. J. Lafontaine, 2000 Dhr1p, a Putative DEAH-Box RNA Helicase, Is Associated with the Box C+D snoRNP U3. Mol. Cell. Biol. 20: 7238–7246. https://doi.org/10.1128/mcb.20.19.7238-7246.2000

Cowen L., T. Ideker, B. J. Raphael, and R. Sharan, 2017 Network propagation: A universal amplifier of genetic associations. Nat. Rev. Genet. 18: 551–562. https://doi.org/10.1038/nrg.2017.38

Demeke M. M., H. Dietz, Y. Li, M. R. Foulquié-Moreno, S. Mutturi, *et al.*, 2013 Development of a D-xylose fermenting and inhibitor tolerant industrial Saccharomyces cerevisiae strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. Biotechnol. Biofuels 6: 89. https://doi.org/10.1186/1754-6834-6-89

Demirbas A., 2017 Tomorrow's biofuels: Goals and hopes. Energy Sources, Part A

Recover. Util. Environ. Eff. 39: 673–679. https://doi.org/10.1080/15567036.2016.1252815

Demmel L., M. Beck, C. Klose, A.-L. Schlaitz, Y. Gloor, *et al.*, 2008 Nucleocytoplasmic Shuttling of the Golgi Phosphatidylinositol 4-Kinase Pik1 Is Regulated by 14-3-3 Proteins and Coordinates Golgi Function with Cell Growth, (B. Glick, Ed.). Mol. Biol. Cell 19: 1046–1061. https://doi.org/10.1091/mbc.e07-02-0134

Ding J., X. Huang, L. Zhang, N. Zhao, D. Yang, *et al.*, 2009 Tolerance and stress response to ethanol in the yeast Saccharomyces cerevisiae. Appl. Microbiol. Biotechnol. 85: 253–263.

Dittmar G. A. G., 2002 Role of a Ubiquitin-Like Modification in Polarized Morphogenesis. Science (80-.). 295: 2442–2446. https://doi.org/10.1126/science.1069989

Fabregat A., S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, *et al.*, 2014 The Reactome Pathway Knowledgebase. Nucleic Acids Res. 44: D481–D487. https://doi.org/10.1093/nar/gkt1102

Feller A., I. Georis, J. J. Tate, T. G. Cooper, and E. Dubois, 2013 Alterations in the Ure2 αCap domain elicit different GATA factor responses to rapamycin treatment and nitrogen limitation. J. Biol. Chem. 288: 1841–55. https://doi.org/10.1074/jbc.M112.385054

Ferrè F., A. Colantoni, and M. Helmer-Citterich, 2016 Revealing protein-lncRNA interaction. Brief. Bioinform. 17: 106–16. https://doi.org/10.1093/bib/bbv031

Finn R. D., A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, *et al.*, 2014 Pfam: The protein families database. Nucleic Acids Res. 42: 222–230. https://doi.org/10.1093/nar/gkt1223

Geisler S., and J. Coller, 2013 RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat. Rev. Mol. Cell Biol. 14: 699–712. https://doi.org/10.1038/nrm3679

Golin J., S. V. Ambudkar, and L. May, 2007 The yeast Pdr5p multidrug transporter: How does it recognize so many substrates? Biochem. Biophys. Res. Commun. 356: 1–5. https://doi.org/10.1016/j.bbrc.2007.02.011

Guarente L., 2010 Forever young. Cell 140: 176–8. https://doi.org/10.1016/j.cell.2010.01.015

Gupta A., and J. P. Verma, 2015 Sustainable bio-ethanol production from agro-

residues: A review. Renew. Sustain. Energy Rev. 41: 550–567.
https://doi.org/10.1016/j.rser.2014.08.032

Gurevich A., V. Saveliev, N. Vyahhi, and G. Tesler, 2013 QUAST: quality
assessment tool for genome assemblies. Bioinformatics 29: 1072–1075.
https://doi.org/10.1093/bioinformatics/btt086

Haas B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, *et al.*, 2013 De
novo transcript sequence reconstruction from RNA-seq using the Trinity platform
for reference generation and analysis. Nat. Protoc. 8: 1494–1512.
https://doi.org/10.1038/nprot.2013.084

He B., S. Zhao, Y. Chen, Q. Cao, C. Wei, *et al.*, 2015 Optimal assembly strategies of
transcriptome related to ploidies of eukaryotic organisms. BMC Genomics 16:
65. https://doi.org/10.1186/s12864-014-1192-7

Howard J. P., J. L. Hutton, J. M. Olson, and G. S. Payne, 2002 Sla1p serves as the
targeting signal recognition factor for NPFX(1,2)D-mediated endocytosis. J. Cell
Biol. 157: 315–26. https://doi.org/10.1083/jcb.200110027

Hua Z., P. Fatheddin, and T. R. Graham, 2002 An essential subfamily of Drs2p-
related P-type ATPases is required for protein trafficking between Golgi complex
and endosomal/vacuolar system., (R. Schekman, Ed.). Mol. Biol. Cell 13: 3162–
3177. https://doi.org/10.1091/mbc.e02-03-0172

Huang X., 1999 CAP3: A DNA Sequence Assembly Program. Genome Res. 9: 868–
877. https://doi.org/10.1101/gr.9.9.868

Iwamoto K., S. Kobayashi, R. Fukuda, M. Umeda, T. Kobayashi, *et al.*, 2004 Local
exposure of phosphatidylethanolamine on the yeast plasma membrane is
implicated in cell polarity. Genes Cells 9: 891–903.
https://doi.org/10.1111/j.1365-2443.2004.00782.x

Jackman S. D., B. P. Vandervalk, H. Mohamadi, J. Chu, S. Yeo, *et al.*, 2017 ABySS
2.0: resource-efficient assembly of large genomes using a Bloom filter. Genome
Res. 27: 768–777. https://doi.org/10.1101/gr.214346.116

Jules M., V. Guillou, J. François, and J.-L. Parrou, 2004 Two distinct pathways for
trehalose assimilation in the yeast Saccharomyces cerevisiae. Appl. Environ.
Microbiol. 70: 2771–8. https://doi.org/10.1128/aem.70.5.2771-2778.2004

Kataoka T., D. Broek, and M. Wigler, 1985 DNA sequence and characterization of
the S. cerevisiae gene encoding adenylate cyclase. Cell 43: 493–505.
https://doi.org/10.1016/0092-8674(85)90179-5

Kim D., B. Langmead, and S. L. Salzberg, 2015 HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12: 357–360. https://doi.org/10.1038/nmeth.3317

Kong L., Y. Zhang, Z.-Q. Ye, X.-Q. Liu, S.-Q. Zhao, *et al.*, 2007 CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res. 35: W345–W349. https://doi.org/10.1093/nar/gkm391

Kuhn C.-D., S. R. Geiger, S. Baumli, M. Gartmann, J. Gerber, *et al.*, 2007 Functional Architecture of RNA Polymerase I. Cell 131: 1260–1272. https://doi.org/10.1016/j.cell.2007.10.051

Kuroda K., S. K. Hammer, Y. Watanabe, J. Montaño López, G. R. Fink, *et al.*, 2019 Critical Roles of the Pentose Phosphate Pathway and GLN3 in Isobutanol-Specific Tolerance in Yeast. Cell Syst. 9: 534-547.e5. https://doi.org/10.1016/j.cels.2019.10.006

Langmead B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nat. Methods 9: 357–9. https://doi.org/10.1038/nmeth.1923

Lee L., S. K. Klee, M. Evangelista, C. Boone, and D. Pellman, 1999 Control of mitotic spindle position by the Saccharomyces cerevisiae formin Bni1p. J. Cell Biol. 144: 947–61. https://doi.org/10.1083/jcb.144.5.947

Lewis J. a., I. M. Elkon, M. a. McGee, A. J. Higbee, and A. P. Gasch, 2010 Exploiting natural variation in Saccharomyces cerevisiae to identify genes for increased ethanol resistance. Genetics 186: 1197–1205. https://doi.org/10.1534/genetics.110.121871

Li W., and A. Godzik, 2006 Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658–9. https://doi.org/10.1093/bioinformatics/btl158

Li Z., W. Zhao, M. Wang, and X. Zhou, 2019 The Role of Long Noncoding RNAs in Gene Expression Regulation, in *Gene Expression Profiling in Cancer*, IntechOpen.

Licata L., L. Briganti, D. Peluso, L. Perfetto, M. Iannuccelli, *et al.*, 2012 MINT, the molecular interaction database: 2012 update. Nucleic Acids Res. 40: D857–D861. https://doi.org/10.1093/nar/gkr930

Love M. I., W. Huber, and S. Anders, 2014 *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*.

Lu Q., S. Ren, M. Lu, Y. Zhang, D. Zhu, *et al.*, 2013 Computational prediction of associations between long non-coding RNAs and proteins. BMC Genomics 14: 651. https://doi.org/10.1186/1471-2164-14-651

Ma M., and Z. L. Liu, 2010 Mechanisms of ethanol tolerance in saccharomyces cerevisiae. Appl. Microbiol. Biotechnol. 87: 829–845.

Mansure J. J ., A. D. Panek, L. M. Crowe, and J. H. Crowe, 1994 Trehalose inhibits ethanol effects on intact yeast cells and liposomes. Biochim. Biophys. Acta - Biomembr. 1191: 309–316. https://doi.org/10.1016/0005-2736(94)90181-3

Mistry J., R. D. Finn, S. R. Eddy, A. Bateman, and M. Punta, 2013 Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41. https://doi.org/10.1093/nar/gkt263

Mittal P., K. Ghule, D. Trakroo, H. K. Prajapati, and S. K. Ghosh, 2020 Meiosis-Specific Functions of Kinesin Motors in Cohesin Removal and Maintenance of Chromosome Integrity in Budding Yeast. Mol. Cell. Biol. 40. https://doi.org/10.1128/MCB.00386-19

Mussatto S. I., G. Dragone, P. M. R. Guimarães, J. P. a Silva, L. M. Carneiro, *et al.*, 2010 Technological trends, global market, and challenges of bio-ethanol production. Biotechnol. Adv. 28: 817–830.

Navarro-Tapia E., R. K. Nana, A. Querol, and R. Pérez-Torrado, 2016 Ethanol Cellular Defense Induce Unfolded Protein Response in Yeast. Front. Microbiol. 7: 1–12. https://doi.org/10.3389/fmicb.2016.00189

Nissan T. A., 2002 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. EMBO J. 21: 5539–5547. https://doi.org/10.1093/emboj/cdf547

Nwaka S., B. Mechler, and H. Holzer, 1996 Deletion of the ATH1 gene in Saccharomyces cerevisiae prevents growth on trehalose. FEBS Lett. 386: 235–8. https://doi.org/10.1016/0014-5793(96)00450-4

Panni S., A. Prakash, A. Bateman, and S. Orchard, 2017 The yeast noncoding RNA interaction network. RNA 23: 1479–1492. https://doi.org/10.1261/rna.060996.117

Parker S., M. G. Fraczek, J. Wu, S. Shamsah, A. Manousaki, *et al.*, 2017 A resource for functional profiling of noncoding RNA in the yeast Saccharomyces cerevisiae. RNA 23: 1166–1171. https://doi.org/10.1261/rna.061564.117

Parker S., M. G. Fraczek, J. Wu, S. Shamsah, A. Manousaki, *et al.*, 2018 Large-

scale profiling of noncoding RNA function in yeast, (M. Snyder, Ed.). PLOS Genet. 14: e1007253. https://doi.org/10.1371/journal.pgen.1007253

Paulusma C. C., and R. P. J. Oude Elferink, 2005 The type 4 subfamily of P-type ATPases, putative aminophospholipid translocases with a role in human disease. Biochim. Biophys. Acta - Mol. Basis Dis. 1741: 11–24. https://doi.org/10.1016/j.bbadis.2005.04.006

Peng Y., H. C. M. Leung, S. M. Yiu, and F. Y. L. Chin, 2010 IDBA – A Practical Iterative de Bruijn Graph De Novo Assembler, pp. 426–440 in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*,.

Peng Y., H. C. M. Leung, S. M. Yiu, and F. Y. L. Chin, 2012 IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28: 1420–1428. https://doi.org/10.1093/bioinformatics/bts174

Pomorski T., R. Lombardi, H. Riezman, P. F. Devaux, G. van Meer, *et al.*, 2003 Drs2p-related P-type ATPases Dnf1p and Dnf2p are required for phospholipid translocation across the yeast plasma membrane and serve a role in endocytosis. Mol. Biol. Cell 14: 1240–54. https://doi.org/10.1091/mbc.e02-08-0501

Pruyne D., and A. Bretscher, 2000 Polarization of cell growth in yeast. J. Cell Sci. 113 (Pt 4: 571–85.

Pruyne D., M. Evangelista, C. Yang, E. Bi, S. Zigmond, *et al.*, 2002 Role of formins in actin assembly: nucleation and barbed-end association. Science 297: 612–5. https://doi.org/10.1126/science.1072309

Pruyne D., L. Gao, E. Bi, and A. Bretscher, 2004 Stable and dynamic axes of polarity use distinct formin isoforms in budding yeast. Mol. Biol. Cell 15: 4971–89. https://doi.org/10.1091/mbc.e04-04-0296

Quinlan A. R., and I. M. Hall, 2010 BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics 26: 841–842. https://doi.org/10.1093/bioinformatics/btq033

Reimand J., T. Arak, P. Adler, L. Kolberg, S. Reisberg, *et al.*, 2016 g:Profiler—a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 44: W83–W89. https://doi.org/10.1093/nar/gkw199

Rice P., I. Longden, and A. Bleasby, 2000 EMBOSS: The European Molecular

Biology Open Software Suite. Trends Genet. 16: 276–277.
https://doi.org/10.1016/S0168-9525(00)02024-2

Roof D. M., P. B. Meluh, and M. D. Rose, 1992 Kinesin-related proteins required for assembly of the mitotic spindle. J. Cell Biol. 118: 95–108.
https://doi.org/10.1083/jcb.118.1.95

Schulz M. H., D. R. Zerbino, M. Vingron, and E. Birney, 2012 Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels.
Bioinformatics 28: 1086–1092. https://doi.org/10.1093/bioinformatics/bts094

Shiratori A., T. Shibata, M. Arisawa, F. Hanaoka, Y. Marakami, *et al.*, 1999
Systematic identification, classification, and characterization of the open reading frames which encode novel helicase-related proteins inSaccharomyces cerevisiae by gene disruption and Northern analysis. Yeast 15: 219–253.
https://doi.org/10.1002/(SICI)1097-0061(199902)15:3<219::AID-YEA349>3.0.CO;2-3

Song G., B. J. A. Dickins, J. Demeter, S. Engel, B. Dunn, *et al.*, 2015 AGAPE (Automated Genome Analysis PipelinE) for Pan-Genome Analysis of Saccharomyces cerevisiae, (J. Schacherer, Ed.). PLoS One 10: e0120671.
https://doi.org/10.1371/journal.pone.0120671

Sopko R., D. Huang, N. Preston, G. Chua, B. Papp, *et al.*, 2006 Mapping pathways and phenotypes by systematic gene overexpression. Mol. Cell 21: 319–30.
https://doi.org/10.1016/j.molcel.2005.12.011

Stanley D., A. Bandara, S. Fraser, P. J. Chambers, and G. a. Stanley, 2010 The ethanol stress response and ethanol tolerance of Saccharomyces cerevisiae. J. Appl. Microbiol. 109: 13–24. https://doi.org/10.1111/j.1365-2672.2009.04657.x

Stevenson L. F., B. K. Kennedy, and E. Harlow, 2001 A large-scale overexpression screen in Saccharomyces cerevisiae identifies previously uncharacterized cell cycle genes. Proc. Natl. Acad. Sci. U. S. A. 98: 3946–51.
https://doi.org/10.1073/pnas.051013498

Straight A. F., J. W. Sedat, and A. W. Murray, 1998 Time-lapse microscopy reveals unique roles for kinesins during anaphase in budding yeast. J. Cell Biol. 143: 687–94. https://doi.org/10.1083/jcb.143.3.687

Supek F., M. Bošnjak, N. Škunca, and T. Šmuc, 2011 REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms, (C. Gibas, Ed.). PLoS One 6: e21800. https://doi.org/10.1371/journal.pone.0021800

Swinnen S., K. Schaerlaekens, T. Pais, J. Claesen, G. Hubmann, *et al.*, 2012 Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. Genome Res. 22: 975–984. https://doi.org/10.1101/gr.131698.111

Tcheperegine S. E., X.-D. Gao, and E. Bi, 2005 Regulation of cell polarity by interactions of Msb3 and Msb4 with Cdc42 and polarisome components. Mol. Cell. Biol. 25: 8567–80. https://doi.org/10.1128/MCB.25.19.8567-8580.2005

Teixeira M. C., L. R. Raposo, N. P. Mira, A. B. Lourenco, I. Sa-Correia, *et al.*, 2009 Genome-Wide Identification of Saccharomyces cerevisiae Genes Required for Maximal Tolerance to Ethanol. Appl. Environ. Microbiol. 75: 5761–5772. https://doi.org/10.1128/AEM.00845-09

Tesnière C., P. Delobel, M. Pradal, and B. Blondin, 2013 Impact of Nutrient Imbalance on Wine Alcoholic Fermentations: Nitrogen Excess Enhances Yeast Cell Death in Lipid-Limited Must, (C. Menezes, Ed.). PLoS One 8: e61645. https://doi.org/10.1371/journal.pone.0061645

The MIT License, Pairfq-Research Computing Center Wiki

Till P., R. L. Mach, and A. R. Mach-Aigner, 2018 A current view on long noncoding RNAs in yeast and filamentous fungi. Appl. Microbiol. Biotechnol. 102: 7319–7331. https://doi.org/10.1007/s00253-018-9187-y

Timmers H. T. M., and L. Tora, 2018 Transcript Buffering: A Balancing Act between mRNA Synthesis and mRNA Degradation. Mol. Cell 72: 10–17. https://doi.org/10.1016/j.molcel.2018.08.023

Tripathi V., J. D. Ellis, Z. Shen, D. Y. Song, Q. Pan, *et al.*, 2010 The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation. Mol. Cell 39: 925–938. https://doi.org/10.1016/j.molcel.2010.08.011

Wang K. C., and H. Y. Chang, 2011 Molecular Mechanisms of Long Noncoding RNAs. Mol. Cell 43: 904–914. https://doi.org/10.1016/j.molcel.2011.08.018

Woudstra E. C., C. Gilbert, J. Fellows, L. Jansen, J. Brouwer, *et al.*, 2002 A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. Nature 415: 929–33. https://doi.org/10.1038/415929a

Wu T. D., and C. K. Watanabe, 2005 GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21: 1859–1875. https://doi.org/10.1093/bioinformatics/bti310

Yamashita A., Y. Shichino, and M. Yamamoto, 2016 The long non-coding RNA world in yeasts. Biochim. Biophys. Acta - Gene Regul. Mech. 1859: 147–154. https://doi.org/10.1016/j.bbagrm.2015.08.003

Yang T.-H., C.-C. Wang, Y.-C. Wang, and W.-S. Wu, 2014 YTRP: a repository for yeast transcriptional regulatory pathways. Database (Oxford). 2014: bau014. https://doi.org/10.1093/database/bau014

Yoshikawa K., T. Tanaka, Y. Ida, C. Furusawa, T. Hirasawa, *et al.*, 2011 Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae. Yeast 28: 349–61. https://doi.org/10.1002/yea.1843

Zerbino D. R., and E. Birney, 2008 Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18: 821–829. https://doi.org/10.1101/gr.074492.107

Ziman M., J. S. Chuang, and R. W. Schekman, 1996 Chs1p and Chs3p, two proteins involved in chitin synthesis, populate a compartment of the Saccharomyces cerevisiae endocytic pathway. Mol. Biol. Cell 7: 1909–1919. https://doi.org/10.1091/mbc.7.12.1909

Financial Support