Lawrence Berkeley National Laboratory
LBL Publications

Title

A transcriptomic atlas of acute stress response to low pH in multiple Issatchenkia orientalis strains.

Permalink

Journal

Authors

Dubinkina, Veronika
Bhogale, Shounak
Hsieh, Ping-Hung
et al.

Publication Date

DOI

A transcriptomic atlas of acute stress response to low pH in multiple *Issatchenkia orientalis* strains

Veronika Dubinkina,[1,2,3] Shounak Bhogale,[4] Ping-Hung Hsieh,[5,6] Payam Dibaeinia,[1,7] Ananthan Nambiar,[1,2] Sergei Maslov,[1,2,8] Yasuo Yoshikuni,[5,6,9,10,11,12] Saurabh Sinha[1,4,7,13,14,15]

AUTHOR AFFILIATIONS See affiliation list on p. 19.

ABSTRACT Because of its natural stress tolerance to low pH, *Issatchenkia orientalis (a.k.a. Pichia kudriavzevii)* is a promising non-model yeast for bio-based production of organic acids. Yet, this organism is relatively unstudied, and specific mechanisms of its tolerance to low pH are poorly understood, limiting commercial use. In this study, we selected 12 *I. orientalis* strains with varying acid stress tolerance (six tolerant and six susceptible) and profiled their transcriptomes in different pH conditions to study potential mechanisms of pH tolerance in this species. We identified hundreds of genes whose expression response is shared by tolerant strains but not by susceptible strains, or vice versa, as well as genes whose responses are reversed between tolerant and susceptible strains. We mapped regulatory mechanisms of transcriptomic responses via motif analysis as well as differential network reconstruction, identifying several transcription factors, including Stb5, Mac1, and Rtg1/Rtg3, some of which are known for their roles in acid response in *Saccharomyces cerevisiae*. Functional genomics analysis of short-listed genes and transcription factors suggested significant roles for energy metabolism and translation-related processes, as well as the cell wall integrity pathway and RTG-dependent retrograde signaling pathway. Finally, we conducted additional experiments for two organic acids, 3-hydroxypropionate and citramalate, to eliminate acid-specific effects and found potential roles for glycolysis and trehalose biosynthesis specifically for response to low pH. In summary, our approach of comparative transcriptomics and phenotypic contrasting, along with a multi-pronged bioinformatics analysis, suggests specific mechanisms of tolerance to low pH in *I. orientalis* that merit further validation through experimental perturbation and engineering.

IMPORTANCE *Issatchenkia orientalis* is a promising industrial chassis to produce biofuels and bioproducts due to its high tolerance to multiple environmental stresses such as low pH, heat, and other chemicals otherwise toxic for the most widely used microbes. Yet, little is known about specific mechanisms of such tolerance in this organism, hindering our ability to engineer this species to produce valuable biochemicals. Here, we report a comprehensive study of the mechanisms of acidic tolerance in this species via transcriptome profiling across variable pH for 12 different strains with different phenotypes. We found multiple regulatory mechanisms involved in tolerance to low pH in different strains of *I. orientalis*, marking potential targets for future gene editing and perturbation experiments.

KEYWORDS *Issatchenkia orientalis*, low pH, acidic stress, gene regulation

Biofuels and bioproducts made from sustainable biomass are becoming increasingly popular as a way to reduce greenhouse gas emissions (1). While current industrial manufacturing methods primarily rely on the use of sugar and starch from food crops, which could threaten food security (2), producing biofuels and bioproducts from

renewable and low-cost lignocellulosic biomass can be a more attractive choice (2–4). However, to build such production pipelines, we have to overcome several critical challenges, such as mitigating low pH stress for the microorganisms used during the bioproduction processes.

In the current pipeline of lignocellulosic biomass-based biofuels and bioproducts production, low pH stress arises due to several reasons. First, acid pretreatment is a popular approach for breaking down lignocellulosic cell wall structures to release sugar fragments for downstream enzymatic hydrolysis and fermentation (5, 6). The acid-pretreated lignocellulosic hydrolysates might theoretically be utilized directly as a carbon source for fermentation, but the low pH limits cell activity, resulting in a lower fermentation yield (7). In practice, a cost-added neutralizing step is required before the downstream enzymatic hydrolysis and fermentation. However, salts formed during the neutralizing step can also have negative impacts on microbes, resulting in reduced fermentation yield (8).

Therefore, adopting microorganisms tolerant to low pH for lignocellulosic biomass-based biofuels and bioproducts production would be a more cost-effective and sustainable direction, bypassing the use of neutralization agents and the further removal of side-product salts. Having such microbes as a biofactory is even more crucial for the efficient production of lignocellulosic biomass-based organic acids, which are among the top value-added bioproducts and have long been recognized as essential building blocks in the chemical industry (9, 10). In this case, in addition to acid pretreatment, production and accumulation of organic acids throughout the fermentation process can decrease pH even further.

Compared to the model bacterial *Escherichia coli* system, yeast species possess many desirable industrial traits for bioproducts, such as lower growth temperature, tolerance to relatively low pH, and cellular compartments for sequestering toxic products (11). Furthermore, the adoption of yeasts tolerant to low pH in large-scale industrial fermentation can help reduce bacterial contamination and minimize the cost of maintaining sterility throughout fermentation by maintaining a lower pH condition or acid washing (12). In addition to the widely used *Saccharomyces cerevisiae*, nonconventional yeasts could be an excellent chassis for bioproducts due to their greater stress tolerance to inhibitors and product toxicity (13). A prominent example is *Issatchenkia orientalis* (previously known as *Pichia kudriavzevii* or *Candida krusei*) (14), which has a high potential for industrial-scale organic acid production due to its ability to tolerate various lignocellulosic inhibitors (15), low pH, high concentrations of organic acids (16–19), salts (20), and temperature as high as 42°C (20–22).

While many studies report the remarkable tolerance of *I. orientalis* to low pH conditions and successfully engineer it to produce various acidic compounds, we still lack a mechanistic understanding of how such tolerance is achieved in this species. Several mechanisms and pathways have been reported as involved in tolerance to low pH, in the context of other microorganisms. These include reducing proton influx by altering cell membrane compositions (23), reducing the sizes of membrane channels (24), generating a chemiosmotic gradient (25), expressing more proton efflux pumps (26, 27), consuming protons via metabolism (28, 29), and producing basic compounds to prevent intracellular protons from accumulation (30, 31). Also, chaperone proteins and DNA repair systems are triggered to safeguard proteins, DNAs, and organelles if intracellular acidic stress persists (32, 33), and metabolic pathways are rewired to cope with the inhibitory effects induced by low pH such as increased glycolytic rates and catabolism of amino acids and RNAs (34).

Knowing which, if any, of the above mechanisms might be involved in the *I. orientalis* tolerance to low pH or whether there are any new ones, will provide insights into how to further improve these strains or other fungal species for more economical and sustainable organic acid production. In a recent study, it was noted that there is a remarkable variability in growth phenotype among different *I. orientalis* strains in extremely acidic environments, which cannot be explained by phylogeny alone (35).

Motivated by this finding, we performed here a systematic comparative transcriptome analysis of a specially selected group of pH-tolerant and pH-susceptible strains of *I. orientalis* in a variety of low pH conditions (including extremely acidic with pH 1.5) in search for molecular mechanisms of tolerance. We compiled sets of differentially expressed genes and functional categories separating tolerant and susceptible strains from each other. Furthermore, we identified a small set of transcription factors (TFs) and their binding motifs that are likely to control the expression of these genes. Our study provides insights into biological processes involved in the *I. orientalis* tolerance to low pH and suggests ways to manipulate it to our advantage by perturbing transcriptional regulatory networks (TRN).

MATERIALS AND METHODS

Growth measurements of the selected strains at pH 1.5

We first streaked the glycerol stock of the selected strains (Fig. 1a) on the yeast peptone dextrose (YPD) plate for growing overnight at 30℃. A single colony was picked up from the plate and inoculated into 2 mL yeast nitrogen base (YNB) broth (with amino acids and 2% glucose) at ~pH 5.5 for growing overnight at 30℃ with constant shaking at 250 rpm on the platform shaker. The 2 mL seed culture was pelleted and diluted in fresh YNB broth at ~pH 5.5 with an OD_{600} of 1.5 for growing at 30℃ with constant shaking at 250 rpm on the platform shaker for 2 h. After 2 h, the culture was pelleted and then diluted to an OD_{600} of 0.1 in pH 1.5 YNB broth adjusted by HCl. Three replicates of 200 µL cultures from each strain were added to the plate wells, and then OD_{600} was measured every 30 min for 48 h at 30℃ with constant shaking in the Synergy H1 Hybrid multimode microplate reader.

Cell growth and RNA sample collection

We first streaked the glycerol stock of the selected strains (Fig. 1a) on the YPD plate for growing overnight at 30℃. A single colony was picked up from the plate and inoculated into 2 mL of YNB broth (with amino acids and 2% glucose) at ~pH 5.5 for growing overnight at 30℃ with constant shaking at 250 rpm on the platform shaker. The 2 mL seed culture was inoculated into 15 mL of YNB broth at ~pH 5.5 for growing overnight at 30℃ with constant shaking at 250 rpm on the platform shaker. The overnight culture was pelleted and diluted in YNB broth at ~pH 5.5 with an OD_{600} of 1.5. We then grew the diluted culture for 2 h at 30℃ with constant shaking at 250 rpm on the platform shaker. After 2 h, the OD_{600} was measured and collected as the 0 h low pH treatment samples. The culture was pelleted and then diluted to an OD_{600} of 1.0 in YNB broth of various pHs adjusted by HCl (i.e., pH 1.5, 2.0, and 2.5) and grown for 1 h at 30℃ with constant shaking at 250 rpm on the platform shaker.

While our growth assays at pH 1.5 (Fig. 1b) were measured over a 48-h period, noticeable growth differences between tolerant and susceptible strains appeared as early as 1.5 h. Given that transcriptomic response to the switch in media conditions is typically quick (36), and that susceptible strains could not withstand long-term exposure to pH 1.5, we needed to collect RNA samples within a short timeframe from the exposure. Moreover, the upregulation of the glycosylphosphatidylinositol-anchored protein gene IoGAS1—a benchmark gene for low-pH tolerance—after just 1 h of exposure to pH 1.5, confirms the efficacy of this time point in capturing significant transcriptomic responses (Fig. S1). Thus, we picked the 1-h mark as a time point for the RNA-seq sample collection.

For organic acid treatments, *I. orientalis* SD108 seed culture was cultivated in YNB broth at pH 3.0 for 2 days prior to the acute 3-hydroxypropanoic acid (3-HP) and citramalate (CM) treatments. The pH of the YNB broth containing 3-HP and citramalate was adjusted to pH 3 using HCl. The 3-hydroxypropionic acid 30% solution (SKU# 792659-1G) and potassium citramalate monohydrate (SKU# 27455-1G-F) utilized were purchased from Sigma-Aldrich.
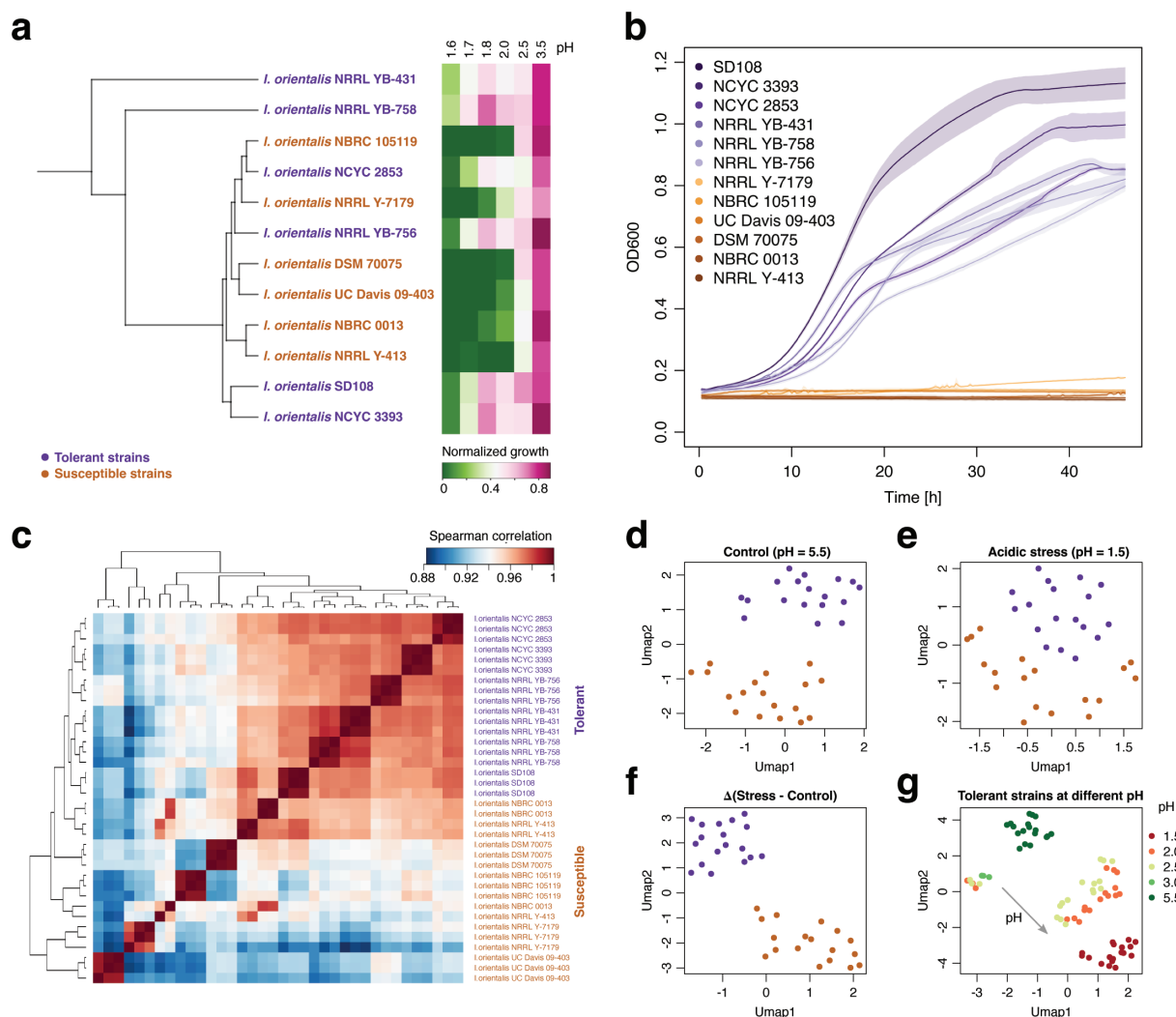
FIG 1 (a) Phylogeny of *I. orientalis* strains selected for this study. Six tolerant strains (purple) and six susceptible strains (orange) were selected. The heatmap shows the relative growth of each strain under low pH conditions compared to the control pH (~5.5). (b) Optical density (OD) curves representing growth profiles of 12 selected strains at pH = 1.5. Purple and orange curves represent tolerant and susceptible strains, respectively. (c) Pairwise Spearman correlation between gene expression profiles [log2(CPM) of all genes, after 1 h of treatment at pH 1.5] of all replicates of all strains. (d-f) Low-dimensional representations of transcriptomic profiles of all replicates of all strains (log2(CPM)) in the control condition (pH ~ 5.5) (d) or pH = 1.5 treatment after 1 h (e). Panel f shows the low-dimensional representations of changes in [log2(CPM)] transcriptomic profiles between 0 (control) and 1 h (treatment) samples, for all strains. Orange and purple points represent susceptible and tolerant strains, respectively. For dimensionality reduction, we first used principal component analysis (PCA) to obtain PCs, which were reduced to two-dimensional representations using UMAP. (g) UMAP plot of log2(CPM) counts for all replicates of tolerant strains subjected to varying pH levels. Colors represent the pH scale. The gray arrow indicates pH change toward more acidic media.

For every strain and culture condition, three replicates of the samples were collected, and the RNA samples were prepared by Direct-zol RNA Miniprep Plus Kit from Zymo Research.

RNA-seq sample collection and sequencing

Construction of libraries and sequencing on the Illumina NovaSeq 6000 were performed at the Roy J. Carver Biotechnology Center at the University of Illinois Urbana-Champaign. In brief, purified DNase-treated total RNAs were run on a Fragment Analyzer (Agilent, CA, USA) to evaluate RNA integrity. The total RNAs were converted into individually barcoded polyadenylated mRNA-seq libraries with the TruSeq Stranded mRNA Sample Prep kit (Illumina, CA, USA). Libraries were barcoded with Unique Dual Indexes (UDIs), which have

been developed to prevent index switching. The adaptor-ligated double-stranded cDNAs were amplified by PCR for eight cycles with the Kapa HiFi polymerase (Roche, IA, USA). The final libraries were quantitated with Qubit (ThermoFisher, MA, USA), and the average cDNA fragment sizes were determined on a Fragment Analyzer. The libraries were diluted to 10 nM and further quantified by qPCR on a CFX Connect Real-Time qPCR system (Biorad, CA, USA) for accurate pooling of barcoded libraries and maximization of number of clusters in the flow cell. The barcoded RNA-seq libraries were loaded on a NovaSeq S4 lane for cluster formation and sequencing. The libraries were sequenced from both ends of the fragments for a total of 150 bp per end. The fastq read files were generated and demultiplexed with the bcl2fastq v2.20 Conversion Software (Illumina, San Diego, CA, USA). The RNA-seq reads were mapped to *Pichia kudriavzevii* CBS573 by using HiSAT2 (37) (v2.2.0) with default arguments. Gene read counts were analyzed by featureCounts v2.0 with arguments: -s 2 C -p --primary -t exon. The reference genome sequences and gene annotation file were obtained from Mycocosm (https://mycocosm.jgi.doe.gov/Pickud1/Pickud1.home.html).

RNA-seq analysis

Differentially expressed gene sets

We first analyzed the differential expressions (DE) of tolerant and susceptible strains under HCl treatments. To this end, for each gene and strain, we used edgeR (38) to characterize the differential expression between low pH (1.5) and control conditions. A combined *P*-value was obtained for each group (tolerant and susceptible) using Fisher's method to combine differential expression *P*-values of the strains in that group. Expression profiles, differential expression *P*-values, and combined *P*-values were collectively used to define the following four gene sets (0.05 is used as the *P*-value significance level):

1. Tolerant-Exclusive: genes with a significant combined *P*-value (<0.05) in the tolerant group, which are significantly differentially expressed in the majority of tolerant strains (*P*-value < 0.05 in at least three out of six strains) with a consistent fold-change direction between low pH and control in the significant strains were selected. We additionally filtered out the genes that are potentially significant in the susceptible group by excluding the genes that have a significant combined *P*-value in the susceptible group and are significantly DE in at least one of the susceptible strains (fold-change directions were also considered when the gene is significant in more than one susceptible strain).

2. Susceptible-Exclusive: genes with a significant combined *P*-value (< 0.05) in the susceptible group, which are significantly differentially expressed in the majority of susceptible strains (*P*-value < 0.05 in at least three out of six strains) with a consistent fold-change direction between low pH and control in the significant strains were selected. We additionally filtered out the genes that are potentially significant in the tolerant group by excluding the genes that have a significant combined *P*-value in the tolerant group and are significant in at least one of the tolerant strains (fold-change directions were also considered when the gene is significant in more than one susceptible strain).

3. Tolerant-Susceptible-Flipped: genes with a significant combined *P*-value in both tolerant and susceptible groups were selected. We additionally demanded genes to be significantly differentially expressed in the majority of the strains in each group (significant *P*-value in at least three out of six strains in each group) with consistent intra-group but different inter-group fold-change directions (only significant strains were considered for fold-change analysis).

4. Stress-correlated: this gene set was obtained based on the expression profiles of tolerant strains in control, intermediate (two pH levels), and low pH conditions. Differential expression of each gene in each of the tolerant strains under different acidic stresses was assessed by Kruskal-Wallis *H*-test. Spearman's correlation between expression and pH levels (for each gene and strain) was additionally used to determine the strength and direction of response to acidic stress. Genes that show significant differential expression *P*-values (< 0.05) and significant correlation ($|corr| > 0.75$) in all the six tolerant strains were selected. We additionally demanded that each selected gene has a consistent response direction (i.e., the same correlation sign) in all the tolerant strains.

Similar to our analysis of HCl treatments, we assessed the differential expressions of SD108 tolerant strain under 3-HP and CM treatments. We used edgeR (38) to conduct four differential expression analyses between treatment and control for each of the 3-HP and CM at two treatment concentrations (40 and 80 g/L). We devised 3HP-CM-shared gene set by extracting the genes that show a significant differential expression (P-value < 0.05) in all of the four comparisons with a consistent fold-change direction.

GO enrichment analysis

The GO terms were obtained from the JGI database. We discarded those GO terms that had less than three gene associations among the set of 5,140 genes included in our study. We performed hypergeometric tests to find significant GO terms that were associated with the six gene sets defined in the previous section. The universe was defined as the set of all the 5,140 genes included in the study. A GO term was called significantly enriched for a gene set if the overlap between the respective gene set and genes associated with the GO term was found to be significant based on the hypergeometric *P*-value.

Gene regulatory network analysis

We used GENIE3 (39) to infer gene regulatory networks (GRNs) from the transcriptomic profiles of tolerant and susceptible groups. Specifically, we collected 87 transcriptomic profiles for the tolerant group including 72 samples from six tolerant strains under HCl treatments (control and pH 2.5, 2.0, and 1.5) and 15 samples from one of the tolerant strains (SD108) before and after treatments by 3-HP and CM. Moreover, we used 36 transcriptomic profiles for the susceptible group obtained from six susceptible strains under HCl treatments (control and pH 2.5, 2.0, and 1.5). GENIE3 was separately run on tolerant and susceptible profiles using 71 pre-selected TFs as regulators. For each group, we devised a GRN by extracting up to three regulators for each target gene with the highest GENIE3 scores among the TFs that show a GENIE3 score of at least 0.05 (exclusive) and are highly correlated (Spearman's correlation) with the target gene expression ($|corr| > 0.5$) in the same group.

We further analyzed the tolerant and susceptible GRNs to obtain differential GRNs. Specifically, for each group, we extracted the TF-gene regulations that are present in that group's GRN but either show relatively small Spearman's correlations (between TF's and target gene's expression) in the other group ($|corr| < 0.5$) or show different regulatory directions between the two groups (evidenced by the different sign of Spearman's correlation). This step results in tolerant-specific and susceptible-specific differential GRNs.

Motif enrichment analysis

To identify enriched motifs in promoter sequences of *I. orientalis*, we used GEMSTAT's (40) sequence annotation tool and motif position weight matrices (PWMs) from the YEASTRACT+ database (41). Promoters were defined as 1 kb upstream of each gene

and truncated shorter if they hit the previous gene (their sequences can be obtained from the MycoCosm portal). Promoter sequences were scored for motif presence using likelihood ratio (LR) (42). LR score for a promoter is obtained by taking the sum of the LR score of putative sites in the promoter. We used a threshold of 0.5 for the annotation tool. Thus, any location on a promoter, which is at least half as strong as the consensus site based on the LR score, is considered as a putative binding site. For a given TF, we assigned approximately the top 10% scoring promoters as putative targets and used them for enrichment analysis by hypergeometric test. The universe of genes consists of all the genes for which a promoter was present in the JGI database.

RESULTS

Strain selection and growth experiments

From a previous study on genotyping and phenotyping of 161 *I. orientalis* strains (35), we strategically selected two groups of pH-tolerant and pH-susceptible strains to study the transcriptional responses of *I. orientalis* under low pH stress (Fig. S2). Our strategy for strain selection included the following considerations: first, we only selected diploids, to exclude the impact of polyploidy; second, we selected two groups of strains with opposite growth patterns, i.e., the susceptible strains cannot grow on the synthetic defined (SD) media at pH 1.7, while tolerant strains grow on it almost as well on this media as in neutral pH; third, we diversified our selection by including strains from different clades, strains that are phylogenetically close and show similar tolerance, and strains that are phylogenetically close but show different tolerance. Based on these considerations, we chose six pH-tolerant and six pH-susceptible strains for physiological characterization and transcriptome analysis (Fig. 1a; Table S1).

We collected our samples from liquid cultures for ease of sample collection. To determine the low-pH condition for RNA-seq analysis, we set up liquid culture growth assays and noted that pH 1.5 is the most distinguishing in terms of the growth differences between tolerant and susceptible strains (Fig. 1b, and data not shown). We grew the cells in YNB media for 2 h with the addition of hydrochloric acid (HCl) to buffer pH up to 5.5, then transferred them to the media with a lower pH of 1.5, grew them for another hour, and collected total RNA for the transcriptome analysis (see Materials and Methods for details).

Global views of transcriptomic diversity under low pH stress

To determine genes and pathways that might be responsible for tolerance or susceptibility, for each strain, we collected three independent biological replicates at two time points (before and after low pH stress) for sequencing and construction of RNA-seq libraries, obtaining 72 samples in total (12 strains × two time points × three biological replicates). For the pH-tolerant strains, we collected additional samples at intermediate pH values (2.0, 2.5) to identify genes that respond gradually to the increase of acidic stress. This resulted in 36 additional RNA-seq samples (six tolerant strains × two conditions × three biological replicates).

First, to check the robustness of the data, we computed the pairwise Spearman correlation between transcriptomic profiles of all strains (all replicates) after 1 h of treatment at pH 1.5 (Fig. 1c, also see Fig. S3). We noted that tolerant strains are more similar to each other than susceptible ones, although two susceptible strains (*I. orientalis* NRRL Y-413 and *I. orientalis* NBRC 0013) have two replicates each, which are much closer to tolerant strains. Note that two tolerant strains (*I. orientalis* NRRL YB-431 and *I. orientalis* NRRL YB-758) are more phylogenetically distant from others (Fig. 1a), and these two strains had transcriptomic profiles distinct from the other tolerant strains (especially *I. orientalis* NCYC 2853, *I. orientalis* NCYC 3393, and *I. orientalis* NRRL YB-756).

We then asked if transcriptomic profiles of tolerant and susceptible strains are separable in a low-dimensional representation (Fig. 1d through f). We noted that the two groups of strains are separable in either control or treatment (pH 1.5 for 1 h) conditions

(Fig. 1d and e) but the separation is more conspicuous when plotting the changes in transcriptomic profiles (Fig. 1f), indicating a global difference in their responses to low pH stress. A 2D representation of tolerant strains treated at varying pH levels (Fig. 1g) suggested that global transcriptomic response presents a quantitative readout of the pH level.

Identification of differentially expressed genes

Our next goal was to identify genes involved in the transcriptomic response to low pH stress. To do this, we first identified genes that are differentially expressed in different pH conditions (Tables S2 and S3) and defined four gene sets based on the patterns of differential expression seen in different contrasts (Table 1; Fig. 2a and b; Table S4a through d) (see Materials and Methods for details). The gene sets "Tolerant-Exclusive" (531 genes) and "Susceptible-Exclusive" (766 genes) were found to be DE between low pH and control conditions exclusively in the tolerant and susceptible strains, respectively. A third gene set, "Tolerant-Susceptible-Flipped," comprises 166 genes that were differentially expressed in strains of either group but in opposite directions. Finally, the fourth gene set, "Stress-correlated," comprises 480 genes that are differentially expressed across the four different levels of acidic stress (pH 1.5, 2.0, 2.5, and control) in all six tolerant strains and whose expression levels are correlated with the pH level (Fig. 2b). We characterize these gene sets in the sections below.

Separately, we performed a direct comparison of expression levels (in pH 1.5 conditions) between tolerant and susceptible strains, identifying a set of 111 genes significantly upregulated at least twofold in one of the groups compared to the other (FDR corrected P-value < 0.01) with 36 genes upregulated in tolerant strains compared to susceptible and 75 upregulated in susceptible compared to tolerant (Fig. 2c). Among these genes, there are many transporters belonging to the major facilitator superfamily, which was recently shown to be involved in acidic stress tolerance in other fungal species (43). We also detected overexpression of genes encoding adhesin proteins, which were previously reported to be involved in pH resistance in *I. orientalis* by exhibiting cell aggregation (filamentous growth and flocculation) in extreme acidic environments (44).

Gene ontology analysis reveals biological processes involved in response to low pH

We tested for the enrichment of Gene Ontology (GO) terms in the four gene sets given in Table 1, relying on GO annotations in *S. cerevisiae* and predicted orthology between genes in *S. cerevisiae* and *I. orientalis* strains (Table 2; Tables S5 and S6; Fig. S4, also see Materials and Methods). Two major insights emerged from this analysis. First, several GO term enrichments pointed to the dysregulation of energy metabolism-related pathways. For instance, the Tolerant-Susceptible-Flipped gene set is associated with glycolysis

TABLE 1 Four gene sets are defined based on the patterns of differential expression across varying pH levels and strains[a]

Gene set	Size	Description
TE	531	Significantly DE in the same direction (between low pH and control conditions) in majority of tolerant strains (P-value < 0.05 in at least three out of six strains), but not in any of the susceptible strains
SE	766	Significantly DE in the same direction in majority of susceptible strains (P-value < 0.05 in at least three out of six strains), but not in any of the tolerant strains
TSF	166	Significantly DE in majority of tolerant as well as susceptible strains but in opposite directions
SC	480	Significantly DE (across varying pH levels) in majority of tolerant strains and expression change is correlated with pH levels

[a]See Materials and Methods for details. Gene sets: Tolerant-Exclusive (TE); Susceptible-Exclusive (SE); Tolerant-Susceptible-Flipped (TSF); and Stress-correlated (SC).
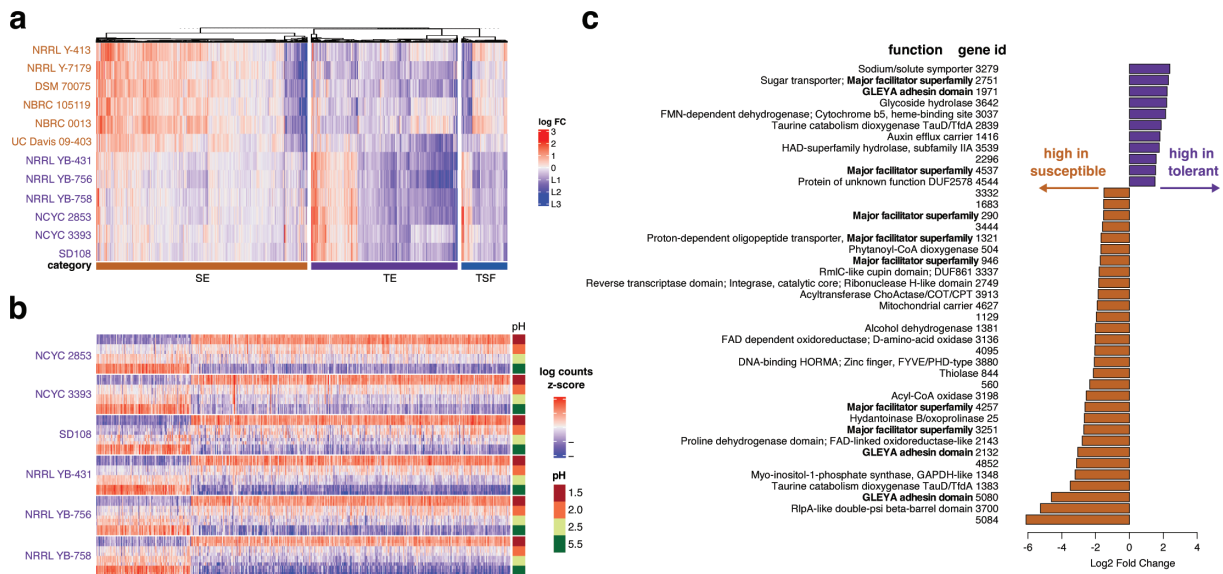
FIG 2 (a) Heatmap showing expression data: log fold change between pH 1.5 and control. Rows = each strain (grouped by susceptible/tolerant, in orange/ purple font, respectively). Columns = genes that are in the union of gene sets Susceptible-Exclusive (SE), Tolerant-Exclusive (TE), and Tolerant-Susceptible-Flipped (TSF). Columns are grouped by gene set. (b) Heatmap showing expression data: log-transformed expression. Columns = genes in Stress-correlated gene set. Rows = Six groups of rows, one group for each tolerant strain; within each group, the 12 replicates are in ascending order of pH. (c) Top genes differentially expressed in a direct comparison between tolerant and susceptible strains under pH 1.5 stress (all replicates of all strains of resistant vs all replicates of all strains of the tolerant group). Only genes that significantly vary (FDR corrected *P*-value < 0.01) with a log fold change of at least 1.5 between the two groups are shown.

(hypergeometric test *P*-value 3.5E-5). A study by Guo and Olsson (45) noted that exposing *S. cerevisiae* to weak acid stress impairs growth and inhibits glycolytic flux. Guan et al. (46) observed proteins of the glycolysis pathway to be key responders to propionic acid stress in the acid-tolerant species *Propionibacterium acidipropionici*. Indeed, Guan and Liu (34) note that microbes strengthen the glycolytic pathway for improved acid

TABLE 2 Significant GO terms associated with one or more of the four gene sets[a]

GO term	TE	SE	TSF	SC
DNA-directed RNA polymerase activity	**0.0001**	0.9892	0.8270	1
rRNA processing	**0.0001**	1	1	1
Small subunit processome	**0.0204**	1	1	1
Glycolysis	1	1	**0.0280**	1
Peptide-methionine-(S)-S-oxide reductase activity	1	1	**0.0349**	1
Oxidoreductase activity, acting on sulfur group of donors, disulfide as acceptor	1	1	**0.0349**	1
Ribosome	1	1	1	**0.0133**
Structural constituent of ribosome	1	1	1	**0.0133**
Electron transport	1	1	1	**0.0133**
Translation	1	1	1	**0.0133**
Tricarboxylic acid cycle	1	1	1	**0.0158**
Oxidoreductase activity, acting on NADH or NADPH	1	1	1	**0.0169**
Proteolysis	1	1	1	**0.0286**
ATP synthesis coupled proton transport	1	1	1	**0.0289**
Hydrogen ion transporting ATP synthase activity, rotational mechanism	1	1	1	**0.0289**
Hydrogen ion transporting ATPase activity, rotational mechanism	1	1	1	**0.0301**

[a]Numbers indicate FDR adjusted *P*-value. Bold font indicates FDR < 0.1. Gene sets: Tolerant-Exclusive (TE); Susceptible-Exclusive (SE); Tolerant-Susceptible-Flipped (TSF); and Stress-correlated (SC).

tolerance. Similarly, GO terms related to the TCA cycle, oxidoreductase activity, and hydrogen ion-transporting ATP synthase activity are enriched in the Stress-correlated gene set. This is consistent with prior research that showed that acid-stressed *S. cerevisiae* maintain pH homeostasis via H+-ATPase-mediated proton efflux (45, 47).

A second theme emerging from the GO analysis is the involvement of ribosome processing- and translation-related genes, which were enriched in Tolerant-exclusive and Stress-correlated gene sets. Ribosomes have been shown to be associated with tolerance to growth inhibitors and extension of cellular lifespan (48, 49), and ribosome-related pathways were previously found to be enriched in genes responding to acid stress in *S. cerevisiae* (50).

The above findings (Table 2) and related observations in the literature thus suggest changes in the energy metabolism pathways and translation machinery to be important aspects of response to low pH stress in tolerant strains of *I. orientalis*.

Transcription factor binding motif analysis identifies regulators of strong acid response

We sought to identify important transcriptional regulators of the large transcriptomic response to low pH observed above. For this, we considered the promoters of each gene set and tested them for the enrichment of binding motifs of transcription factors characterized in *S. cerevisiae* (41). Note that the motifs represent TF binding preferences in *S. cerevisiae* and hence may be noisy approximations for the orthologous TFs in *I. orientalis*; in such a case, a loss of sensitivity is possible, but statistically significant enrichments should point to biologically relevant TFs. Such cross-species use of motifs has been demonstrated in the literature (51).

As shown in Tables S7 and S8, several TF motifs were found to be enriched in the promoters of the Stress-correlated gene set (P-value < 0.001, FDR < 0.02), the most statistically significant motif being Mot3 (P-value 2.4E-6), which is known to be linked to osmotic stress response in *S. cerevisiae* (52). Another TF with motif enrichment in this gene set, Stb5, was previously reported to participate in resistance to weak acids in *S. cerevisiae*, with deletion of the transcription factor resulting in lower resistance to decanoic acid (53). Stb5 was also found to be among the three TFs with the largest number of regulatory targets among the genes required for acetic acid tolerance (54).

Motif enrichments in the other three gene sets were less prominent, so we examined the motifs significant at a nominal P-value of 0.05 (Table 3). The Mac1 motif was noted to be enriched in the Tolerant-Exclusive gene set (P-value = 3E-4). Mac1 is a copper ion-sensing TF (55), and copper homeostasis has been shown to be closely linked to acetic acid response in the acid-tolerant yeast *Zygosaccharomyces bailii* (56). The Azf1,

TABLE 3 Significant motif enrichments in any of the three DE gene sets TE, SE, and TSF[a]

TFs	Consensus	TE	SE	TSF	SC
Mac1	CGGNNNNNNNNCGG	**0.0003**	0.8234	0.9526	0.7082
Reb1	CAGCCAC	0.4092	**0.0109**	**0.0397**	0.1493
Lys14	WWNNNYARNHNNNVVCGR	**0.0109**	0.4484	0.6021	0.3989
Rtg1	NCCDTYNVNCCG	**0.0109**	0.9995	0.7017	**0.0121**
Rtg3	NCCDTYNVNCCG	**0.0109**	0.9995	0.7017	**0.0121**
Cup2/Haa1	CACATGC	0.8440	0.9290	**0.0130**	0.6501
Sfl1	CCGTTAACGG	0.1295	**0.0204**	0.3945	0.4613
Stb5	YCCNYTNRRCCG	**0.0314**	0.8234	0.9526	**0.0356**
Azf1	CACGTG	**0.0314**	0.7013	0.0649	0.7082
Ino4	CAGCGTG	**0.0314**	0.7457	0.3945	0.6501
Cup2/Haa1	HTHNNGCTGD	**0.0431**	0.1825	0.2185	0.1882
Cad1/Yap1	CGGCNNNNNNCGGC	**0.0431**	0.6040	0.1524	0.8514

[a]Shown are the P-values of association between the selected TF motifs (rows) and all four DE gene sets (columns). Bold font indicates FDR < 0.2. Criteria for inclusion of motifs: at least one of the three gene sets has an association with P-value < 0.05. Gene sets: Tolerant-Exclusive (TE); Susceptible-Exclusive (SE); Tolerant-Susceptible-Flipped (TSF); and Stress-correlated (SC).

Rtg1, and Rtg3 motifs are also enriched in this gene set. Rajkumar et al. (57) have previously shown that inserting Azf1 motifs is an effective way to induce a synthetic promoter to respond to low pH stress in *S. cerevisiae*, suggesting that Azf1 plays an important role in such response. Mitochondrial Retrograde (RTG) pathway is known for its involvement in acetic acid stress response (58). The Cup2/Haa1 motif was found enriched in the promoters of the Tolerant-Susceptible-Flipped gene set. This motif has been shown to increase low pH-induced gene expression (57).

In summary, motif analysis of tolerance-related gene sets revealed their potential regulation by TFs previously implicated in acid tolerance/response [Stb5 (53), Azf1 (57), Rtg1/Rtg3(58), and Cup2/Haa1(59)] and osmotic stress response (Mot3) in *S. cerevisiae*, and thus provided a first glimpse into the underlying GRN in tolerant strains of *I. orientalis*.

Gene regulatory networks underlying response to low pH stress in tolerant strains

Our data reveal a large variation in gene expression under different treatment conditions, including varying levels of acid stress and control conditions, in six tolerant strains of *I. orientalis*. A transcriptional regulatory network, comprising edges representing regulatory relationships between TFs and target genes, is a widely popular construct to simplify the presentation of such transcriptomic variation, while also providing testable hypotheses about gene regulation. Such TRNs can be reconstructed systematically from the expression profiles of varying biological conditions, by examining the co-expression patterns between TFs and genes (60). We used a state-of-the-art computational method for this task, called GENIE3 (39), which fits a random forest model to explain target gene expression as a non-linear function of TFs' expression and infers TF-gene edges based on which TFs were most useful in this modeling task. Specifically, we considered 87 transcriptomic profiles corresponding to six tolerant strains under low pH (1.5, 2.0, and 2.5) and control conditions (as well as organic acid stress conditions) and used GENIE3 to identify the putative regulators of each gene (see Materials and Methods). The resulting TRN comprises 11,650 edges involving 71 TFs and 4,671 target genes (Table S9). While the TRN was reconstructed based entirely on our data from *I. orientalis* strains, we refer to the TFs of the network by the names of their *S. cerevisiae* orthologs for easier contextualization with the literature.

We examined the TFs that had the most predicted targets (so-called "hubs") in the TRN (Fig. 3a; Table S9) and the differential expression status of their targets (Fig. 3b). The TF with the most targets is Dal80, known for its involvement in resistance to nitrogen starvation in *S. cerevisiae* (61). Nearly a third of its predicted targets are differentially expressed exclusively in tolerant strains, while over half belong to one of the four DE gene sets examined above. Three of the top 10 regulators—Kar4, Arg81, and Nrg1/2— primarily target DE genes of the Stress-correlated set. Of these, Kar4 and Nrg1/2 additionally had their predicted regulons supported by the presence of their respective motifs in the promoters of target genes (Fig. 3c). Nrg1 has been shown to be an important regulator in response to acetic acid (62) and to repress expression in response to low pH (57). Three other hub TFs—Put3, Fzf1, and Swi4—were observed to have predicted targets primarily in the Susceptible-exclusive gene set. This appears counterintuitive at first since the TRN that led to the identification of these regulators was derived from tolerant strains. However, this TRN is not meant to exclusively explain tolerant strain gene expression, and in fact, these three TFs are hub TFs in a corresponding TRN derived from susceptible strains as well (Fig. S5a; Table S10). Our interpretation thus is that these three TFs are key regulators of expression variation in *I. orientalis*, but not specifically for the stress-induced regulation in tolerant strains. We revisit this point in the next section, with the presentation of differential TRNs.

We reconstructed the TRN for susceptible strains following an identical procedure as above, deriving a network with 12,965 edges involving 71 regulators and 4,885 targets (Table S10; Fig. S5a). Comparing individual regulons from TRNs inferred from susceptible
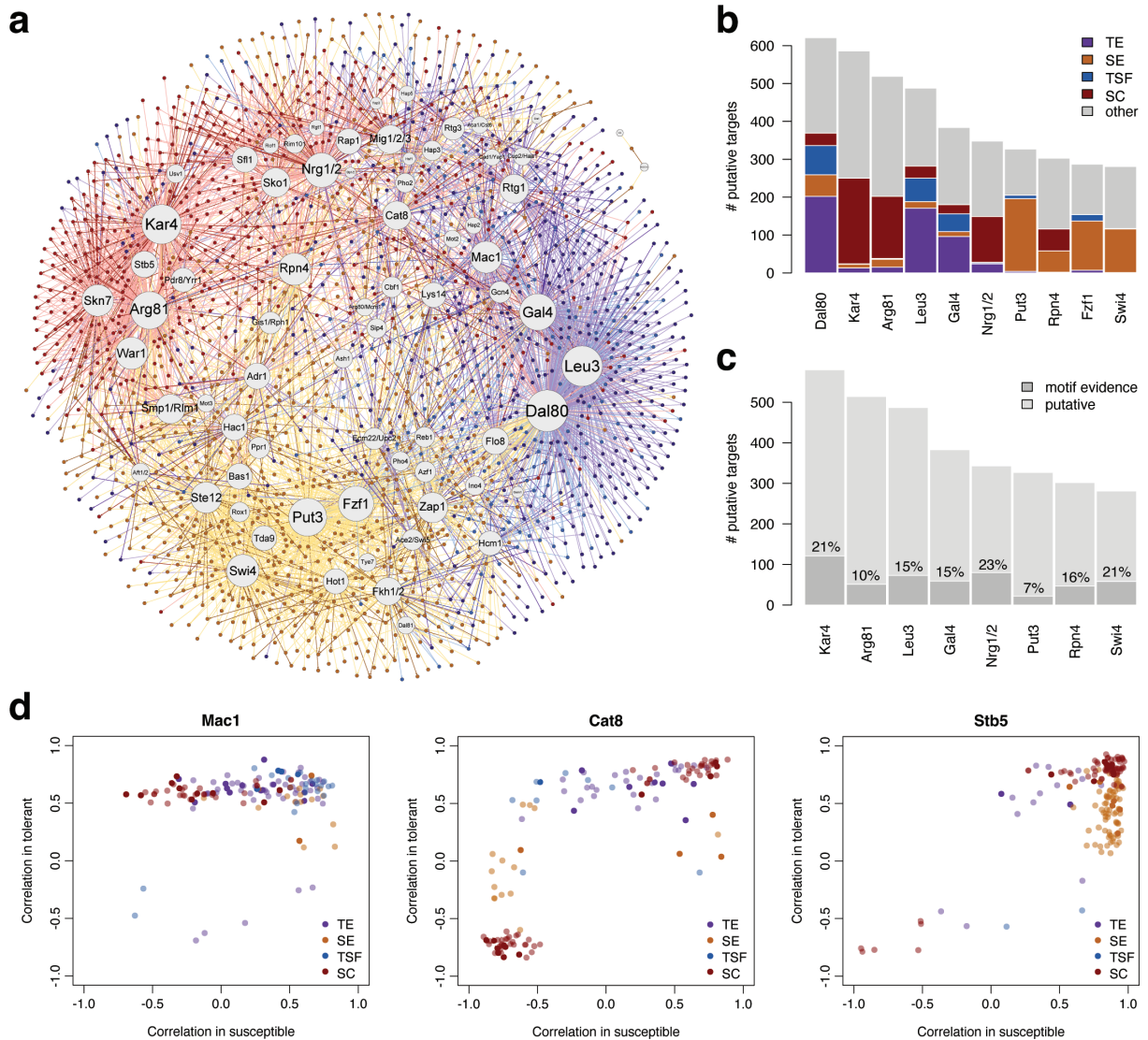
FIG 3 (a) GENIE3 reconstructed TRN for all tolerant strains (see Table S9). Only genes from the DE gene sets are shown. Nodes are colored according to the corresponding gene set: TE (Tolerant-Exclusive)—purple, SE (Susceptible-Exclusive)—orange, TSF (Tolerant-Susceptible-Flipped)—blue, SC (Stress-correlated)—red, and TFs—gray, while their node size is proportional to the number of putative targets. Edges are colored according to the target gene set, light edges are GENIE3 predicted targets, while dark edges have additional motif presence evidence. (b) Top regulators in TRN and the gene sets they primarily regulate. (c) Motif evidence for discovered putative targets of top regulators. Dal80 is not shown here as none of the known motifs were detected. (d) Stb5, Mac1, and Cat8 targets in the TRNs for tolerant and susceptible strains and their expression correlation to the corresponding TF. Colors correspond to different DE gene sets; darker colors indicate the presence of the TF's motifs in the target promoter region (see text).

and tolerant strains (Fig. 3d; Fig. S5 and S6), we noted several trends. First, strong hubs from one TRN are typically highly correlated with their targets in the corresponding set of strains (tolerant or susceptible) but not in the other, suggesting a degree of exclusivity in regulatory control of tolerant versus susceptible strains. Second, gene targets from the Stress-correlated gene set are highly correlated (positively or negatively) with the expression of their regulator in both strain groups (note red points in Fig. 3d; Fig. S6), while gene targets from Tolerant-Exclusive or Susceptible-Exclusive gene sets tend to be correlated with the expression of their regulator only in the corresponding set of strains (e.g., Stb5 in Fig. 3d). This is in line with the exclusivity of control between the two types of strains, as suggested above.

TABLE 4 Top five hubs of the differential TRN for tolerant strains[a]

TF	No. of targets
Leu3	76
Dal80	40
Cat8	40
Mac1	37
Ecm22/Upc2	26

[a]Edges of differential TRN were sorted in descending order of absolute difference between TF-gene correlation in tolerant strains and in susceptible strains, the top 500 edges were examined, and TFs were sorted by the number of targets.

Differential gene regulatory networks of tolerant and susceptible strains

To focus on the regulatory mechanisms specific to tolerant strains, we derived a differential TRN that comprises regulatory relationships supported only by tolerant strains' transcriptomic profiles (Table S11; see Materials and Methods). For this, we examined the top 500 edges in order of their specificity to tolerant strains and found the TFs Leu3, Dal80, Cat8, Mac1, and Ecm22/Upc2 to be the major hubs of the network (Table 4). Of special interest among these hubs is Mac1, a copper ion-sensing TF (55) whose binding motif was noted to be significantly over-represented in the promoters of the Tolerant-Exclusive gene set (Table 3), further supporting a regulatory role unique to the tolerant strains.

Closer inspection of the targets of Mac1 in the global TRN (Fig. 3d) revealed the genes primarily involved in amino acid metabolism, glycolysis, TCA cycle, lipid and peptidoglycan biosynthesis, as well as Porphyrin, Vitamin B6, Thiamine, Riboflavin, and Glutamate metabolism. These are known to play a role in response to oxidative and low pH stress, which might be evidence of the crucial role of Mac1 TF in the mitigation of acidic stress in *I. orientalis*. Among its targets, there are also a few genes involved in metal/copper/iron ion transport, binding, oxidation, and protein kinase CK2 activity, whose linkage to acidic stress is corroborated by the previous study of the acid-tolerant yeast *Z. bailii* (56).

Other hubs of the differential TRN include Cat8 (Fig. 3d), which is known to be required for evasion from acetic acid-induced programmed cell death when *S. cerevisiae* is grown in raffinose (58), and Ecm22/Upc2, previously shown to be associated with the regulation of transcriptomic response of *S. cerevisiae* during fermentation under oleic acid and ergosterol depletion (63). Interestingly, in our data, Cat8 primarily associated with targets from tolerant-exclusive and stress-correlated gene sets, sharing a noticeable fraction (~30%) with Mac1, suggesting tight cooperation of these TFs in acidic stress response mitigation. In contrast, the Ecm22/Upc2 regulator seems to be Susceptible-Exclusive. See Table S12 for a similar examination of the differential TRN of susceptible strains.

Note that our TRN reconstruction identifies TF-gene relationships solely based on coordinate expression between the TF and target gene across varying conditions, without consideration of the pH treatment levels of those conditions. Thus, we next focused our attention on TFs whose regulatory targets in the differential TRNs are enriched for genes differentially expressed between low pH and control conditions. For this, we identified TFs whose target genes in the differential TRN (of tolerant strains) include the largest fraction of genes from the Tolerant-Exclusive gene set. As shown in Table 4, the top TFs thus identified included Stb5, which has 25 target genes in the differential TRN for tolerant strains, and 12 of these 25 targets (48%) are in the Tolerant-Exclusive gene set. The TF's correlation with the 25 target genes is 0.61 (median) in the data on tolerant strains and only 0.31 (median) in susceptible strains, indicating its tolerant-specific regulatory effect; furthermore, the Stb5 motif was found to be enriched in the promoters of the Tolerant-Exclusive gene set (P-value < 0.05, Table 3). Examination of Stb5 targets in the global TRN (Fig. 3d) shows that it might be involved in the regulation of other parts of glycan and glycerolipid biosynthesis pathways as well as arginine and proline metabolism, which were previously associated with response to

acidic stress. Another prominent activity of Stb5 in *I. orientalis* is that it is associated with many protein kinases (NF-kappaB, SAP, JUN, casein, etc.), the most notable being MAP kinase, known to be a key player in stress response in yeast (64). This TF has been previously shown to be involved in resistance to octanoic acid in yeast (53) and was among the three TFs with the greatest regulatory role on genes required for acetic acid tolerance (54).

Another important TF revealed in part (a) of Table 5, Rtg1, has a regulon of 128 target genes; these genes have a median expression correlation of 0.57 with the TF when considering tolerant strains but only 0.16 when analyzing susceptible strains, thus underscoring the differential nature of those regulatory relationships; moreover, 41% of these target genes belong to the Tolerant-Exclusive DE gene set. The RTG pathway is known for its involvement in acetic acid stress response (58), along with Adr1, a hub TF of the TRN for susceptible strains (Table S10) and Cat8p, a hub TF of the differential TRN for tolerant strains (Table 4).

Similar analyses revealed the hub TFs of the differential TRN whose target genes belong to the Tolerant-Susceptible-Flipped gene set (Table S13); these included Rtg3, also of the RTG pathway noted above, and Haa1, whose binding sites have been shown to increase low pH-induced gene expression (57). The Haa1 motif was found above to be enriched in the Tolerant-Susceptible-Flipped gene set (*P*-value 0.01, Table S7), and its regulatory effects have been recorded in *S. cerevisiae* adaptation to weak acids (59). Hub TFs whose regulons are enriched for the stress-correlated gene set (Table S14) include Reb1, whose motif is strongly enriched in the Tolerant-Susceptible-Flipped gene set (*P*-value = 0.04).

We next examined hubs of the differential TRN for susceptible strains, focusing on TFs whose regulons are statistically enriched for DE gene sets (part (b) of Table 5; Tables S15 and S16), and found several of the same TFs we had noted as being hubs of the differential TRN for tolerant strains, e.g., Stb5, Nrg1/Nrg2, Skn7, Hac1, and Mac1. This observation suggests that certain TFs have context-specific regulatory roles on different modules of pH-responsive genes in tolerant and susceptible strains. At the same time, the differential TRN for susceptible strains also revealed a hub TF not observed in the analyses above: Rlm1/Smp1, which has 103 targets in the TRN. These targets have a median correlation of 0.82 with the TF in data from the susceptible strains and a far lower median correlation of 0.29 in tolerant strains. Moreover, 50% of these targets

TABLE 5 Top regulators in the differential GRN of tolerant (a) and susceptible (b) strains, sorted by association with Tolerant-Exclusive (a) and Susceptible-Exclusive (b) gene sets

(a)

TF	No. of targets in tolerant GRN[a]	Median correlation in tolerant strains	Median correlation in susceptible strains	Fraction of targets in TE gene set
Stb5	25	0.61	0.31	0.48
Rtg1	128	0.57	0.16	0.41
Skn7	26	0.59	0.38	0.38
Leu3	371	0.64	0.22	0.37
Dal80	399	0.64	0.28	0.35

(b)

TF	No. of targets in susceptible GRN[a]	Median correlation in tolerant strains	Median correlation in susceptible strains	Fraction of targets in SE gene set
Smp1/Rlm1	103	0.29	0.82	0.50
Stb5	98	0.37	0.86	0.40
Nrg1/Nrg2	164	0.30	0.80	0.36
Skn7	78	0.28	0.79	0.36
Arg81	117	0.30	0.75	0.36

[a]Number of targets is the number of genes predicted to be regulated by a TF in the differential GRN of tolerant strains. Median Spearman correlation coefficient between the TF and its target genes, computed from expression data in the tolerant and susceptible strains, respectively. TE—Tolerant-Exclusive gene set and SE—Susceptible-Exclusive gene set.

are differentially expressed (between low pH and control) exclusively in the susceptible strains. Rlm1 binding sites have been previously shown to increase response to low pH in yeast (57) and this TF is a key player in the CWI (cell wall integrity) pathway, which is a mechanism of response to low pH stress in *S. cerevisiae*.

In summary, differential TRN reconstruction and intersection with differential expression analysis revealed key TFs underlying transcriptomic responses of tolerant and susceptible strains to low pH stress.

Discriminating transcriptomic response to low pH from that of organic acids

I. orientalis was selected as a target organism in this study because it can tolerate low pH, potentially reducing production costs of organic acids such as 3-hydroxypropanoic acid and citramalate as well as other industrially relevant chemicals. Since our growth assays demonstrated (Fig. 1c) the tolerance of certain strains of *I. orientalis* to low pH media with HCl, we sought to determine the molecular and gene regulatory basis of such tolerance over and above the response to weak organic acids such as 3-HP and CM. Thus, for the strain with the highest acid tolerance, *I. orientalis* SD108, we used transcriptomic profiling (see Materials and Methods) to study the effect of different concentrations of organic acids 3-HP or CM in the media, in search for molecular mechanisms unique to HCl-induced low pH tolerance.

Examining the new transcriptomic profiles globally, we observed that the effect of these acids is different from that of low pH stress (Fig. 4a). We then identified genes that were differentially expressed between treatment and control at either concentration (40 and 80 g/L) and defined a gene set called "3HP-CM-shared" that comprises DE genes common to both treatments (see Materials and Methods and Table S17). This gene set includes 25.1% of the genes in the Tolerant-Exclusive gene set, 20.6% of the Stress-correlated set, and 25.9% of the Tolerant-Susceptible-Flipped set from the HCl-treatment experiments (Fig. 4b), while its overlap with the Susceptible-Exclusive gene set (9.4% of the latter) is much smaller, as expected since the organic acid response was profiled in a tolerant strain. The genes in these overlaps are the primary targets for future strain engineering experiments.

The 3HP-CM-shared gene set was significantly enriched for several GO terms related to energy metabolism and ion transport (Table S18; Fig. S7). This is consistent with prior
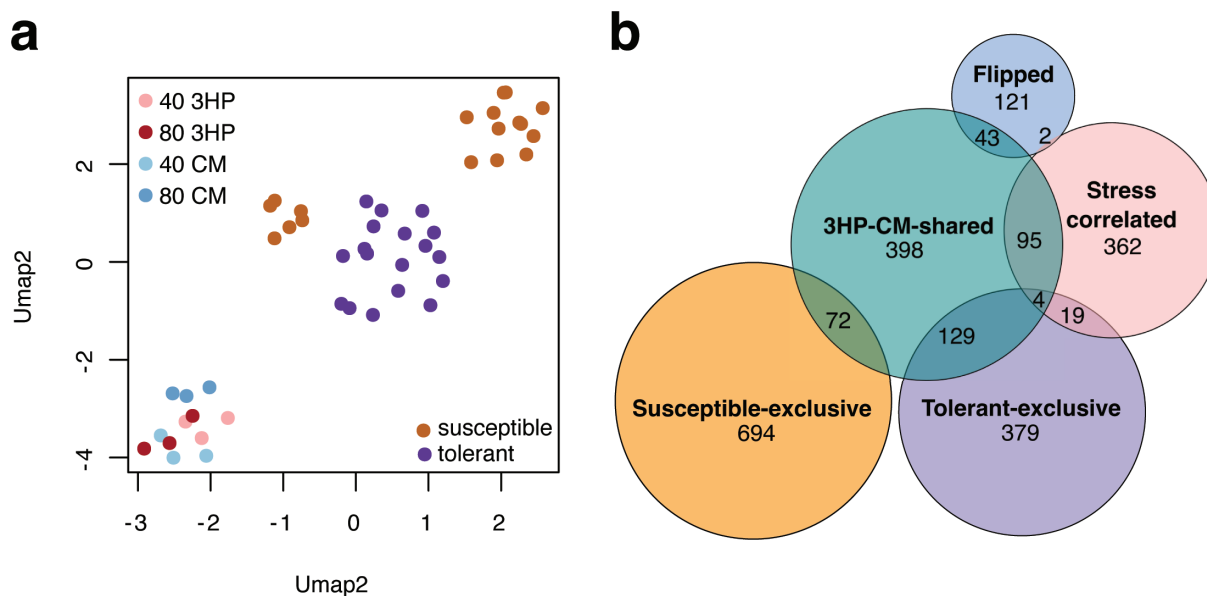

FIG 4 (a) UMAP with 3HP- and citramalate (CM)-treated samples for *I. orientalis* SD108 (samples of pH = 1.5-treated tolerant and susceptible strains are also included for the reference). The control background of pH = 5.5 is subtracted for each strain. (b) Venn diagram showing overlap of 3HP-CM-shared with Tolerant-Exclusive, Susceptible-Exclusive, Tolerant-Susceptible-Flipped, and Stress-Correlated gene sets.

studies that have found *S. cerevisiae* to counteract changes in cytoplasmic pH and proton gradient by activating the plasma membrane proton pump (65). The recorded response to organic acids also involved iron-sulfur cluster binding, a phenomenon observed in the response of the bacteria *Bacillus megaterium* to acid stress (65). The 3HP-CM-shared gene set also exhibited strong statistical enrichment for many TF binding motifs (Table S19), which included some TFs that were noted above in the analysis of HCl-responsive gene sets, e.g., Nrg1/Nrg2, Stb5, and also several TFs not observed above but known for their involvement in organic acids response, e.g., Stp1 (66), Mig1 (67), and Tda9 (68). This suggests to us that careful examination of transcriptomic responses has the potential to reveal molecular mechanisms that differ between the weak organic acid response and the response to HCl-induced low pH conditions.

To pursue the above possibility further, we defined a set of differentially expressed genes that are unique to the HCl response. This set, called the low-pH-Specific gene set, comprises genes that belong to any of the four HCl-responsive DE gene sets defined above but not to the 3HP-CM-shared gene set. The gene set was found to be significantly enriched (Table S20) not only for broad biological processes such as RNA processing and nucleic acid binding (FDR < 0.1) but also for highly specific processes such as protein disulfide isomerase activity (*P*-value 0.003), trehalose biosynthetic process (*P*-value 0.013), and glycolysis (*P*-value 0.018). Interestingly, trehalose has been previously demonstrated to increase tolerance to extremely low pH in the probiotic strain *Lactobacillus acidophilus* NCFM (69), and this finding had not emerged in the GO analysis of HCl-response without considering the difference from organic acid response.

DISCUSSION

Developing a microbe robustly tolerant to low pH is a critical step for more cost-effective and eco-friendly production of bio-based organic acids. *I. orientalis* is a promising chassis to engineer such systems on an industrial scale. To improve our understanding of specific mechanisms involved in tolerance to low pH stress of this species, we performed an extensive RNA-seq profiling of 12 strains of *I. orientalis* and detected and analyzed reproducible gene expression patterns associated with acidic conditions. The strains in our study were strategically selected to provide a sharp contrast between tolerant and susceptible phenotypes not explainable by their shared phylogeny. Overall, the transcriptomes of tolerant and susceptible strains are clearly separable, and the distance between transcriptomes correlates well with the pH level of the treatment. We identified several hundreds of genes that respond exclusively in tolerant strains or exclusively in susceptible strains. Over 150 genes responded in the opposite directions in tolerant and susceptible strains. Furthermore, the expression of nearly 500 genes is significantly correlated with the pH stress level of treatment in tolerant strains.

Gene Ontology analysis reveals that the different transcriptomic responses of tolerant and susceptible strains involve many energy- and metabolism-related processes. Systemic changes in cell's energy and metabolism are perfectly understandable since low pH causes a significant increase in energy demand to maintain intracellular pH homeostasis (34). The importance of energy and metabolism-related processes in low pH tolerance has been previously reported in *S. cerevisiae* (45), *P. acidipropionici* (46), and other organisms [see reference (34) for a recent review]. We also observed changes in the ribosome and translation-related genes, consistent with earlier findings in *S. cerevisiae* (49, 50).

We then explored transcription factors responsible for the transcriptomic shifts in tolerant strains, explaining the differences thereof between tolerant and susceptible strains. One way to do that is via the analysis of known *cis*-regulatory motifs upstream of differentially expressed genes. Since these motifs are known to be generally conserved across species, we were able to perform enrichment analysis for *S. cerevisiae* motifs in *I. orientalis* promoter regions. More specifically, we found motifs of *S. cerevisiae* transcription factors Mot3 and Stb5 to be overrepresented upstream of the set of genes correlated with the pH level; motifs Mac1, Rtg1, and Rtg3 upstream of the set of genes

differentially expressed exclusively in tolerant strains and motifs Cup2/Haa1 upstream of genes shifting in the opposite direction in tolerant and susceptible strains. These TFs were previously described to be involved in acidic tolerance in *S. cerevisiae* (53, 57–59, 64), suggesting that their function might be conserved across species, i.e., their orthologs in *I. orientalis* also regulate genes important for acidic stress relief.

The other method relied on a random forest-based network inference algorithm GENIE3 (39) to infer TRNs directly from transcriptomic profiles of tolerant and susceptible groups. We then used orthology to match identified *I. orientalis* TFs to known *S. cerevisiae* TFs and analyzed their respective regulons. This method is complementary to motif-based analysis since it predicts TF-regulated gene pairs without explicit knowledge of the binding motif (or several motifs) mediating these regulatory interactions. Going one step further and comparing TF-gene relationships in TRNs for tolerant and susceptible strains, we found the TFs Mac1, Stb5, and Rtg1/Rtg3, noted also in the motif-based approach above, to be among the "hub TFs" in the differential TRN. This lends additional credibility to their role in response to low pH stress since two complementary approaches (motif-based and expression-based) reported these three TFs among the top predictions. The implied involvement of the RTG pathway is especially noteworthy, given its well-studied role in tolerance to acid stress in *S. cerevisiae* (58, 70).

Further work is needed to experimentally test the specific role of these TFs through knockdown of the TF followed by either RNA-seq or phenotypic profiling. Specific TF-gene edges may be tested by augmenting the TF knockdown assays with genome-wide cis-regulatory profiling using ChIP-seq assays. In particular, Dal80 ortholog could be a promising candidate for future investigation. It is a top regulator of tolerant strains' TRN and differential TRN (Table 4; Tables S9 and S11). Remarkably, a third of Dal80's target genes belong to the Tolerant-Exclusive gene set (see Fig. 3b and part (a) of Table 5). Furthermore, Dal80 expression is flipped between Tolerant and Susceptible strains (Table S4), and its contrasting behavior could greatly influence target genes, resulting in differences between tolerant and susceptible strains at low pH. Such validation will be a crucial step in bridging our work with strain engineering based on the validated findings.

However, due to the complex and context-dependent regulatory nature of specific hub TFs, predicting the consequences of engineering them for enhanced low-pH tolerance can be challenging. An alternative strategy to enhance low-pH tolerance might focus on downstream targets directly linked to physiological functions, including genes involved in pH homeostasis and modifications of cell membranes and walls. For example, the genes of hydrogen ion-transporting ATP synthase and ATPase activities found in the Stress-correlated set are good targets for evaluation because of their potential functionality without the need for strain- or species-specific partner proteins or cofactors. Furthermore, genes that show differential expression when directly comparing expression levels under pH 1.5 conditions between tolerant and susceptible strains highlight the inherent transcriptional variations between the two strain groups (Fig. 2c). The enrichment of major facilitator superfamily transporters and cell aggregation-related proteins in this comparison suggests these genes could be preferable targets for evaluation.

The existing literature is relatively sparse regarding genes and regulatory mechanisms involved in low pH tolerance in *I. orientalis*. One of the few such studies identified a putative glycosylphosphatidylinositol-anchored protein gene IoGAS1 to be important for low pH tolerance, likely through maintaining cell wall integrity during low pH conditions (71). IoGAS1 is induced by low pH conditions, and the overexpression of IoGAS1 in *S. cerevisiae* enables growing under highly acidic conditions (i.e., pH 2.0) for a variety of different acids (71, 72). The authors suggest that this effect is mediated by the CWI pathway and the Rlm1 transcription factor. Interestingly, we found Rlm1 to be the top TF in the differential TRN of susceptible strains in terms of targets that are differentially expressed exclusively in susceptible strains, suggesting that acid tolerance in *I. orientalis* may involve rewiring of the CWI response to acid stress and transcriptomic changes of genes in this pathway. Another recent study showed that while *I. orientalis* shows greater

tolerance to low pH conditions than *S. cerevisiae*, it might be due to the changes in membrane composition, induction of arginine catabolism, switch to filamentous growth, and formation of tightly aggregated colonies (44). The last finding corroborates our observation that susceptible strains of *I. orientalis* increase the expression of adhesins while using ammonia as a proton acceptor.

While general acidic tolerance of *I. orientalis* is crucial to the industrial use of this strain, we note that specific mechanisms involved might vary depending on the acid since it imposes two types of stress on the organism: low-pH and solvent toxicity. To address this issue, we performed an additional set of experiments for the most pH-tolerant strain in our collection, *I. orientalis* SD108, in various concentrations of two potential bioproduct organic acids, 3-hydroxypropanoic acid and citramalate. We then identified genes differentially expressed in response to HCl but not differentially expressed in response to the organic acids in SD108. These genes, likely more specific to the low-pH stress of acids, showed enrichment in pathways associated with RNA processing, glycolysis, and trehalose biosynthesis. Higher levels of trehalose have been previously observed in *S. cerevisiae* under acid stress caused by acetic, formic, levulinic, and cinnamic acids (45), while increased glycolytic levels have been reported as a mechanism to compensate for inhibition in enzyme activity due to low pH and thus rescue normal metabolism (34).

However, for future efforts toward bioengineering *I. orientalis* strains, genes differentially expressed in response to both HCl and organic acid are of greater interest and should be the primary targets. In particular, we identified 43 genes that are not only induced in 3-HP and citramalate media but also flip their expression when comparing tolerant and susceptible strains. Notably, this set includes several TFs, which are orthologs of Hcm1, Zap1, and Dal80 in *S. cerevisiae*.

Our study has several limitations. First, we note that bioinformatic reconstruction of GRNs is an area of active research and can make frequent errors in predicting TF-gene relationships (73). For this reason, our analysis interpretation underplayed the individual GRN edges and emphasized the "hub" TFs of various networks, since aggregating the GRN information at the level of TFs improves the reliability of inference. Identifying differential GRNs and their hub TFs also adopts the same strategy of aggregating evidence (in this case from two GRNs) to reduce erroneous predictions. Similarly, the motif analysis based on position-weight-matrix scanning of gene promoters is likely to be fraught with false positives, a recognized problem (74) that is mitigated partially by doing enrichment tests on gene sets harboring the same motif. Future work profiling genome-wide TF-DNA binding through experimental assays such as ChIP-seq and ATAC-seq may provide greater resolution and reliability to that analysis.

Finally, in our study, we focus on acute stress response to low pH since it enables us to compare tolerant strains to the baseline of susceptible ones, which cannot tolerate acidic environments and die. Future studies should explore long-term low-pH stress experiments on tolerant strains. This would elucidate the regulatory differences between acute and prolonged low-pH stress adaptations.

AUTHOR AFFILIATIONS

[1]Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

[2]Department of Bioengineering, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

[3]The Gladstone Institute of Data Science and Biotechnology, San Francisco, California, USA

[4]Center for Biophysics and Quantitative Biology, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

[5]Center for Advanced Bioenergy and Bioproducts Innovation, Lawrence Berkeley National Laboratory, Berkeley, California, USA

[6]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

[7]Department of Computer Science, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

[8]Department of Physics, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

[9]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA

[10]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

[11]Global Institution for Collaborative Research and Education, Hokkaido University, Hokkaido, Japan

[12]Institute of Global Innovation Research, Tokyo University of Agriculture and Technology, Tokyo, Japan

[13]Cancer Center at Illinois, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

[14]Department of Biomedical Engineering at Georgia Tech and Emory University, Atlanta, Georgia, USA

[15]Department of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

AUTHOR ORCIDs

Veronika Dubinkina http://orcid.org/0000-0002-4844-6795
Ping-Hung Hsieh http://orcid.org/0000-0003-3528-8617
Sergei Maslov http://orcid.org/0000-0002-3701-492X
Yasuo Yoshikuni http://orcid.org/0000-0002-8372-640X
Saurabh Sinha http://orcid.org/0000-0001-6033-7746

AUTHOR CONTRIBUTIONS

DATA AVAILABILITY

Raw RNA-seq expression data have been uploaded to SRA under accession number PRJNA950397. Processed TPM gene expression counts are available in Table S2.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Fig. S1 to S7 (Spectrum02536-23-S0001.docx). Supplemental figures.
Tables S1 and S3 to S20 (Spectrum02536-23-S0002.xlsx). All supplemental tables except Table S2.
Table S2 (Spectrum02536-23-S0003.xlsx). TPM gene expression counts for all samples in this study.

REFERENCES

1. Hassan SS, Williams GA, Jaiswal AK. 2019. Moving towards the second generation of lignocellulosic biorefineries in the EU: drivers, challenges, and opportunities. Renewable and Sustainable Energy Reviews 101:590–599. https://doi.org/10.1016/j.rser.2018.11.041
2. Lange J-P. 2007. Lignocellulose conversion: an introduction to chemistry, process and economics. Biofuels, Bioprod. Bioref 1:39–48. https://doi.org/10.1002/bbb.7
3. Balan V. 2014. Current challenges in commercially producing biofuels from lignocellulosic biomass. ISRN Biotechnol 2014:463074. https://doi.org/10.1155/2014/463074
4. Singhvi MS, Gokhale DV. 2019. Lignocellulosic biomass: hurdles and challenges in its valorization. Appl Microbiol Biotechnol 103:9305–9320. https://doi.org/10.1007/s00253-019-10212-7
5. Solarte-Toro JC, Romero-García JM, Martínez-Patiño JC, Ruiz-Ramos E, Castro-Galiano E, Cardona-Alzate CA. 2019. Acid pretreatment of lignocellulosic biomass for energy vectors production: a review focused on operational conditions and techno-economic assessment for bioethanol production. Renewable and Sustainable Energy Reviews 107:587–601. https://doi.org/10.1016/j.rser.2019.02.024
6. Galbe M, Zacchi G. 2012. Pretreatment: the key to efficient utilization of lignocellulosic materials. Biomass and Bioenergy 46:70–78. https://doi.org/10.1016/j.biombioe.2012.03.026
7. Caspeta L, Castillo T, Nielsen J. 2015. Modifying yeast tolerance to inhibitory conditions of ethanol production processes. Front Bioeng Biotechnol 3:184. https://doi.org/10.3389/fbioe.2015.00184
8. Yang M, Wang J, Nan Y, Zhang J, Li L, Liu G, Vepsäläinen J, Kuittinen S, Pappinen A. 2019. Effect of salts formed by neutralization for the enzymatic hydrolysis of cellulose and acetone-butanol-ethanol fermentation. RSC Adv 9:33755–33760. https://doi.org/10.1039/c9ra06869d
9. Werpy T, Petersen G. 2004. Top value added chemicals from biomass: volume I -- results of screening for potential candidates from sugars and synthesis gas. DOE/GO-102004-1992. National Renewable Energy Lab. https://doi.org/10.2172/15008859

10. Liu J, Li J, Shin H-D, Liu L, Du G, Chen J. 2017. Protein and metabolic engineering for the production of organic acids. Bioresour Technol 239:412–421. https://doi.org/10.1016/j.biortech.2017.04.052

11. Liu L, Redden H, Alper HS. 2013. Frontiers of yeast metabolic engineering: diversifying beyond ethanol and saccharomyces. Curr Opin Biotechnol 24:1023–1030. https://doi.org/10.1016/j.copbio.2013.03.005

12. Seo S-O, Park S-K, Jung S-C, Ryu C-M, Kim J-S. 2020. Anti-contamination strategies for yeast fermentations. Microorganisms 8:274. https://doi.org/10.3390/microorganisms8020274

13. Thorwall S, Schwartz C, Chartron JW, Wheeldon I. 2020. Stress-tolerant non-conventional microbes enable next-generation chemical biosynthesis. Nat Chem Biol 16:113–121. https://doi.org/10.1038/s41589-019-0452-x

14. Douglass AP, Offei B, Braun-Galleani S, Coughlan AY, Martos AAR, Ortiz-Merino RA, Byrne KP, Wolfe KH. 2018. Population genomics shows no distinction between pathogenic *Candida krusei* and environmental *Pichia kudriavzevii:* one species, four names. PLoS Pathog 14:e1007138. https://doi.org/10.1371/journal.ppat.1007138

15. Kwon Y-J, Ma A-Z, Li Q, Wang F, Zhuang G-Q, Liu C-Z. 2011. Effect of lignocellulosic inhibitory compounds on growth and ethanol fermentation of newly-isolated thermotolerant *Issatchenkia orientalis*. Bioresour Technol 102:8099–8104. https://doi.org/10.1016/j.biortech.2011.06.035

16. Seong Y-J, Lee H-J, Lee J-E, Kim S, Lee DY, Kim KH, Park Y-C. 2017. Physiological and metabolomic analysis of *Issatchenkia orientalis* MTY1 with multiple tolerance for cellulosic bioethanol production. Biotechnol J 12. https://doi.org/10.1002/biot.201700110

17. Xiao H, Shao Z, Jiang Y, Dole S, Zhao H. 2014. Exploiting *Issatchenkia orientalis* SD108 for succinic acid production. Microb Cell Fact 13:121. https://doi.org/10.1186/s12934-014-0121-4

18. Ruyters S, Mukherjee V, Verstrepen KJ, Thevelein JM, Willems KA, Lievens B. 2015. Assessing the potential of wild yeasts for bioethanol production. J Ind Microbiol Biotechnol 42:39–48. https://doi.org/10.1007/s10295-014-1544-y

19. Oberoi HS, Babbar N, Sandhu SK, Dhaliwal SS, Kaur U, Chadha BS, Bhargav VK. 2012. Ethanol production from alkali-treated rice straw via simultaneous saccharification and fermentation using newly isolated thermotolerant *Pichia kudriavzevii* HOP-1. J Ind Microbiol Biotechnol 39:557–566. https://doi.org/10.1007/s10295-011-1060-2

20. Isono N, Hayakawa H, Usami A, Mishima T, Hisamatsu M. 2012. A comparative study of ethanol production by *Issatchenkia orientalis* strains under stress conditions. J Biosci Bioeng 113:76–78. https://doi.org/10.1016/j.jbiosc.2011.09.004

21. Gallardo JCM, Souza CS, Cicarelli RMB, Oliveira KF, Morais MR, Laluce C. 2011. Enrichment of a continuous culture of *Saccharomyces cerevisiae* with the yeast *Issatchenkia orientalis* in the production of ethanol at increasing temperatures. J Ind Microbiol Biotechnol 38:405–414. https://doi.org/10.1007/s10295-010-0783-9

22. Kitagawa T, Tokuhiro K, Sugiyama H, Kohda K, Isono N, Hisamatsu M, Takahashi H, Imaeda T. 2010. Construction of a β-glucosidase expression system using the multistress-tolerant yeast *Issatchenkia orientalis*. Appl Microbiol Biotechnol 87:1841–1853. https://doi.org/10.1007/s00253-010-2629-9

23. Shabala L, Ross T. 2008. Cyclopropane fatty acids improve *Escherichia coli* survival in acidified minimal media by reducing membrane permeability to H+ and enhanced ability to extrude H+. Res Microbiol 159:458–461. https://doi.org/10.1016/j.resmic.2008.04.011

24. Sohlenkamp C. 2017. Biogenesis of fatty acids. Lipids and Membranes 1–13. https://doi.org/10.1007/978-3-319-43676-0

25. Baker-Austin C, Dopson M. 2007. Life in acid: pH homeostasis in acidophiles. Trends Microbiol 15:165–171. https://doi.org/10.1016/j.tim.2007.02.005

26. Mira NP, Teixeira MC, Sá-Correia I. 2010. Adaptive response and tolerance to weak acids in *Saccharomyces cerevisiae*: a genome-wide view. OMICS 14:525–540. https://doi.org/10.1089/omi.2010.0072

27. Krulwich TA, Sachs G, Padan E. 2011. Molecular aspects of bacterial pH sensing and homeostasis. Nat Rev Microbiol 9:330–343. https://doi.org/10.1038/nrmicro2549

28. He A, Penix SR, Basting PJ, Griffith JM, Creamer KE, Camperchioli D, Clark MW, Gonzales AS, Chávez Erazo JS, George NS, Bhagwat AA, Slonczewski JL, Elliot MA. 2017. Acid evolution of *Escherichia coli* K-12 eliminates amino acid decarboxylases and reregulates catabolism. Appl Environ Microbiol 83:e00442-17. https://doi.org/10.1128/AEM.00442-17

29. Shabayek S, Spellerberg B. 2017. Acid stress response mechanisms of group B streptococci. Front Cell Infect Microbiol 7:395. https://doi.org/10.3389/fcimb.2017.00395

30. Pennacchietti E, D'Alonzo C, Freddi L, Occhialini A, De Biase D. 2018. The Glutaminase-dependent acid resistance system: qualitative and quantitative assays and analysis of its distribution in enteric bacteria. Front Microbiol 9:2869. https://doi.org/10.3389/fmicb.2018.02869

31. Miller EF, Maier RJ. 2014. Ammonium metabolism enzymes aid helicobacter pylori acid resistance. J Bacteriol 196:3074–3081. https://doi.org/10.1128/JB.01423-13

32. Hong W, Wu YE, Fu X, Chang Z. 2012. Chaperone-dependent mechanisms for acid resistance in enteric bacteria. Trends Microbiol 20:328–335. https://doi.org/10.1016/j.tim.2012.03.001

33. Voth W, Jakob U. 2017. Stress-activated chaperones: a first line of defense. Trends Biochem Sci 42:899–913. https://doi.org/10.1016/j.tibs.2017.08.006

34. Guan N, Liu L. 2020. Microbial response to acid stress: mechanisms and applications. Appl Microbiol Biotechnol 104:51–65. https://doi.org/10.1007/s00253-019-10226-1

35. Sasaki Y, PingHH, ZiaF. 2023. A population phenomics study of *Issatchenkia orientalis* Illuminates the possibility of creating super chassis for next-generation biotechnology. in preparation

36. Hackett SR, Baltz EA, Coram M, Wranik BJ, Kim G, Baker A, Fan M, Hendrickson DG, Berndl M, McIsaac RS. 2020. Learning causal networks using inducible transcription factors and transcriptome-wide time series. Mol Syst Biol 16:e9174. https://doi.org/10.15252/msb.20199174

37. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915. https://doi.org/10.1038/s41587-019-0201-4

38. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. https://doi.org/10.1093/bioinformatics/btp616

39. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. 2010. Inferring regulatory networks from expression data using tree-based methods. PLoS One 5:e12776. https://doi.org/10.1371/journal.pone.0012776

40. He X, Samee MAH, Blatti C, Sinha S. 2010. Thermodynamics-based models of transcriptional regulation by enhancers: the roles of synergistic activation, cooperative binding and short-range repression. PLoS Comput Biol 6:e1000935. https://doi.org/10.1371/journal.pcbi.1000935

41. Monteiro PT, Oliveira J, Pais P, Antunes M, Palma M, Cavalheiro M, Galocha M, Godinho CP, Martins LC, Bourbon N, Mota MN, Ribeiro RA, Viana R, Sá-Correia I, Teixeira MC. 2020. YEASTRACT+: a portal for cross-species comparative genomics of transcription regulation in yeasts. Nucleic Acids Res 48:D642–D649. https://doi.org/10.1093/nar/gkz859

42. He X, Duque T, Sinha S. 2012. Evolutionary origins of transcription factor binding site clusters. Mol Biol Evol 29:1059–1070. https://doi.org/10.1093/molbev/msr277

43. Xu X, Chen J, Xu H, Li D. 2014. Role of a major facilitator superfamily transporter in adaptation capacity of penicillium funiculosum under extreme acidic stress. Fungal Genet Biol 69:75–83. https://doi.org/10.1016/j.fgb.2014.06.002

44. Ji H, Xu K, Dong X, Sun D, Peng R, Lin S, Zhang K, Jin L. 2020. Transcriptional profiling reveals molecular basis and the role of arginine in response to low-pH stress in *Pichia kudriavzevii*. J Biosci Bioeng 130:588–595. https://doi.org/10.1016/j.jbiosc.2020.07.014

45. Guo Z-P, Olsson L. 2016. Physiological responses to acid stress by *Saccharomyces cerevisiae* when applying high initial cell density. FEMS Yeast Res 16:fow072. https://doi.org/10.1093/femsyr/fow072

46. Guan N, Shin H, Chen RR, Li J, Liu L, Du G, Chen J. 2014. Understanding of how propionibacterium acidipropionici respond to propionic acid stress at the level of proteomics. Sci Rep 4:6951. https://doi.org/10.1038/srep06951

47. Carmelo V, Santos H, Sá-Correia I. 1997. Effect of extracellular acidification on the activity of plasma membrane ATPase and on the cytosolic and vacuolar pH of *Saccharomyces cerevisiae*. Biochim Biophys Acta 1325:63–70. https://doi.org/10.1016/s0005-2736(96)00245-3

48. Steffen KK, MacKay VL, Kerr EO, Tsuchiya M, Hu D, Fox LA, Dang N, Johnston ED, Oakes JA, Tchao BN, Pak DN, Fields S, Kennedy BK, Kaeberlein M. 2008. Yeast life span extension by depletion of 60s ribosomal subunits is mediated by Gcn4. Cell 133:292–302. https://doi.org/10.1016/j.cell.2008.02.037

49. Torello Pianale L, Rugbjerg P, Olsson L. 2021. Real-time monitoring of the yeast intracellular state during bioprocesses with a toolbox of biosensors. Front Microbiol 12:802169. https://doi.org/10.3389/fmicb.2021.802169

50. Johnston NR, Nallur S, Gordon PB, Smith KD, Strobel SA. 2020. Genome-wide identification of genes involved in general acid stress and fluoride toxicity in Saccharomyces cerevisiae. Front Microbiol 11:1410. https://doi.org/10.3389/fmicb.2020.01410

51. Kim J, Cunningham R, James B, Wyder S, Gibson JD, Niehuis O, Zdobnov EM, Robertson HM, Robinson GE, Werren JH, Sinha S. 2010. Functional characterization of transcription factor motifs using cross-species comparison across large evolutionary distances. PLoS Comput Biol 6:e1000652. https://doi.org/10.1371/journal.pcbi.1000652

52. Martínez-Montañés F, Rienzo A, Poveda-Huertes D, Pascual-Ahuir A, Proft M. 2013. Activator and repressor functions of the Mot3 transcription factor in the osmostress response of Saccharomyces cerevisiae. Eukaryot Cell 12:636–647. https://doi.org/10.1128/EC.00037-13

53. Legras JL, Erny C, Le Jeune C, Lollier M, Adolphe Y, Demuyter C, Delobel P, Blondin B, Karst F. 2010. Activation of two different resistance mechanisms in Saccharomyces cerevisiae upon exposure to octanoic and decanoic acids. Appl Environ Microbiol 76:7526–7535. https://doi.org/10.1128/AEM.01280-10

54. Sousa MJ, Ludovico P, Rodrigues F, Leo H, Côrte-Real C. 2012.Cell metabolism - cell homeostasis and stress response. PLoS One 7:e52402. https://doi.org/10.5772/27726

55. Yonkovich J, McKenndry R, Shi X, Zhu Z. 2002. Copper ion-sensing transcription factor Mac1P post-translationally controls the degradation of its target gene product Ctr1P. J Biol Chem 277:23981–23984. https://doi.org/10.1074/jbc.C200203200

56. Antunes M, Palma M, Sá-Correia I. 2018. Transcriptional profiling of Zygosaccharomyces bailii early response to acetic acid or copper stress mediated by ZbHaa1. Sci Rep 8:14122. https://doi.org/10.1038/s41598-018-32266-9

57. Rajkumar AS, Liu G, Bergenholm D, Arsovska D, Kristensen M, Nielsen J, Jensen MK, Keasling JD. 2016. Engineering of synthetic, stress-responsive yeast promoters. Nucleic Acids Res 44:e136. https://doi.org/10.1093/nar/gkw553

58. Laera L, Guaragnella N, Ždralević M, Marzulli D, Liu Z, Giannattasio S. 2016. The transcription factors ADR1 or CAT8 are required for RTG pathway activation and evasion from yeast acetic acid-induced programmed cell death in raffinose. Microb Cell 3:621–631. https://doi.org/10.15698/mic2016.12.549

59. Fernandes AR, Mira NP, Vargas RC, Canelhas I, Sá-Correia I. 2005. Saccharomyces cerevisiae adaptation to weak acids involves the transcription factor Haa1P and Haa1P-regulated genes. Biochem Biophys Res Commun 337:95–103. https://doi.org/10.1016/j.bbrc.2005.09.010

60. Schlitt T, Brazma A. 2007. Current approaches to gene regulatory network modelling. BMC Bioinformatics 8:S9. https://doi.org/10.1186/1471-2105-8-S6-S9

61. Święciło A. 2016. Cross-stress resistance in Saccharomyces cerevisiae yeast—new insight into an old phenomenon. Cell Stress Chaperones 21:187–200. https://doi.org/10.1007/s12192-016-0667-7

62. Mira NP, Becker JD, Sá-Correia I. 2010. Genomic expression program involving the Haa1P-regulon in Saccharomyces cerevisiae response to acetic acid. OMICS 14:587–601. https://doi.org/10.1089/omi.2010.0048

63. Zara G, van Vuuren HJJ, Mannazzu I, Zara S, Budroni M. 2019. Transcriptomic response of Saccharomyces cerevisiae during fermentation under oleic acid and ergosterol depletion. Fermentation 5:57. https://doi.org/10.3390/fermentation5030057

64. Kawahata M, Masaki K, Fujii T, Iefuji H. 2006. Yeast genes involved in response to lactic acid and acetic acid: acidic conditions caused by the organic acids in Saccharomyces cerevisiae cultures induce expression of intracellular metal metabolism genes regulated by Aft1p. FEMS Yeast Res 6:924–936. https://doi.org/10.1111/j.1567-1364.2006.00089.x

65. Lucena RM, Dolz-Edo L, Brul S, de Morais MA, Smits G. 2020. Extreme low cytosolic pH is a signal for cell survival in acid stressed yeast. Genes -basel 11:656. https://doi.org/10.3390/genes11060656

66. Teixeira MC, Fernandes AR, Mira NP, Becker JD, Sá-Correia I. 2006. Early transcriptional response of Saccharomyces cerevisiae to stress imposed by the Herbicide 2,4-dichlorophenoxyacetic acid. FEMS Yeast Res 6:230–248. https://doi.org/10.1111/j.1567-1364.2006.00041.x

67. Cottier F, Tan ASM, Yurieva M, Liao W, Lum J, Poidinger M, Zolezzi F, Pavelka N. 2017. The transcriptional response of Candida albicans to weak organic acids, carbon source, and MIG1 inactivation unveils a role for HGT16 in mediating the fungistatic effect of acetic acid. G3 Genes Genomes Genetics 7:3597–3604. https://doi.org/10.1534/g3.117.300238

68. Walkey CJ, Luo Z, Madilao LL, van Vuuren HJJ. 2012. The fermentation stress response protein Aaf1P/Yml081Wp regulates acetate production in Saccharomyces cerevisiae. PLoS One 7:e51551. https://doi.org/10.1371/journal.pone.0051551

69. Wang R-M, Li N, Zheng K, Hao J-F. 2018. Enhancing acid tolerance of the probiotic bacterium lactobacillus acidophilus NCFM with trehalose. FEMS Microbiol Lett. https://doi.org/10.1093/femsec/fny217

70. Guaragnella N, Stirpe M, Marzulli D, Mazzoni C, Giannattasio S. 2019. Acid stress triggers resistance to acetic acid-induced regulated cell death through Hog1 activation which requires Rtg2 in yeast. Oxid Med Cell Longev 2019:4651062. https://doi.org/10.1155/2019/4651062

71. Matsushika A, Negi K, Suzuki T, Goshima T, Hoshino T. 2016. Identification and characterization of a novel Issatchenkia orientalis GPI-anchored protein, Iogas1, required for resistance to low pH and salt stress. PLoS One 11:e0161888. https://doi.org/10.1371/journal.pone.0161888

72. Wada K, Fujii T, Akita H, Matsushika A. 2020. Iogas1, a GPI-anchored protein derived from Issatchenkia orientalis, confers tolerance of Saccharomyces cerevisiae to multiple acids. Appl Biochem Biotechnol 190:1349–1359. https://doi.org/10.1007/s12010-019-03187-8

73. Passemiers A, Moreau Y, Raimondi D. 2022. Fast and accurate inference of gene regulatory networks through robust precision matrix estimation. Bioinformatics 38:2802–2809. https://doi.org/10.1093/bioinformatics/btac178

74. Kim J, He X, Sinha S. 2009. Evolution of regulatory sequences in 12 drosophila species. PLoS Genet 5:e1000330. https://doi.org/10.1371/journal.pgen.1000330