Transcriptional Regulation and the Diversification of Metabolism in Wine Yeast Strains

Debra Rossouw, Dan Jacobson, and Florian F. Bauer[1]
Institute for Wine Biotechnology, University of Stellenbosch, Stellenbosch, 7601 South Africa

ABSTRACT Transcription factors and their binding sites have been proposed as primary targets of evolutionary adaptation because changes to single transcription factors can lead to far-reaching changes in gene expression patterns. Nevertheless, there is very little concrete evidence for such evolutionary changes. Industrial wine yeast strains, of the species *Saccharomyces cerevisiae*, are a geno- and phenotypically diverse group of organisms that have adapted to the ecological niches of industrial winemaking environments and have been selected to produce specific styles of wine. Variation in transcriptional regulation among wine yeast strains may be responsible for many of the observed differences and specific adaptations to different fermentative conditions in the context of commercial winemaking. We analyzed gene expression profiles of wine yeast strains to assess the impact of transcription factor expression on metabolic networks. The data provide new insights into the molecular basis of variations in gene expression in industrial strains and their consequent effects on metabolic networks important to wine fermentation. We show that the metabolic phenotype of a strain can be shifted in a relatively predictable manner by changing expression levels of individual transcription factors, opening opportunities to modify transcription networks to achieve desirable outcomes.

SACCHAROMYCES cerevisiae is the yeast species most widely used in the fermentation industry (oenology, bread making, and brewing). Most genetic studies of *S. cerevisiae* have been carried out using a handful of strains (Mortimer *et al.* 1957; Mortimer and Johnston 1986) that were selected for their ease of use under laboratory conditions.

By contrast, industrial yeast strains are geno- and phenotypically highly diverse (Frezier and Dubourdieu 1992; Schütz and Gafner1994; Rossouw *et al.* 2009), having adapted to the ecological niches provided by industrial or semi-industrial environments. In the wine industry a large number of such strains are commercially produced, most of which were originally isolated from spontaneous wine fermentations (Johnston *et al.* 2000). Although the original or natural ecological niche of *S. cerevisiae* is subject to conjecture, industrial environments have undoubtedly sculpted the recent evolution of the strains currently used in industry, offering an excellent opportunity for comparative studies to investigate evolutionary relationships and the molecular mechanisms underlying phenotypic differentiation.

Wine yeast strains were primarily selected for their ability to completely ferment (to ferment to dryness) very high levels (>200 g/liter) of sugars in a largely anaerobic environment. Beyond this fundamental trait, strains have been selected for specific and diverse purposes, for example to support the production of different styles of wine or to produce different aroma profiles. These strains therefore represent a wide range of phenotypic traits, which is a reflection of significant genetic diversity.

A number of studies have focused on evolutionary adaptations of wine yeast strains. It has been suggested that the diploid status of most wine yeast strains may confer an advantage in terms of rapid adaptation to variable external environments and provide a way to increase the dosage of genes important for fermentation (Bakalinsky and Snow 1990; Salmon 1997). Furthermore, subtelomeric chromosomal regions are subject to duplications and rearrangements via ectopic exchanges (Bidenne *et al.* 1992; Rachidi *et al.* 1999). Another reported mode of evolution of *Saccharomyces* is the formation of interspecific hybrids. The resulting genome plasticity promotes faster adaptation

doi: 10.1534/genetics.111.132720
Manuscript received July 20, 2011; accepted for publication October 23, 2011
Supporting information is available online at http://www.genetics.org/content/suppl/2011/10/31/genetics.111.132720.DC1.
Sequence data from this article have been deposited with the NCBI/GEO Data Library under accession nos. GSE11651 and GSE26929.
[1]Corresponding author: Institute for Wine Biotechnology, Faculty of AgriSciences, University of Stellenbosch, Private Bag X1, Matieland, 7602 South Africa. E-mail: fb2@sun.ac.za

in response to environmental changes (Puig and Perez-Ortin 2000; Libkind *et al.* 2011) by providing the genetic diversity upon which natural selection operates.

Adaptations of these strains to the specific oenological environment and their selection for specific biotechnological purposes are also reflected in global transcriptomic, proteomic, and metabolomic profiles. Studies of wine yeast strains have correlated differences in fermentation phenotypes to gene expression, protein levels, and metabolic regulation (Rossouw *et al.* 2008, 2009, 2010). These studies focused on the aroma-relevant exometabolome as produced by different wine yeast strains, since this metabolome largely determines the aromatic perception of fruitiness and complexity of wines, and is therefore of particular interest to winemakers.

It has been proposed that some of the primary evolutionary targets of strain diversification are transcription factors and their binding sites (Dermitzakis and Clark 2002). Data show that although *S. cerevisiae* and *S. mikatae* have similar genome sequences, they are significantly different in their transcription-factor binding profiles (Borneman *et al.*, 2007a,b). It has been hypothesized that the extensive binding site differences observed between the different species reflect rapid specialization of *Saccharomyces* for distinct ecological environments (Borneman *et al.* 2007a,b).

For this study, the production of volatile aroma compounds was correlated to previously established transcriptional profiles of five different wine yeast strains under simulated winemaking conditions. We were able to identify transcription factors (TFs) whose expression profiles may contribute to the different metabolism-related phenotypes observed in different strains. In particular, we assessed whether the metabolic phenotype of one strain could be engineered to more closely resemble that of another strain by adjusting the expression of key transcription factors. This would support the hypothesis that changes in expression of specific transcription factors are responsible for the evolutionary adaptation of different *Saccharomyces* strains. The identification of such key TFs promises targeted improvement of fermentation performance (Hou *et al.* 2009).

Methods

Strains, media, and culture conditions

The yeast strains used in this study are listed in Table1. All are diploid *Saccharomyces cerevisiae* strains used in industrial wine fermentations. Yeast cells (inoculated from single, characterized colonies) were cultivated at 30° in YPD synthetic media 1% yeast extract (Biolab, South Africa), 2% peptone (Fluka, Germany), 2% glucose (Sigma, Germany). Solid medium was supplemented with 2% agar (Biolab).

Fermentation media

Fermentation experiments were carried out with synthetic must MS300, which approximates a natural must as previously described (Bely *et al.* 1990). The medium contained 125 g/liter glucose and 125 g/liter fructose, and the pH was adjusted to 3.3 with NaOH.

Table 1 Strains used in this study

Strain	Source
VIN13	Anchor Yeast, South Africa
BM45	Lallemand, Inc., Montréal, Canada
DV10	Lallemand, Inc., Montréal, Canada
SOK2-VIN13	This study
RAP1-VIN13	This study

Fermentation conditions

All fermentations were carried out under microaerobic conditions in 100-ml glass bottles (containing 80 ml of the medium) sealed with rubber stoppers with a CO_2 outlet. All fermentations were carried out in triplicate, *i.e.*, independent biological repeats. The fermentation temperature was approximately 22° and no continuous stirring was performed during the course of the fermentation. Fermentation bottles were inoculated with YPD cultures in the logarithmic growth phase (around OD_{600} = 1) to an OD_{600} of 0.1 (*i.e.*, a final cell density of approximately 10^6 cfu/ml). The cells from the YPD precultures were briefly centrifuged and resuspended in MS300 to avoid carryover of YPD to the fermentation media. The fermentations followed a time course of 14 days and the bottles were weighed daily to assess the progress of fermentation. Samples of the fermentation media and cells were taken at days 2, 5, and 14 as representative of the exponential, early logarithmic, and late logarithmic growth phases, respectively.

Growth measurement

Cell proliferation (*i.e.*, growth) was determined spectrophotometrically ($Powerwave_X$, Bio-Tek Instruments) by measuring the optical density (at 600 nm) of 200-μl samples of the suspensions over the 14-day experimental period.

Analytical methods High-performance liquid chromatography (HPLC):

Culture supernatants were obtained from the cell-free upper layers of the fermentation media. For the purposes of glucose determination and carbon recovery, culture supernatants and starting media were analyzed by HPLC on an AMINEX HPX-87H ion exchange column using 5 mM H_2SO_4 as the mobile phase. Agilent RID and UV detectors were used in tandem for peak detection and quantification. Analysis was carried out using the HPChemstation software package.

Gas chromatograph–flame ionization detector (GC-FID): Each 5-ml sample of synthetic must taken during fermentation was spiked with an internal standard of 4-methyl-2-pentanol to a final concentration of 10 mg/liter. To each of these samples 1 ml of solvent (diethyl ether) was added and the tubes sonicated for 5 min. The top layer in each tube was

separated by centrifugation at 3000 rpm for 5 min and the extract analyzed. Three microliters of each sample was injected into the GC. All extractions were done in triplicate.

The analysis of volatile compounds was carried out on a Hewlett Packard 5890 Series II GC coupled to an HP 7673 auto-sampler and injector and an HP 3396A integrator. The column used was a Lab Alliance organic-coated, fused silica capillary with dimensions of 60 m × 0.32 mm internal diameter with a 0.5-μm coating thickness. The injector temperature was set to 200°, the split ratio to 20:1, and the flow rate to 15 ml/min, with hydrogen used as the carrier gas for a flame ionization detector held at 250°. The oven temperature was increased from 35° to 230° at a ramp of 3°/min.

Internal standards (Merck, Cape Town) were used to calibrate the machine for each of the compounds measured.

Microarray analysis: Sampling of cells from fermentation and total RNA extraction was performed as described by Abbott *et al.* (2007). Samples were taken from independent fermentations in triplicate on days 2, 5, and 14. For a complete description of the hybridization conditions refer to Rossouw *et al.* (2008). Transcript data can be downloaded from the GEO repository under the following accession numbers: GSE11651 (for the original VIN13, BM45, EC1118, 285, and DV10 data sets analyzed in Rossouw *et al.* 2009) and GSE26929 (for the SOK2-overexpressing strain and VIN13 control data sets).

Transcriptomics data analysis: The microarray data were background corrected and normalized with robust multichip average (Irizarry *et al.* 2003) and the resultant $\log^2$ transformed data were mean centered for each probe set. Determination of differential gene expression between experimental parameters was conducted using SAM (significance analysis of microarrays) version 2 (Tusher *et al.* 2001). The two-class, unpaired setting was used and genes with a Q-value <0.5 ($P < 0.0005$) and a fold change greater than 2 (positive or negative) were taken into consideration as differentially expressed genes.

The sequences for each of the individual probes of the Affymetrix Yeast 2.0 Genechip were mapped to the yeast genome by the use of blastn (Altschul *et al.* 1990). A Perl program was written to perform the following tasks: (1) 100% identity matches (over the full length of the probe) were extracted from the blastn results; (2) the probes were subsequently assembled into probe sets and the resultant probe set to gene relationships modeled as a graph; (3) ambiguous probe sets, *i.e.*, those that were found to map to more than one gene (node degree >1), were removed from the input gene list for the subsequent random forest analyses.

Random forest analysis (Breiman 2001) was carried out on the normalized and mean centered expression data by the use of the randomForest R package (Liaw and Wiener 2002). A random forest classification model was created using the strains as classes, regardless of time point. Fifteen thousand trees were generated in the creation of the model with 73 randomly selected variables (probe sets) used at each split. The out-of-bag (OOB) estimate of error rate was 4.65%. The mean decrease of accuracy measure of variable importance was extracted from the random forest model and used to rank the contribution of all probe sets according to their ability to discriminate between different strains. The probe sets occurring within the 200 most important variables from the random forest model described above were selected for further in depth analysis and evaluation.

Gene expression profiles were clustered using the short time series expression miner (STEM; Ernst and Bar-Joseph 2006).

Multivariate data analysis: The patterns within the different sets of data were investigated by principal-component analysis (PCA; Qlucore Omics Explorer v. 2.2). PCA is a bilinear modeling method, which gives a visually interpretable overview of the main information in large, multidimensional data sets. By plotting the principal components it is possible to view statistical relationships between different variables in complex data sets and detect and interpret sample groupings, similarities, or differences, as well as the relationships between the different variables (Mardia *et al.* 1979).

Univariate statistics and visualization: The levels of aroma compounds from target strains and transcription factor overexpression strains were compared to their respective control strains and the statistical significance of the changes evaluated with a *t*-test at a 95% confidence interval. To better visualize the statistical relationships in the data set the following algorithm was implemented in Perl: a mathematical graph was created with a node for each control strain (VIN13 or BM45). Subsequently, those compounds that showed a statistically significant difference from the control in either the target or the overexpression strain were added as a node to the graph and an edge created to the control strain node. For each strain showing a significant change a node was added to the graph and an edge created between it and the previously mentioned compound node. Fold change was calculated for each compound in each strain as a simple ratio between the compound level in the strain and that of its control. If the ratio was less than one its negative reciprocal was taken. This fold change information as well as descriptive information for each node was then written into an annotation file. Cytoscape v. 2.8.1 (Smoot *et al.* 2011) was used to visualize the resulting graph and annotation. The nodes were shaded according to fold change on a red (positive) or blue (negative) color scale.

Overexpression constructs and transformation: The two plasmids constructed for use in this study are pDM-PhR-RAP1 (genotype, 2μ LEU2 $TEF1_P$ PhR322 $TEF1_T$ PGK_P RAP1 PGK_T) and pDM-PhR-SOK2 (genotype, 2μ LEU2 $TEF1_P$ PhR322 $TEF1_T$ PGK_P SOK2 PGK_T). Primers used for amplification of

transcription factor encoding genes are listed in Supporting Information, Table S1. Standard procedures for the isolation of DNA were used throughout this study (Ausubel *et al.* 1994). Standard DNA techniques were also carried out as described by Sambrook *et al.* (1989). All enzymes for cloning, restriction digest, and ligation reactions were obtained from Roche Diagnostics (Randburg, South Africa) and used according to supplier specifications. Sequencing of all plasmids was carried out on an ABI PRISM automated sequencer. All plasmids contain the dominant marker PhR conferring phleomicin resistance (PhR) and were transformed into host VIN13 and BM45 cells via electroporation (Wenzel *et al.* 1992; Lilly *et al.* 2006).

Quantitative real-time PCR analysis (QRT-PCR): RNA extractions from fermenting yeasts were carried out as per the microarray analyses. Primer design for QRT-PCR analysis was performed using the Primer Express software v. 3 (Applied Biosystems) and reagents were purchased from KAPA Biosystems. Spectral data were captured by the 7500 cycler (Applied Biosystems). Data analyses were conducted using Signal Detection Software (SDS) v. 1.3.1. (Applied Biosystems) to determine the corresponding Ct values and PCR efficiencies, respectively, for the samples analyzed (Ramakers *et al.* 2003). The genes selected for QRT-PCR, as well as the primer sequences used for amplification are described in Table S2.

Transcriptomic analysis of overexpressing strains: Fermentations of the *SOK2*-overexpressing strain as well as the VIN13 control were carried out in triplicate in synthetic must as described previously. Samples for transcriptomic analysis were taken from three independent biological repeats at day 2 of fermentation, during the exponential growth phase. The microarray data can be viewed at the GEO repository under the accession number (GSE26929).

Results

Transcription factor enrichment

In our previous work (Rossouw *et al.* 2008), the transcriptome of five distinct industrial wine yeast strains was analyzed at three time points in synthetic wine must fermentations, day 2 (exponential growth phase), day 5 (early stationary growth phase), and day 14 (late stationary-growth phase). Strains were also monitored for sugar utilization and production of ethanol, glycerol, and 32 volatile aroma compounds (Rossouw *et al.* 2008, 2009).

Normalized expression values for the different strains and time points were analyzed by random forest analysis (Breiman 2001), and the top 200 strain discriminatory genes were ranked according to their ability to differentiate between the different strains. These genes were subsequently subjected to transcription factor enrichment as described by Teixeira *et al.* (2006) to identify the main regulatory structures present in the data. Transcription factors that reportedly regulate most of the highly discriminatory genes from the random forest outputs were thus identified and ranked according to the percentage of genes identified by the random forest, which are regulated by these transcription factors. Enrichment of transcription factors was performed on the total set of 200 genes, as well as on a smaller subset of 42 genes from the random forest output, which are thought to be involved in metabolism based on GO functional annotations. From Table 2 it is clear that a few key transcription factors may account for the majority of genes responsible for the differential transcriptional response between strains.

Table 2 Top 10 hits for transcription factor enrichment analysis of random forest outputs (% of total) for strain discriminatory genes in the total gene list and in the metabolism-specific subset

All genes	%	Metabolic genes	%
Ste12	39	Ste12p	45
Sfp1p	36	Sok2p	29
Yap1p	33	Rap1p	26
Rap1p	26	Yap6p	24
Aft1p	26	Cin5p	24
Sok2p	25	Phd1p	19
Msn2p	24	Skn7p	17
Met4p	24	Tec1p	17
Msn4p	18	Nrg1p	17
Rpn4p	18	Ino4p	17

Most of the identified transcription factors are involved in the synchronization of stress responses, the regulation of carbon utilization and the modulation of cell membrane and cell wall properties. Genes in these categories can be directly linked to the major changes that yeast experience during fermentation and presumably also reflect the evolutionary framework of domesticated strains.

The transcriptome data were screened to identify the transcription factors in Table 2 that showed differences in either expression level and/or expression pattern between different strains over time. Some of the TF genes did show significant differences in expression levels between one or more strains at particular time points, but overall expression trends and patterns over time were similar. Importantly, six of the transcription factor-encoding genes and notably some of the top-scoring candidates of the TF enrichment, namely *YAP1*, *YAP6*, *SOK2*, *PHD1*, *STE12*, and *RAP1*, did show significant differences between strains in terms of relative transcript abundance and expression patterns over time (Figure 1).

Interestingly, strains with similar physiological properties regarding metabolite profiles and cell wall properties as described in Rossouw *et al.* (2009) (*e.g.*., EC1118 and DV10, as well as BM45 and 285) also presented similar profiles regarding the expression patterns of these six transcriptional regulators. These transcription factors play important roles in cellular metabolism and regulation, although their specific functions are not fully characterized, and information regarding regulatory networks and specific targets is limited.

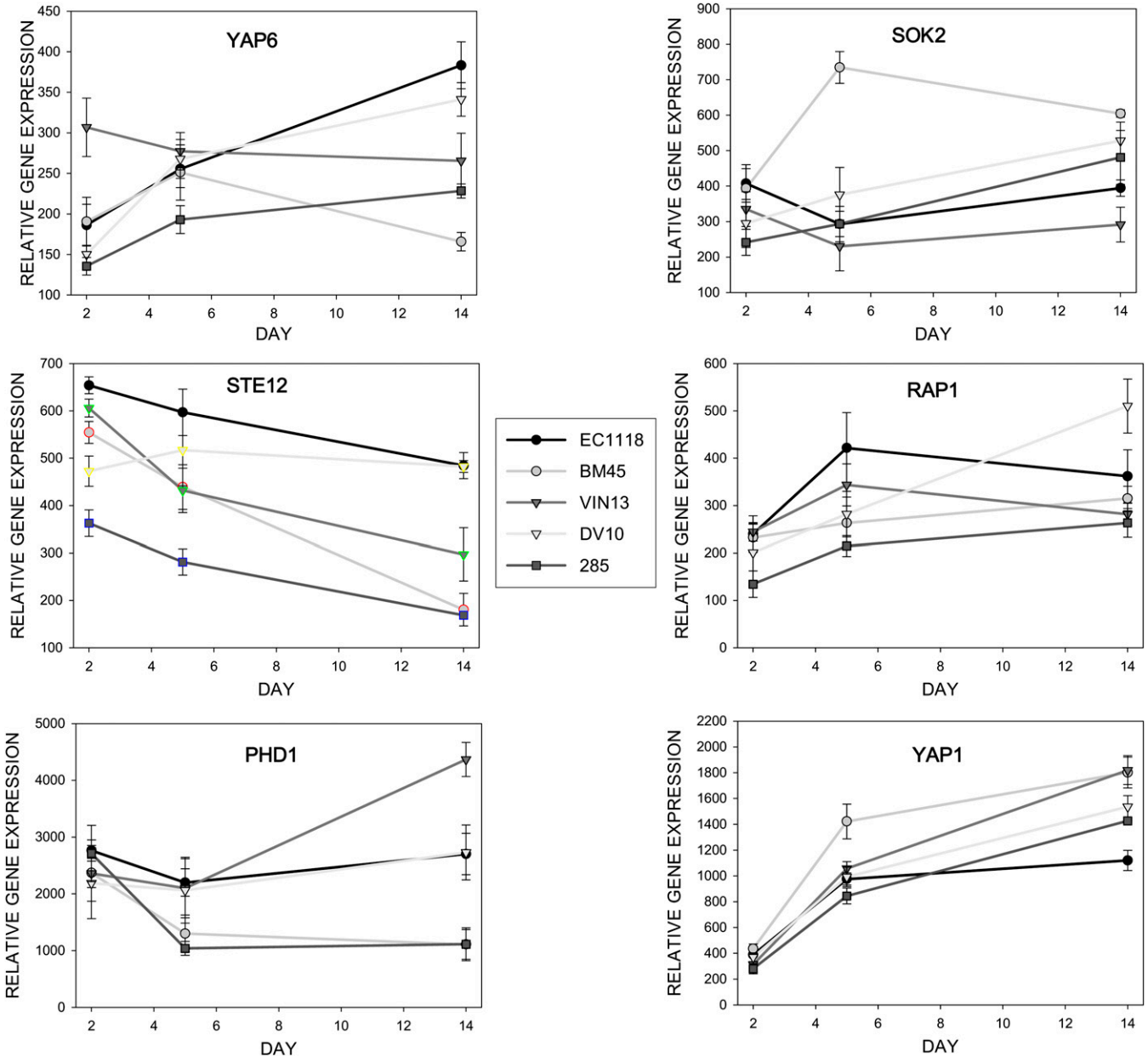

Figure 1 Expression patterns of six genes encoding key transcription factors based on transcription factor enrichment of 200 top-scoring strain-discriminatory genes from random forest analysis. The expression values are derived from microarray experiments and are the average of three biological repeats ±SD.

Yap1p is induced in response to oxidative stress conditions (Okazaki *et al.* 2007) and is believed to regulate the expression of several genes involved in protein mannosylation as well as the invasive growth response (Haugen *et al.* 2004; Thorsen *et al.* 2007). Yap6p is involved in a variety of stress-related programs, including the response to DNA damage and oxidative, osmotic, and toxic metal stresses (Tan *et al.* 2008). Three other key transcription factor encoding genes in the enrichment analysis, namely *SOK2*, *PHD1*, and *STE12*, show highly variable expression patterns between strains (Figure 1). Their protein products are all involved in pseudohyphal growth and regulation of key mannoproteins such as Flo11p (Gimeno and Fink 1994; Pan and Heitman 2000), as well as a host of other metabolic processes. Finally, Rap1p is a multipurpose DNA-binding protein that functions in transcriptional activation, silencing, and replication in yeast. Genes containing Rap1p binding sites include genes encoding proteins involved in amino acid biosynthesis and regulation of carbon metabolism (Yarragudi *et al.* 2007).

Overexpression of selected transcription factors

To determine whether the different expression patterns of these key regulators could be reconciled with the metabolic and phenotypic differences observed between the strains,

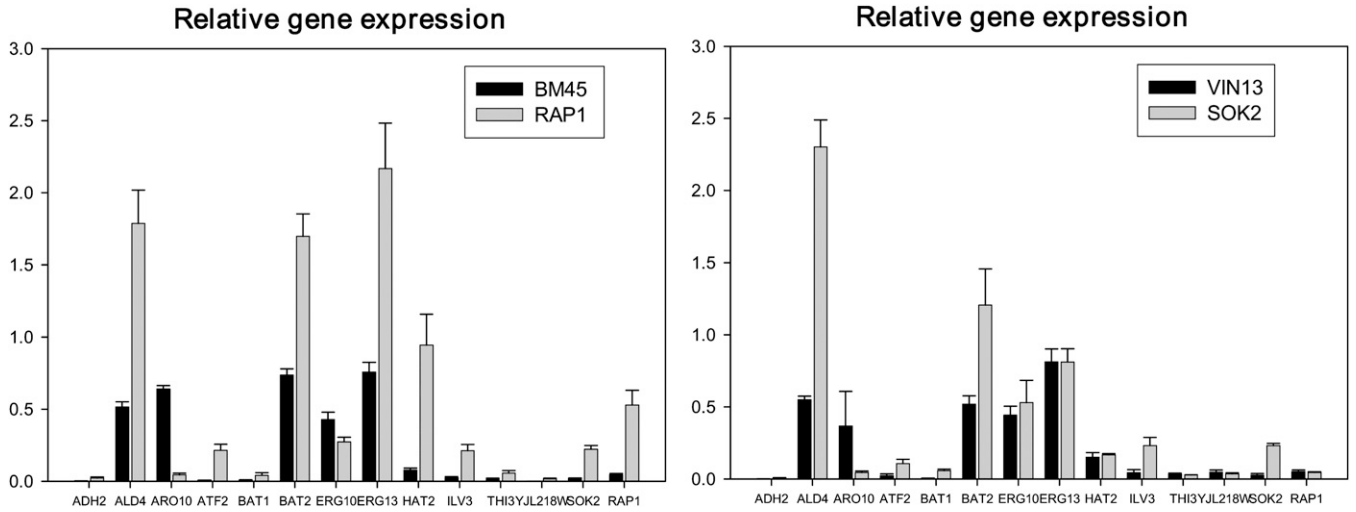

Figure 2 Relative gene expression (normalized to PDA1 expression) of RAP1, SOK2, and selected target genes. Values are the average of three biological repeats ±SD.

we selected two of these genes, namely *SOK2* and *RAP1*, for overexpression analysis. The *SOK2* gene was cloned from the BM45 strain and overexpressed in VIN13, while the *RAP1* gene was cloned from DV10 and overexpressed in BM45. Our goal was to elevate the expression levels of these transcription factors in the overexpression strains to more closely match the expression levels observed in the "donor" strains.

Figure 2 clearly shows that the expression levels of *SOK2* and *RAP1* in the transformed strains were successfully and significantly increased in comparison to their respective controls. To assess whether the overexpression of these factors had an impact on genes under their control, several known or suggested target genes of Sok1p and of Rap1p (Table S3) were selected for expression analysis using real-time PCR, while two genes, *ERG10* and *THI3*, were included as negative controls.

Both negative controls (*THI3* and *ERG10*) showed no change in expression for the transformants, while most of the known or suggested target genes of the two transcription factors, such as *ERG13*, *BAT2*, and *ALD4*, were increased in expression (Figure 2). Of these suggested targets, only *ARO10* did not show any increase in both the *RAP1* and *SOK2* overexpression strains. Considering that the identification of target genes in databases is not always based on direct biological evidence (Li *et al.* 2008), these data provide strong evidence that the transformed strains show expression patterns that indeed reflect increased levels of the two transcription factors.

Fermentation properties of the overexpressing strains

The three original strains (DV10, VIN13, and BM45), as well as the two transformants were inoculated into synthetic wine must and the fermentations monitored over the 14-day fermentation period. All fermentations completed to dryness and the levels of ethanol and glycerol production were similar for the two transformed strains and their respective controls (data not shown).

The impact of changes in transcription factor expression levels on the wine aroma-relevant metabolite profile produced by the different strains was assessed. For this purpose, the concentrations of 22 exometabolites were measured at days 2, 5, and 14 of fermentation, in keeping with our original sampling scheme (Rossouw *et al.* 2008). The results are summarized in Table S4, Table S5, Table S6, and Figure 3.

Clearly, significant differences in the production of volatile aroma compounds at all three stages of fermentation when transformed and untransformed parental strains are compared. The differences were most pronounced for the *SOK2* transformant, but significant differences were also evident for the *RAP1*-overexpressing strain. By the end of fermentation, more than half of the aroma compounds measured were present at substantially different concentrations in the *SOK2*-overexpressing strain in comparison to the parental VIN13 strain, similar to the BM45 target strain. In the case of the *RAP1* transformant, four compounds were significantly increased, and two compounds decreased with reference to the control BM45 strain (Table S6 and Figure 3). For certain volatiles (such as propanol and isoamyl alcohol) the increased concentrations observed in the transformed strains exceeds that of the target strains. Levels of overexpression of the transcription factors in our experiments are not controlled in a precise manner and therefore are not identical to the levels in the original target strains. The impact of overexpression on individual metabolite levels is thus likely to differ (being either more or less) from the exact concentrations determined for the original strains.

Importantly, for the *SOK2*-overexpressing strain, most of the specific metabolic changes as shown in Figure 3 can be directly accounted for by the observed differences in gene expression as determined by transcriptomic analysis. Although we did not assess the transcriptional response of

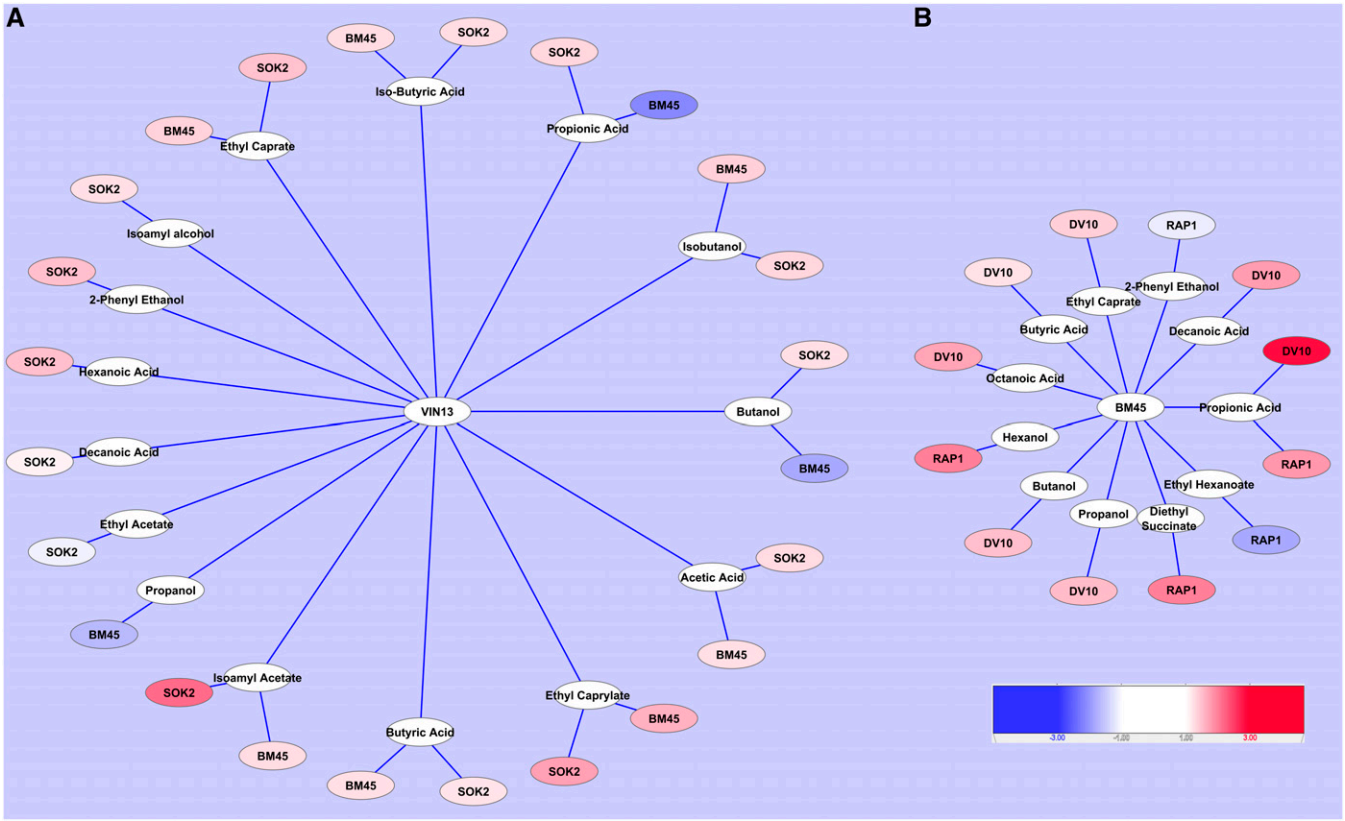

Figure 3 Statistically significant changes in aroma compounds among control, target, and transformed strains on day 14 of fermentation. (A) The levels of aroma compounds from both the target strain (BM45) and the SOK2-overexpression strain that were shown to be statistically significantly different from the control (VIN13). (B) The levels of aroma compounds from both the target strain (DV10) and the RAP1-overexpression strain that were shown to be statistically significantly different from the control (BM45). The degree of fold change is represented by red (positive) and blue (negative) color scales.

the *RAP1*-overexpressing strain, changes in metabolite levels in this strain also correlate well with known targets of Rap1p and the enzymatic activities of these enzymes. For example, one of the target genes, *ERG13* (Kasahara *et al.* 2007), is involved in the production of diethyl succinate, which is present at much higher concentrations at the end of fermentation in the transformed strain compared to the BM45 reference strain (Table S6 and Figure 3).

Transcriptomic analysis of a SOK2-overexpressing strain

Samples from the *SOK2* overexpression fermentations were taken for transcriptomic analysis at day 2 of fermentation, during the exponential growth phase. Close to 1000 transcripts were found to be significantly differentially expressed with a fold change of >2 or <-2. Of these, 258 transcripts were upregulated and 677 downregulated. In terms of alignment with the real-time data, the trends for the 13 transcripts quantified in the real-time analysis were similar to the data derived from the transcriptome analysis, but for *ILV3*, *ALD4*, and *BAT2*, where the significant increases in expression evident in the real-time data (Figure 2) were not reflected in the microarray data. This difference may be explained by different SOK2 expression levels in the two experiments, *i.e.*, a sixfold increase in the real-time experiments *vs.* a twofold increase in the microarray data. It is well established that such differences are commonly seen when 2μ-based multiple copy plasmids are used to amplify gene expression and that many targets of transcription factors are responsive to the precise concentration of the activator (Sauer and Jäckle 1991; Ni *et al.* 2009; Zheng *et al.* 2010).

Of the differentially expressed transcripts (>2- or <-2-fold), 20% were targets of Sok2p as previously described in the literature (Borneman *et al.* 2006; Borneman *et al.* 2007a, b; Horak *et al.* 2002; Lee *et al.* 2002). The remaining 80% of differentially expressed genes may be accounted for by secondary effects of the overexpression or indeed reflect unidentified downstream targets of Sok2p.

When comparing the *SOK2*-overexpressing strain with the VIN13 control, the upregulated genes showed enrichment for the GO categories of metabolism, specifically amino acid metabolism (Table S7). This aligns with the known metabolic regulation of Sok2p. In the case of the downregulated genes, GO processes such as autophagy and energy reserve metabolic processes were the most strongly represented (Table S8).

In the context of the aroma profile changes seen in the transformed strains, gene expression differences in fermentation pathways and pathways related to amino acid metabolism are the most important as amino acids are the precursors for the higher alcohols and esters produced during

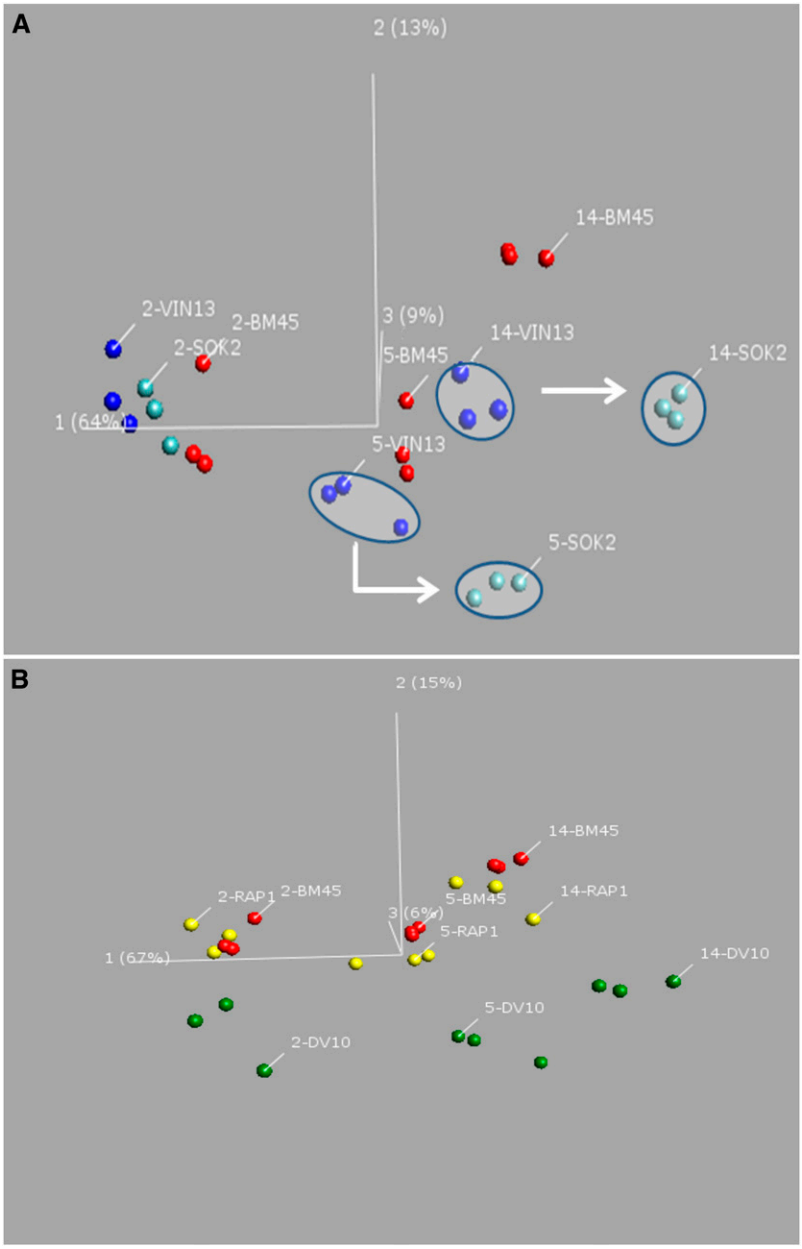

Figure 4 Principal component analysis of aroma compound concentrations in strains overexpressing individual transcription factors as compared to the corresponding untransformed parental as well as to the strain with naturally higher levels of expression of the same transcription factor. (A) A PC1 *vs.* PC2 *vs.* PC3 plot of the VIN13 SOK2-overexpression strain (light blue), the VIN13 control strain (dark blue), and the BM45 target strain (red). Component 1 accounts for 64%, component 2 for 13%, and component 3 for 9% of model variation. (B) The BM45 RAP1-overexpressing strain (yellow), control BM45 strain (red), and target DV10 strain (green) are shown. In this case component 1 accounts for 67% of model variation, component 2 for 15%, and component 3 for 6% of model variation. Samples are labeled according to timepoint (day 2, 5, or 14) and strain.

alcoholic fermentation. Table S9 shows the fold changes for genes in these pathways for fold changes >1.5 or <−1.5. A major increase in expression (fold change >4) is evident for *ATF2* (a known target of Sok2p; Workman *et al.* 2006). The Atf2p enzyme is responsible for the production of a number of volatile esters from their corresponding alcohols, such as ethyl acetate, isoamyl acetate, and phenylethyl acetate (Vestrepen *et al.* 2003). Isoamyl acetate concentrations in the overexpression strain were significantly higher at all time points considered (Table S4, Table S5, and Table S6), corroborating the effect of elevated gene expression on the amount of a metabolite produced. Likewise, activation of *ALD6* by Sok2p (Borneman *et al.* 2006; Chua *et al.* 2006) could explain the increase in acetic acid concentrations (Table S4, Table S5, and Table S6) in the *SOK2*-overexpressing strain as acetate is the direct product of the reaction catalyzed by the aldehyde dehydrogenase isomer encoded by *ALD6* (Saint-Prix *et al.* 2004). Increased expression levels of three *ILV* genes (1, 2, and 5) involved in branched-chain amino acid metabolism (Holmberg and Petersen 1988) may account for the dramatic increase in the two end-products of this pathway, namely isobutanol and isobutyric acid. Increased expression of *ADH5* and *ADH4* in particular also account for the higher concentrations of several higher alcohols and esters (such as 2-phenylethanol) as these enzymes carry out key dehydrogenation reactions in the Ehrlich pathway (Dickinson *et al.* 2002).

Multivariate and univariate analysis

Our original question pertained to whether the metabolic phenotype of one strain could be shifted in the direction of

another by adjusting the expression of a key transcription factor. This would suggest that changes in the regulation/expression of specific transcription factors could be responsible for major phenotypic divergence and adaptation of different *Saccharomyces* species or different strains within a species. To address this issue we followed both a multivariate approach (PCA) and created a statistical graph to visualize the overall structure of the volatile metabolite data set in a qualitative and quantitative manner.

Figure 3 clearly shows that the vast majority of metabolites in the transcription factor overexpression strain have shifted in the same direction as the target strain, either matching closely or in some cases overshooting the target. Only a few compounds shift in an opposite direction to that of the target strain. The PCA analysis in Figure 4 shows the overall shift in the metabolic profiles at each time point in fermentations performed with each of the five strains (DV10, VIN13, BM45, SOK2-VIN13, and RAP1-VIN13). On day 2 of fermentation the differences between the sample groupings of the *SOK2*-overexpressing VIN13 and the reference VIN13 strain is still small. The same is true of the *RAP1*-overexpressing strain and its BM45 control strain. However, by day 5 of fermentation the two transformed industrial strains form clearly distinct clusters that are separated from their control samples along the first three principal components. The same is true for day 14, when the distances between distinct sample groupings are even greater for the first two principal components.

As can be seen in Figure 4B, the overall exometabolite composition of the *RAP1*-overexpressing strain has shifted from the BM45 control strain in the direction of the DV10 target cluster for days 5 and 14 of fermentation. Similarly, *SOK2*-overexpressing samples shift from the VIN13 control cluster (Figure 4A), in the direction of the target BM45 cluster, even shifting beyond the target cluster in days 5 and 14.

Discussion

The adjustment of key transcription factor expression levels in a wine yeast strain can indeed alter metabolism on a large scale. More specifically, we were able to moderate metabolism in a qualitatively reasonably defined manner by engineering the expression levels of transcription factors identified by the analysis of high-quality comparative gene expression data. This was achieved despite the complexity of the regulation of aroma compound metabolism, which is affected by many other parameters, such as the prevailing redox balance, the concentration of intermediates, and the flux through upstream and downstream pathways, which affects the rates and directionality of many promiscuous enzymes that catalyze the reactions of higher alcohol and ester synthesis.

The data clearly support the hypothesis that microevolution, which has provided us with the plethora of industrial *Saccharomyces* strains known today, could use transcription factor moderation and/or binding site alteration to effect a large-scale rewiring of metabolic and regulatory circuits in the cell. The possibility thus exists to modify or enhance industrial wine yeasts in a holistic manner by carefully selecting and modifying high-level master regulatory systems, instead of instituting numerous single gene changes at the effector level.

Acknowledgments

We thank Jo McBride and the Cape Town Centre for Proteomic and Genomic Research for the microarray analysis and the staff and students at the Institute for Wine Biotechnology for their support and assistance in numerous areas. Funding for the research presented in this article was provided by the National Research Foundation of South Africa and research funding organization of the South African wine industry, Winetech.

Literature Cited

Abbott, D. A., T. A. Knijnenburg, L. M. De Poorter, M. J. Reinders, J. T. Pronk *et al.*, 2007 Generic and specific transcriptional responses to different weak organic acids in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. FEM. Yeast Res. 7: 819–833.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, 1990 Basic local alignment search tool. J. Mol. Biol. 215: 403–410.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, and J. G. Seidman, 1994 *Current Protocols In Molecular Biology.* John Wiley, New York.

Bakalinsky, A. T., and R. Snow, 1990 The chromosomal constitution of wine strains of *Saccharomyces cerevisiae*. Yeast 6: 367–382.

Bely, L., J. Sablayrolles, and P. Barre, 1990 Description of alcoholic fermentation kinetics: its variability and significance. Am. J. Enol. Vitic. 40: 319–324.

Bidenne, C., B. Blondin, S. Dequin, and F. Vezinhet, 1992 Analysis of the chromosomal DNA polymorphism of wine strains of *Saccharomyces cerevisiae*. Curr. Genet. 22: 1–7.

Borneman, A. R., J. A. Leigh-Bell, H. Yu, P. Bertone, M. Gerstein *et al.*, 2006 Target hub proteins serve as master regulators of development in yeast. Genes Dev. 20: 435–448.

Borneman, A. R., Z. D. Zhang, J. Rozowsky, M. R. Seringhaus, M. Gerstein *et al.*, 2007a Transcription factor bindind site identification in yeast: a comparison of high-density oligonucleotide and PCR-based microarray platforms. Funct. Integr. Genomics 4: 335–345.

Borneman, A. R., T. A. Gianoulis, Z. D. Zhang, H. Yu, J. Rozowsky *et al.*, 2007b Divergence of transcription factor binding sites across related yeast species. Science 317: 815–819.

Breiman, L., 2001 Random forests. Mach. Learn. 45: 5–22.

Chua, G., Q. D. Morris, R. Sopko, M. D. Robinson, O. Ryan *et al.*, 2006 Identifying transcription factor functions and targets by phenotypic activation. Proc. Natl. Acad. Sci. USA 103: 12045–12050.

Dermitzakis, E. T., and A. G. Clark, 2002 Evolution of transcription factor binding sites in mammalian gene regulatory regions: conservation and turnover. Mol. Biol. Evol. 19: 1114–1121.

Dickinson, J. R., L. Eshantha, J. Salgado, and M. J. E. Hewlins, 2002 The catabolism of amino acids to long chain and com-

plex alcohols in *Saccharomyces cerevisiae*. J. Biol. Chem. 278: 8028–8034.
Ernst, J., and Z. Bar-Joseph, 2006 STEM: a tool for the analysis of short time series gene expression data. BMC Bioinformatics 7: 191.
Frezier, V., and D. Dubourdieu, 1992 Ecology of yeast strain *Saccharomyces cerevisiae* during spontaneous fermentation in a Bordeaux winery. Am. J. Enol. Vitic. 43: 375–380.
Gimeno, C. J., and G. R. Fink, 1994 Induction of pseudohyphal growth by overexpression of PHD1, a *Saccharomyces cerevisiae* gene related to transcriptional regulators of fungal development. Mol. Cell. Biol. 14: 2100–2112.
Haugen, A. C., R. Kelley, J. B. Collins, C. J. Tucker, C. Deng *et al.*, 2004 Integrating phenotypic and expression profiles to map arsenic-response networks. Genome Biol. 5: R95.
Holmberg, S., and J. G. Petersen, 1988 Regulation of isoleucine-valine biosynthesis in *Saccharomyces cerevisiae*. Curr. Genet. 13: 207–217.
Horak, C. E., N. M. Luscombe, J. Qian, P. Bertone, S. Piccirrillo *et al.*, 2002 Complex transcriptional circuitry at the G1/S transition in *Saccharomyces cerevisiae*. Genes Dev. 16: 3017–3033.
Hou, L., X. Cao, M. Wang, and M. Lu, 2009 Effect of overexpression of transcription factors on the fermentation properties of *Saccharomyces cerevisiae* industrial strains. Lett. Appl. Microbiol. 49: 14–19.
Irizarry, R. A., B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs *et al.*, 2003 Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res. 31: e15.
Johnston, J. R., C. Baccari, and R. K. Mortimer, 2000 Genotypic characterization of strains of commercial wine yeasts by tetrad analysis. Res. Microbiol. 151: 583–590.
Kasahara, K., K. Ohtsuki, and S. Ki, K. Aoyama, H. Takahashi *et al.*, 2007 Assembly of regulatory factors on rRNA and ribosomal protein genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 27: 6686–6705.
Lee, T. I., N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph *et al.*, 2002 Transcriptional regulatory networks in *Saccharomyces cerevisiae*. Science 298: 799–804.
Li, X., S. Macartur, R. Bourgon, D. Nix, D. A. Pollard *et al.*, 2008 Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. PLoS Biol. 6: e190.
Liaw, A., and M. Wiener, 2002 Classification and Regression by randomForest. R News 2: 18–22.
Libkind, D., C. T. Hittinger, E. Valério, C. Gonçalves, J. Dover *et al.*, 2011 Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. Proc. Natl. Acad. Sci. USA 108: 14539–14544.
Lilly, M., F. F. Bauer, G. Styger, M. G. Lambrechts, and I. S. Pretorius, 2006 The effect of increased branched-chain amino acid transaminase activity in yeast on the production of higher alcohols and on the flavour profiles of wine and distillates. FEM. Yeast Res. 6: 726–743.
Mardia, K. V., J. T. Kent, and J. H. Bibby, 1979 *Multivariate Analysis*. Academic Press, London.
Mortimer, R., and J. Johnston, 1986 Geneology of principal strains of the Yeast Genetics Stock Center. Genetics 113: 35–43.
Mortimer, R., R. Lerner, and J. Barr, 1957 Ultraviolet-induced biochemical mutants of *Saccharomyces cerevisiae*. U.S. Afon.Energy. Comm. Doc. UCRL 3746: 1–10.
Ni, L., C. Bruce, C. Hart, J. Leigh-Bell, D. Gelperin *et al.*, 2009 Dynamic and complex transcription factor binding during an inducible response in yeast. Genes Dev. 23: 1351–1363.
Okazaki, S., T. Tachibana, A. Naganuma, N. Mano, and S. Kuge, 2007 Multistep disulfide bond formation in Yap1 is required for sensing and transduction of H2O2 stress signal. Mol. Cell 27: 675–688.
Pan, X., and J. Heitman, 2000 Sok2 regulates yeast pseudohyphal differentiation via a transcription factor cascade that regulates cell–cell adhesion. Mol. Cell. Biol. 20: 8364–8372.
Puig, S., and J. E. Perez-Ortin, 2000 Stress response and expression patterns in wine fermentations of yeast genes induced at the diauxic shift. Yeast 16: 139–148.
Rachidi, N., P. Barre, and B. Blondin, 1999 Multiple Ty-mediated chromosomal translocation lead to karyotype changes in a wine strain of *Saccharomyces cerevisiae*. Mol. Gen. Genet. 261: 841–850.
Ramakers, C., J. M. Ruijter, R. H. Deprez, and A. F. Moorman, 2003 Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci. Lett. 339: 62–66.
Rossouw, D., T. Naes, and F. F. Bauer, 2008 Linking gene regulation and the exo-metabolome: A comparative transcriptomics approach to identify genes that impact on the production of volatile aroma compounds in yeast. BMC Genomics 9: 530–548.
Rossouw, D., R. Olivares-Hernandes, J. Nielsen, and F. F. Bauer, 2009 Comparative transcriptomic approach to investigate differences in wine yeast physiology and metabolism during fermentation. Appl. Environ. Microbiol. 75: 6600–6612.
Rossouw, D., A. H. Van Den Dool, D. Jacobson, and F. F. Bauer, 2010 Comparative transcriptomic and proteomic profiling of industrial wine yeast strains. Appl. Environ. Microbiol. 10.1128/AEM.00586–10.
Saint-Prix, F., L. Bönquist, and S. Dequin, 2004 Functional analysis of the *ALD* gene family of *Saccharomyces cerevisiae* during anaerobic growth on glucose: the $NADP^+$ -dependent Aldp6p and Aldp5p isoforms play a major role in acetate formation. Microbiology 150: 2209–2220.
Salmon, J. M., 1997 Enological fermentation kinetics of an isogenic ploidy series derived form an industrial *Saccharomyces cerevisiae* strain. J. Ferment. Bioeng. 83: 253–260.
Sambrook, J., E. F. Fritsch, and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory Press, New York.
Sauer, F., and H. Jäckle, 1991 Concentration-dependent transcriptional activation or repression by Krüppel from a single binding site. Nature 10: 563–566.
Schütz, M., and J. Gafner, 1994 Dynamics of the yeast strain population during spontaneous alcoholic fermentation determined by CHEF gel electrophoresis. Lett. Appl. Microbiol. 19: 253–257.
Smoot, M., K. Ono, J. Ruscheinski, P.-L. Wang, and T. Ideker, 2011 Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27: 431–432.
Tan, K., H. Feizi, C. Luo, S. H. Fan, T. Ravasi *et al.*, 2008 A systems approach to delineate functions of paralogous transcription factors: role of the Yap family in the DNA damage response. Proc. Natl. Acad. Sci. USA 105: 2934–2939.
Teixeira, M. C., P. Monteiro, P. Jain, S. Tenreiro, A. R. Fernandes *et al.*, 2006 The YEASTRACT database: a tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*. Nucleic Acids Res. 34: D446–D451.
Thorsen, M., G. Lagniel, E. Kristiansson, C. Junot, O. Nerman *et al.*, 2007 Quantitative transcriptome, proteome, and sulfur metabolite profiling of the *Saccharomyces cerevisiae* response to arsenite. Physiol. Genomics 30: 35–43.
Tusher, C. G., R. Tibshirani, and G. Chu, 2001 Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA 98: 5116–5121.
Vestrepen, K. J., S. D. Van Laere, B. M. Vanderhaegen, G. Derdelinckx, J. P. Dufour *et al.*, 2003 Expression levels of the yeast alcohol acetyltransferase genes ATF1, Lg-ATF1, and ATF2 control the formation of a broad range of volatile esters. Appl. Environ. Microbiol. 69: 5228–5237.

Wenzel, T. J., A. Migliazza, H. Y. Steensma, and J. A. Van Den Berg, 1992 Efficient selection of phleomycin-resistant *Saccharomyces cerevisiae* transformants. Yeast 8: 667–668.

Workman, C. T., H. C. Mak, S. Mccuine, J. B. Tagne, M. Agarwal *et al.*, 2006 A systems approach to mapping DNA damage response pathways. Science 312: 1054–1059.

Yarragudi, A., L. W. Parfrey, and R. H. Morse, 2007 Genome-wide analysis of transcriptional dependence and probable target sites for Abf1 and Rap1 in *Saccharomyces cerevisiae*. Nucleic Acids Res. 35: 193–202.

Zheng, W., H. Zhao, E. Mancera, L. M. Steinmetz, and M. Snyder, 2010 Genetic analysis of variation in transcription factor binding in yeast. Nature 464: 1187–1191.

Communicating editor: M. Johnston

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2011/10/31/genetics.111.132720.DC1

Transcriptional Regulation and the Diversification of Metabolism in Wine Yeast Strains

Debra Rossouw, Dan Jacobson, and Florian F. Bauer

DOI: 10.1534/genetics.111.132720

Table S1 Primers used for amplification of target genes.

Primer Name	Sequence (5'-3')
PhR322F	GATCCACGTCGGTACCCGGGGGATC
PhR322R	GATCGCGATCGCAAGCTTGCAAATTAAAGCC
RAP1f	TTAAGCGGCCGCATACGCAACCGCCCTACATAA
RAP1r	TCTACATATGCGTGAATCAGTGAAATAAAGG
SOK2f	TTAAGCGGCCGCTATAACCCTGGTAAGGTCCTT
SOK2r	TCTACATATGGGCGGTAGGGTTTTGATTAA

Table S2 Target genes and primers for QRT-PCR

Gene ID	Forward primer (5'-3')	Reverse primer (5'-3')
ADH2	TTCAAGCCGCTCACATTCC	CACAAGATTGGCGCGACTT
ALD4	TTGTGGGTGAGGCCATTACA	ACCCTGTGAAGGCAACCTTTT
ARO10	AGTGTTGAATCAGCTGGCCTAAG	CATAAGCGGCGTTCAGTTCAT
ATF2	GTTCGGCCTAAACGTTTGCT	CCACGCTCATGTCCATGTTC
BAT1	CCATGTTCCGTCCGGATAAG	CAAACAAATTCTAGCGGCAG
BAT2	AATCTGTTTGCCAACGTTCGA	TGCTGGATCAGTTTCCCAATT
ERG10	CGTGCGGGTGCCAAAT	CCATCTCTTTCGACACCATCAA
ERG13	GATCGGTCCTGATGCTCCAA	CGTAGGCGTGTTCCATGTAAGA
HAT2	TGCCCGCAACCTTTCAA	GGCCGCAAGGAGGTTTG
ILV3	CGTCCCAGGCCATGCTT	CCCGACTTGAGGCTTCTTGA
RAP1	ATTGGATCCGAGTATGGTCGTT	TCCGATGGCGCTGTGACT
SOK2	TCAACCTCTGATGCCCGTATC	GCGGGTACGGCCACTGT
THI3	GGCGTGGCCGGATCTTA	GGCGGCATACCCACTATGTG
YJL218W	GGTCATCCAATTGACGTGGAA	GGTCACAGGCATGGCATATTC

Table S3 Sok2p and Rap1p activity with reference to the target genes in figure 2.

	SOK2	RAP1
ADH2	√	X
ALD4	√	√
ARO10	√	√
ATF2	√	√
BAT1	√	X
BAT2	√	√
ERG10	X	X
ERG13	X	√
HAT2	X	√
ILV3	√	X
THI3	X	X
YJL218W	X	√
RAP1	X	n/a
SOK2	n/a	√

Transcription factor activity is based on reported interaction studies by Vachova *et al.* (2004), Chua *et al.* (2006), Workman *et al*. (2006), Kasahara *et al*. (2007) and Yarragudi *et al*. (2007). Tic marks indicate evidence for regulation whereas X's are used where no evidence for regulation of the target gene by the transcription factor in question has been reported.

Table S4 Volatile alcohols and esters present in the fermentation media at day 2 of fermentation.

DAY2	VIN13	SOK2-VIN13	BM45	RAP1-BM45	DV10
Ethyl Acetate	5.53 ± 1.40	5.70 ± 1.20	7.60 ± 0.71	6.14 ± 2.16	8.10 ± 2.13
Propanol	33.24 ± 4.38	34.25 ± 3.19	32.81 ± 1.18	27.25 ± 1.37	28.39 ± 5.21
Isobutanol	5.78 ± 0.71	**8.71 ± 0.74**	9.26 ± 0.70	8.20 ± 1.70	6.20 ± 1.62
Isoamyl Acetate	0.10 ± 0.11	**0.31 ± 0.02**	0.18 ± 0.11	0.24 ± 0.11	0.17 ± 0.17
Butanol	0.16 ± 0.2	**0.41 ± 0.12**	Bd	Bd	Bd
Isoamyl alcohol	32.58 ± 5.74	37.27 ± 3.82	37.80 ± 2.90	35.85 ± 3.30	32.78 ± 3.61
Ethyl Hexanoate	Bd	bd	Bd	Bd	0.17 ± 0.17
Hexanol	Bd	bd	Bd	Bd	Bd
Ethyl Caprylate	0.05 ± 0.04	0.09 ± 0.01	0.10 ± 0.02	0.11 ± 0.03	0.11 ± 0.03
Acetic Acid	449.5 ± 17.8	**525.2 ± 26.2**	715.3 ± 18.9	*658.8 ± 7.0*	618.4 ± 15.4
Propionic Acid	2.23 ± 0.15	2.47 ± 0.18	2.04 ± 0.19	2.15 ± 0.23	2.38 ± 0.31
Iso-Butyric Acid	0.78 ± 0.04	0.71 ± 0.02	0.79 ± 0.06	0.68 ± 0.04	0.80 ± 0.06
Butyric Acid	0.55 ± 0.04	0.52 ± 0.01	0.58 ± 0.05	0.57 ± 0.01	0.67 ± 0.02
Ethyl Caprate	0.08 ± 0.016	0.09 ± 0.02	0.12 ± 0.04	0.16 ± 0.06	0.10 ± 0.02
Iso-Valeric Acid	0.45 ± 0.03	0.37 ± 0.01	0.47 ± 0.08	*0.33 ± 0.04*	0.38 ± 0.06
Diethyl Succinate	Bd	bd	Bd	Bd	Bd
Valeric Acid	Bd	bd	Bd	Bd	Bd
2-Phenylethyl Acetate	Bd	bd	Bd	Bd	Bd
Hexanoic Acid	0.73 ± 0.03	**0.85 ± 0.07**	0.94 ± 0.13	1.05 ± 0.15	1.39 ± 0.07
2-Phenyl Ethanol	6.42 ± 0.47	7.11 ± 0.69	9.64 ± 0.35	*7.57 ± 0.78*	7.49 ± 0.50
Octanoic Acid	0.76 ± 0.15	1.14 ± 0.26	1.25 ± 0.64	1.03 ± 0.10	3.05 ± 0.92
Decanoic Acid	2.54 ± 0.19	2.34 ± 0.26	2.73 ± 0.12	2.95 ± 0.38	3.33 ± 0.09

All values are expressed in $mg.L^{-1}$ and are the average of 4 biological repeats ± standard deviation. Metabolites present at concentrations below the detection limit are indicated by "Bd". Values in bold indicate a statistically significant increase in concentration for a given metabolite relative to the untransformed control, whereas values in italics indicate a significant decrease in concentration.

Table S5 Volatile alcohols and esters present in the fermentation media at day 5 of fermentation.

DAY5	VIN13	SOK2-VIN13	BM45	RAP1-BM45	DV10
Ethyl Acetate	19.74 ± 2.48	22.52 ± 2.65	20.52 ± 1.13	19.38 ± 0.95	28.38 ± 1.69
Propanol	70.22 ± 2.34	**82.06 ± 4.97**	48.65 ± 3.43	44.66 ± 3.02	66.88 ± 5.64
Isobutanol	12.97 ± 1.95	**18.17 ± 1.84**	20.14 ± 1.96	17.29 ± 1.55	16.42 ± 1.81
Isoamyl Acetate	0.30 ± 0.09	**0.70 ± 0.17**	0.36 ± 0.02	0.36 ± 0.04	0.38 ± 0.09
Butanol	0.59 ± 0.09	**0.88 ± 0.04**	0.52 ± 0.03	0.58 ± 0.04	0.69 ± 0.05
Isoamyl alcohol	78.74 ± 4.54	**106.80 ± 8.47**	85.54 ± 4.69	89.53 ± 2.01	95.87 ± 7.52
Ethyl Hexanoate	0.11 ± 0.18	0.15 ± 0.01	0.16 ± 0.16	0.10 ± 0.08	0.18 ± 0.05
Hexanol	Bd	bd	Bd	bd	Bd
Ethyl Caprylate	0.11 ± 0.04	0.12 ± 0.01	0.14 ± 0.00	*0.10 ± 0.01*	0.15 ± 0.03
Acetic Acid	792.6 ± 16.4	**1047.3 ± 72.87**	1131.1 ± 44.0	1159.9 ± 113.6	1093.2 ± 81.7
Propionic Acid	4.58 ± 0.42	**6.56 ± 0.44**	2.62 ± 0.10	2.78 ± 0.28	5.05 ± 0.39
Iso-Butyric Acid	0.83 ± 0.03	0.87 ± 0.05	0.90 ± 0.05	0.81 ± 0.06	0.89 ± 0.06
Butyric Acid	0.65 ± 0.08	0.67 ± 0.05	0.68 ± 0.04	0.71 ± 0.05	0.80 ±0.08
Ethyl Caprate	0.24 ± 0.05	**0.35 ± 0.03**	0.30 ± 0.04	0.33 ± 0.07	0.46 ± 0.03
Iso-Valeric Acid	0.65 ± 0.07	0.66 ± 0.04	0.62 ± 0.09	0.50 ± 0.06	0.64 ± 0.06
Diethyl Succinate	0.03 ± 0.05	**0.14 ± 0.03**	0.10 ± 0.00	**0.15 ± 0.00**	0.11 ± 0.01
Valeric Acid	0.02 ± 0.03	0.06 ± 0.01	0.02 ± 0.02	**0.07 ± 0.00**	0.05 ± 0.00
2-Phenylethyl Acetate	0.01 ± 0.67	0.04 ± 0.00	0.03 ± 0.60	0.02 ± 0.01	0.03 ± 0.04
Hexanoic Acid	1.11 ± 0.17	1.40 ± 0.16	1.37 ± 0.28	1.56 ± 0.32	2.19 ±0.24
2-Phenyl Ethanol	10.74 ± 0.68	**14.62 ± 0.84**	12.66 ± 0.66	13.10 ± 2.10	13.52 ± 1.25
Octanoic Acid	1.38 ± 0.08	**1.65 ± 0.13**	1.34 ± 0.21	1.28 ± 0.09	2.65 ± 0.12
Decanoic Acid	2.80 ±0.17	3.28 ± 0.21	2.98 ± 0.39	3.80 ± 0.14	4.50 ± 0.29

All values are expressed in $mg.L^{-1}$ and are the average of 4 biological repeats ± standard deviation. Metabolites present at concentrations below the detection limit are indicated by "bd". Values in bold indicate a statistically significant increase in concentration for a given metabolite relative to the untransformed control, whereas values in italics indicate a significant decrease in concentration.

Table S6 Volatile alcohols and esters present in the fermentation media at day 14 of fermentation.

DAY14	VIN13	SOK2-VIN13	BM45	RAP1-BM45	DV10
Ethyl Acetate	31.39 ± 0.66	28.09 ± 1.41	27.11 ± 2.85	23.88 ± 0.99	33.18 ± 0.43
Propanol	76.48 ± 3.09	83.37 ± 6.25	45.60 ± 1.21	41.53 ± 4.01	69.52 ± 5.30
Isobutanol	19.00 ± 1.74	**24.96 ± 0.53**	25.88 ± 2.81	22.42 ± 1.65	21.27 ± 3.07
Isoamyl Acetate	0.34 ± 0.04	**0.73 ± 0.02**	0.43 ± 0.03	0.40 ± 0.04	0.43 ± 0.11
Butanol	1.07 ± 0.07	**1.33 ± 0.04**	0.58 ± 0.06	**0.70 ± 0.05**	0.87 ± 0.06
Isoamyl alcohol	106.8 ± 9.37	**132.74 ± 7.57**	104.61 ± 3.42	108.09 ± 7.01	113.69 ± 11.49
Ethyl Hexanoate	0.22 ± 0.19	0.36 ± 0.01	0.35 ± 0.01	*0.19 ± 0.02*	0.39 ± 0.03
Hexanol	Bd	0.01 ± 0.01	Bd	**0.35 ± 0.02**	Bd
Ethyl Caprylate	0.15 ± 0.02	**0.26 ± 0.02**	0.24 ± 0.04	0.23 ± 0.05	0.29 ± 0.03
Acetic Acid	926.9 ± 50.2	**1182.9 ± 87.8**	1154.6 ± 112.7	1263.2 ± 85.9	1261.0 ± 47.1
Propionic Acid	6.05 ± 0.48	**7.93 ± 0.63**	2.81 ± 0.17	**5.07 ± 0.42**	8.01 ± 0.22
Iso-Butyric Acid	0.76 ± 0.03	**0.96 ± 0.04**	0.96 ± 0.07	0.86 ± 0.03	1.02 ± 0.10
Butyric Acid	0.49 ± 0.04	**0.59 ± 0.02**	0.61 ± 0.04	0.63 ± 0.06	0.75 ± 0.01
Ethyl Caprate	0.32 ± 0.04	**0.47 ± 0.05**	0.43 ± 0.04	0.50 ± 0.09	0.59 ± 0.04
Iso-Valeric Acid	0.84 ± 0.01	0.87 ± 0.04	0.79 ± 0.09	0.67 ± 0.07	0.91 ± 0.12
Diethyl Succinate	Bd	0.07 ± 0.03	Bd	**0.11 ± 0.04**	0.05 ± 0.05
Valeric Acid	Bd	bd	Bd	**0.22 ± 0.15**	0.01 ± 0.01
2-Phenylethyl Acetate	0.03 ± 0.02	0.06 ± 0.00	0.04 ± 0.01	0.02 ± 0.01	0.04 ± 0.01
Hexanoic Acid	1.53 ± 0.08	**2.28 ± 0.28**	2.56 ± 0.60	2.76 ± 0.30	3.28 ± 0.51
2-Phenyl Ethanol	13.68 ± 0.88	**20.43 ± 1.54**	15.16 ± 0.74	*12.93 ± 0.83*	16.07 ± 0.69
Octanoic Acid	1.15 ± 0.06	1.31 ± 0.11	1.13 ± 0.19	1.25 ± 0.29	1.93 ± 0.15
Decanoic Acid	2.18 ± 0.04	2.38 ± 0.11	1.95 ± 0.16	2.34 ± 0.21	3.45 ± 0.12

All values are expressed in $mg.L^{-1}$ and are the average of 4 biological repeats ± standard deviation. Metabolites present at concentrations below the detection limit are indicated by "bd". Values in bold indicate a statistically significant increase in concentration for a given metabolite relative to the untransformed control, whereas values in italics indicate a significant decrease in concentration.

Table S7 GO categorisation of differentially expressed transcripts with a fold change greater than or equal to 2 (i.e. up-regulated genes).

Category	p-value	In Category from Cluster	k	f
ribosome biogenesis [GO:0042254]	<1e-14	RPS9B ENP1 SPB1 PWP2 NOP1 NOP14 NHP2 RLI1 UTP4 UTP5 SNU13 TMA20 NSA2 SPB4 LOC1 CGR1 DBP3 ROK1 SLX9 UTP8 CIC1 RRP3 GAR1 IMP3 RIX1 UTP18 UTP10 ALB1 MRT4 URB1 EBP2 RPL40B SOF1 RLP24 SDO1 DIP2 CBF5 EMG1 NOP56 UTP13 DBP9 UTP21 ERB1 UTP15 RRB1 HAS1 NOP2 DBP2 NOG2 ESF2 NOP12 BRX1 NOC2 PUS7 YTM1 RRS1 NOP58 RRP12 RPS9A NOG1 NIP7 RRP9	62	184
regulation of translation [GO:0006417]	<1e-14	RBG1 URA7 ILS1 RPS11B RPG1 GRS1 RPS9B ARO4 SRO9 THR4 RPL13A SSB1 RPL4B RPS11A RLI1 RPL12A RPS8B YGR054W VAS1 TIF4631 ADE3 RPS0A MES1 RPL14B IMD2 RPS24B THS1 URA2 RPL17B SUI2 TEF4 EAP1 GCN3 DPS1 FRS1 SAM1 PWP1 IMD3 RPS1A RPL6B IMD4 RPL6A NIP1 DBP2 SSB2 RPL18B WRS1 PRT1 RPS9A CDC60 NEW1 TIF5 TKL1	53	172
cellular amino acid and derivative metabolic process [GO:0006519]	3.5E-14	ILS1 ADH5 HIS7 ARO4 ILV6 LYS21 GGC1 KRS1 TRP4 HOM3 HIS1 ILV1 TRP2 TRP5 ARO2 ARO8 ASN2 LYS1 ACO2 TRP3 SAM1 ATR1 ARG7 ADE4 ARG1 LEU9 ORT1 HIS3 SAM4 YMC1 ASN1	31	71
cellular amino acid biosynthetic process [GO:0008652]	1.7E-13	LYS2 HIS7 ARO4 ILV6 THR4 LYS21 LYS4 TRP4 PRO3 HOM3 HIS1 SER3 ILV1 TRP2 IRC7 TRP5 ARO2 ASN2 ADE3 SER2 THR1 LYS12 LYS1 ARG3 MDE1 TRP3 ILV5 YML096W ARG7 LYS9 ARG1 LEU9 ORT1 HIS3 PRO2 SAM4 ARO7 ASN1	38	107
nitrogen compound metabolic process [GO:0006807]	3.2E-13	ILS1 ADH5 HIS7 ARO4 ILV6 LYS21 GGC1 KRS1 TRP4 HOM3 HIS1 ILV1 TRP2 TRP5 ARO2 ARO8 ASN2 ADE3 NIT1 LYS1 URA2 ACO2 TRP3 SHM2 SAM1 ATR1 ARG7 ADE4 ARG1 LEU9 ORT1 HIS3 SAM4 YMC1 ASN1	35	94
cellular aromatic compound metabolic process [GO:0006725]	6.3E-13	ADH5 HIS7 ARO4 ILV6 LYS21 GGC1 KRS1 TRP4 HOM3 HIS1 ILV1 TRP2 TRP5 YGL039W ARO2 ARO8 ASN2 LYS1 ACO2 TRP3 ATR1 ARG7 ARG1 LEU9 ORT1 HIS3 SAM4 YMC1 ASN1	29	68
organic acid metabolic process [GO:0006082]	2.0E-11	ADH5 HIS7 ARO4 ILV6 LYS21 GGC1 KRS1 TRP4 HOM3 HIS1 ILV1 TRP2 TRP5 ARO2 ARO8 ASN2 LYS1 ACO2 MAE1 TRP3 SAM1 ATR1 ARG7 ARG1 LEU9 ORT1 HIS3 SAM4 YMC1 ASN1	30	81
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000462]	4.3E-10	RPS11B RPS9B PWP2 NOP1 NOP14 RPS11A UTP4 UTP5 SNU13 RPS8B PRP43 SLX9 UTP8 EFG1 RRP3 RPS24B UTP25 UTP10 SOF1 DIP2 NOP56 UTP13 TSR2 RPS1A UTP15 RRP12 RPS9A	27	75
tRNA aminoacylation for protein translation [GO:0006418]	6.3E-09	ILS1 GRS1 SES1 KRS1 FRS2 ARC1 VAS1 TYS1 MES1 DED81 YHR020W THS1 DPS1 YNL247W WRS1 GLN4 CDC60	17	36
metabolic process [GO:0008152]	2.7E-08	BNA4 LYS2 ADH5 HIS7 ARO4 MAL32 ILV6 THR4 SLC1 QRI1 LYS21 PHO13 LYS4 EXG2 TRP4 URH1 URA3 UTR2 HPA3 PRO3 SAH1 HOM3 SER3 ILV1 TRP5 SCW11 YGL039W ARI1 CRH1 ADE3 SER2 ARD1 PAN5 IMD2 RHR2 SUC2 LYS1 RPE1 URA2 ACO2 BNA1 PGU1 YKL027W MAE1 MCD4 TRP3 URA1 ACS2 DPH5 ILV5 DUS3 YLR426W IMD3 ERG6 IMD4 ERG13 PLB2 ERG12 ADE4 SCW10 NRK1 SPS19 LYS9 DSE4 GPD2 LEU9 PRO2 ERG10 ALD6 TKL1	70	377
sterol biosynthetic process [GO:0016126]	1.4E-06	ERG25 ERG11 ERG3 ERG6 HMG1 ERG13 ERG5 ERG2 ERG12 CYB5 MVD1 HES1 IDI1	13	30
steroid biosynthetic process [GO:0006694]	1.5E-06	ERG25 ERG11 ERG3 ERG6 HMG1 ERG13 ERG5 ERG2 ERG12 MVD1 HES1 IDI1	12	26
peptidyl-amino acid modification [GO:0018193]	2.4E-06	NCL1 PRS4 HMT1 ARO4 DUT1 FEN1 DTD1 PHO13 RLI1 UTP5 UTR2 MAK10 ARD1 PHO90 HAM1 LIA1 SUR4 OST6 APT1 URA5 GAS3 SEC63 ALG5	23	84
alcohol metabolic process [GO:0006066]	3.7E-06	YAT1 ERG25 ATF2 ERG11 YEH1 ERG3 ACS2 ERG6 HMG1 ERG13 ERG5 ERG2 ERG12 CYB5 MVD1 ERG10 IDI1	17	52
endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA	1.0E-05	ENP1 PWP2 NOP14 LOC1 RPS0A NOP9 UTP18 UTP10 DIP2 EMG1 UTP13 RPS18B ESF2 RRS1 NOP58	15	45

transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000447]				
translational elongation [GO:0006414]	2.7E-05	BNA4 AAC3 ADH5 CTP1 ILV6 APA1 BSC1 PRM7 RPP1B GGC1 PAD1 GLY1 HPA3 DLD3 HOM3 PTC2 TRP2 WWM1 IRC7 MF(ALPHA)2 SCS3 SUT1 MIG2 BUD9 PTI1 INM1 TIM44 POR2 PAN6 LYS1 PRY3 ARG3 ESS1 SFK1 TEF4 HOT13 EAP1 SSA2 NEJ1 DIC1 HXT2 SSO2 DSK2 RNH201 DSE4 ARG1 GPD2 HPF1 NRT1 HIS3 FIT2 ALD6 DIP5 DPM1	54	322
RNA modification [GO:0009451]	4.0E-05	MAK16 HMT1 ENP1 SPB1 RSA4 RPC53 TRM82 DBP3 YVH1 ALB1 CBF5 DUS3 TSR2 GCD10 TRM11 PUS7 RRS1 PUS1	18	67
pyrimidine nucleotide biosynthetic process [GO:0006221]	4.5E-05	URA7 URA3 DCD1 URA2 URA8 URA1 URA4 URA5	8	16
glutamine metabolic process [GO:0006541]	5.4E-05	URA7 HIS7 ASN2 URA2 URA8 TRP3 YML096W GUA1 ADE4 ASN1	10	25
lipid biosynthetic process [GO:0008610]	9.1E-05	FEN1 ERG25 ERG11 SFK1 ERG3 SUR4 ERG6 HMG1 ERG13 ERG5 ERG2 ERG12 MVD1 HES1 IDI1	15	53
ergosterol biosynthetic process [GO:0006696]	2.3E-04	ERG25 ERG11 ERG3 ERG6 HMG1 ERG13 ERG5 ERG2 ERG12 ERG10	10	29
GMP biosynthetic process [GO:0006177]	3.9E-04	IMD2 IMD3 IMD4 GUA1	4	5
lysine biosynthetic process [GO:0009085]	4.3E-04	LYS2 LYS21 LYS4 LYS12 LYS1 LYS9	6	12
cellular amino acid metabolic process [GO:0006520]	4.3E-04	THR4 ASP1 GLY1 ILV1 IRC7 ARG3 URA2 MAE1 AAT1 ASP3-1	10	31
translation [GO:0006412]	4.4E-04	MAK16 FUN12 ILS1 RPS11B RPG1 GRS1 RPS9B SRO9 RPL13A RPP1B RPL41A DTD1 SSB1 RPL4B SES1 RPS11A KRS1 RLI1 RPS17B RPL12A TMA20 RPL34A RPS8B RPS26B FRS2 RPL22B RPL9A YGR054W VAS1 RPL24B TIF4631 TYS1 RPS0A MES1 RPL14B DED81 YHR020W RPS24B THS1 RPL17B RPS22A SUI2 RPL43B TEF4 GCN3 RPL40B DPS1 RLP24 FRS1 RPL31B RPS1A RPL6B RPS18B RPL6A TIF34 RPL36A RPS10B NIP1 SSB2 YNL247W RPL18B WRS1 RPL18A RPS19A RPS28A GLN4 RPS10A PRT1 RPS9A CDC60 TIF5 TIF3	72	511
aromatic amino acid family biosynthetic process [GO:0009073]	3.5E-03	ARO4 TRP4 TRP2 TRP5 ARO2 TRP3 ARO7	7	22
lysine biosynthetic process via aminoadipic acid [GO:0019878]	3.6E-03	LYS2 LYS21 LYS4 LYS1 LYS9	5	12
de novo NAD biosynthetic process from tryptophan [GO:0034354]	7.6E-03	BNA4 BNA1 BNA2	3	5

k refers to the number of differentially expressed genes annotated with the GO term.
f refers to the total number of genes present in the yeast genome annotated with the GO term.

Table S8 GO categorisation of differentially expressed transcripts with a fold change less than or equal to -2 (i.e. down-regulated genes).

Category	p-value	In Category from Cluster	k	f
vacuolar protein catabolic process [GO:0007039]	<1e-14	ACS1 BDH2 VID24 TPS1 NGR1 SDS24 HBT1 GYP7 NTH1 RCR2 REG1 UBC5 TPS2 KIN1 GGA1 TSA2 PEX29 GLC3 UBC8 EDC2 PIC2 GIP2 SSA4 SPI1 HSP12 RIM15 CMK1 PMC1 AMS1 ATG1 MDS3 VID30 STF2 CTT1 XKS1 SOL4 COQ6 HUA1 SOD2 GRE3 VID28 PIG2 FYV10 PFK26 GUT2 GTT1 BBC1 TPK1 PTK2 KNS1 HSP104 VPS13 TFS1 GSY2 TSL1 YPK2 PGM2 GID8 RIM11 SIP5 ALD3 ALD2 DDR48 PAI3 GAD1 TPS3 UBP15 MAM3 ATG19 DCS2 GSP2 YDC1 PIN3 GPH1 GDB1	75	129
autophagy [GO:0006914]	<1e-14	SEC17 ATG8 ATG14 ATG12 SAF1 ATG15 RPN4 ATG20 IWR1 ATG9 RRI1 GYP7 RCR2 CIS1 NRG1 UBC5 DOA4 KIN1 MFB1 RMD5 PIB1 UBX5 PEX29 UBC8 EDC2 AST2 RAD4 ATG18 PMC1 MON1 AMS1 CUP2 ATG1 VID30 NQM1 TAM41 RAD2 HSE1 ATG7 NVJ1 PIG2 FYV10 PFK26 HOS4 SNX4 TAX4 FBP26 TPK1 RAD26 IRS4 ECM4 ATG10 APC9 CLF1 PUS5 TFS1 ATG26 VPS34 NDL1 ATG17 RSF1 HFD1 GID8 PSO2 ATG16 ALD2 PAI3 RAD14 ATG4 ATG2 ATG19 PKH2 AHC1 SHE4 PEP12 GLO4 VAM3 DCS2 DGA1 ATG21 ATG29 ATG11 ATG13	83	151
cell death [GO:0008219]	5.1E-14	COR1 ATP1 ATP3 COX9 INH1 SDH4 ATP5 ATP17 QCR7 RIP1 QCR6 COX4 QCR9 QCR10 COX6 UBX6 QCR8 ATP2 ATP7 SDH1 SDH2 COX8 COX5A CYT1 PIN3 QCR2	26	29
response to temperature stimulus [GO:0009266]	2.0E-13	BDH2 HSP26 TPS1 SDS24 NTH1 UBC5 TPS2 HSP42 HSP78 SPI1 HSP12 CTT1 SOL4 SPL2 ECM4 HSP104 TFS1 TMA10 TSL1 MSC1 PGM2 ALD3	22	23
energy reserve metabolic process [GO:0006112]	3.1E-13	GIP4 GIP1 VPS15 IRA1 ICS2 LRE1 GPR1 STP4 BDF2 MRK1 MSH5 YDL180W SNF3 PRR2 RRI1 MFB1 MTH1 DOT6 RIM15 SAP155 EDC1 CPD1 MGA1 HXT4 RPI1 POG1 TAX4 GSH1 ASG7 HXT8 YJR115W HAP4 YKR017C SPO75 COX19 VPS13 GAL2 RFX1 CRR1 YLR446W FMP27 SRT1 CAT8 YNL144C ATG2 YNR034W-A AEP3 RDS2 PUF2	49	89
biological_process [GO:0008150]	5.7E-13	FUN14 FUN19 AIM2 BDH2 UIP3 YAR028W YAR029W YBL029C-A YBL029W MOH1 SEF1 YBL086C YBL095W EDS1 RFS1 YBR053C YRO2 YBR063C YBR085C-A YBR090C AIM3 OPY1 RTC2 APD1 ICS2 YBR182C-A YBR200W-A YBR204C YBR219C YBR221W-A YBR225W OM14 YBR230W-A ERT1 MTC4 AIM5 FMP21 BIT2 YBR284W YBR285W YBR287W YCL012C YCL021W-A YCL057C-A YCR007C CTR86 YCR076C HMRA2 YCR108C YDL027C STP4 BDF2 UBX3 TMA17 YDL114W SNA4 YDL133W UGX2 YDL180W RTN2 YDL206W YDL218W YDL233W YDR034W-B YDR042C RTR2 FMP16 ALT2 YDR124W YDR169C-A YDR182W-A YDR186C YDR246W-A YDR249C YDR262W BSC2 YDR282C HRQ1 YDR357C YDR366C YDR379C-A YDR381C-A YDR391C THI74 YDR444W JIP4 SPG3 PSP1 YDR506C SNA2 KRE28 HSP31 YDR541C VAB2 YEL020C UTR5 RMD6 YEL073C YEL076C PHM8 HVG1 YER039C-A SAP1 RRT13 RGI1 YER078W-A YER079W AST2 YER137C YER158C YER175W-A FMP10 YER184C YFL012W YFL034W YFL041W-A YFL042C COS4 YFL064C AIM13 YFR012W-A YFR016C YFR017C PES4 RMD8 YGL006W-A YGL007C-A YGL010W YGL081W TOS8 SNT2 RRT6 AIM14 YGL176C MTC3 SHE10 YGL235W YGL258W-A YPS5 YGR016W YGR021W NQM1 FMP48 YGR053C YGR066C YGR067C YGR079W YGR093W YGR121W-A YGR125W YGR126W YGR127W YGR130C ECL1 YGR146C-A YGR153W YGR174W-A YGR201C YGR204C-A YGR205W YGR235C SPG1 YGR237C YGR250C HUA1 YHL012W YLF2 YHL015W-A YHL018W YHL026C YHL048C-A YHR007C-A YSC83 YHR033W YHR035W YHR050W-A YHR078W YHR080C YHR097C YPT35 ANS1 YHR140W YSP1 LIN1 YHR159W AIM18 AIM46 YHR202W YIL024C YIL029C YIL055C YIL060W YIL077C AIM19 YIL089W YIL102C OM45 YIR014W YIR016W YIR018C-A PET130 IKS1 MPM1 YJL070C ICS3 YJL077W-B PRY1 IML2 AIM23 YJL132W YJL136W-A YJL147C DAS1 FMP33 YJL163C YJL181W YJL185C YJL193W YJL206C REE1 YJR005C-A YJR008W YJR039W HIT1 YJR056C YJR061W AIM24 YJR085C YJR112W-A YJR115W IML1 YJR149W DAN4 YJR151W-A YJR154W YKL018C-A YKL023W YKL050C YKL068W-A YKL070W YKL071W STB6 CUE2 YKL091C MTC2 YKL100C YKL106C-A YKL107W DGR2 RMA1 YKL133C MRP8 YKL151C YKL162C KKQ8 YKL222C YKR005C YKR011C YKR017C YKR018C FMP46 YKR051W YKR096W UBP11 YLL056C YLL066W-B YLR001C YLR030W YLR031W YLR046C YLR049C YLR053C YLR125W YLR149C PCD1 YLR152C RRT15 YLR164W YLR173W YLR177W SKG3 IRC20 YLR257W BOP2 YLR283W YLR297W YLR312C TMA10 STP3 ART10 YLR422W YLR445W YLR446W FMP27 YLR466C-B YML002W YML003W YML007C-A YML020W AIM31 YML037C PRM6 AIM32 TCB3 AIM33 YML100W-	447	1237

		A NAB6 NGL3 YML131W YMR018W YMR031C YMR034C YET2 YMR084W YMR085W YMR090W YMR105W-A SPG4 YMR114C YMR118C YMR124W YMR155W YMR158C-A YMR160W SIP18 YMR175W-A YMR178W YMR181C YMR182W-A YMR185W SPG5 ICY1 YMR196W YMR206W YMR252C YMR253C YMR258C YMR262W YMR265C PGM3 YMR295C SNO4 ERR3 YNL011C YNL033W YNL040W YNL042W-B APJ1 YNL092W AIM37 YNL115C YNL144C YNL146C-A YNL155W YNL165W YNL176C YNL193W YNL195C SLZ1 YNL200C VID27 YNL234W YTP1 RTC4 YNL260C BSC4 YNL277W-A YNL295W YNL305C SWM2 YNR034W-A YNR068C BSC5 IRC10 YOL024W YOL036W RRT8 AIM39 YOL073C PHM7 YOL087C RTC1 ZPS1 YOL159C YOR011W-A YOR032W-A YOR034C-A IRC23 YOR052C RTS2 TCB1 YOR097C YOR152C PNS1 YOR161C-C YOR186W YOR192C-C YOR214C AIM41 RCN2 YOR223W YOR228C RDL1 YOR289W YOR292C YOR293C-A YOR316C-A MNE1 YOR365C YOR376W-A YOR381W-A FRE5 YOR389W PAU21 YPL038W-A YPL039W LEE1 CWC27 YTA6 YPL077C YPL107W YPL109C YPL119C-A PRM4 YPL162C SET6 YPL168W UIP4 FMP40 GRE1 YPL236C YPL247C YPL257W YPL260W YPL277C YPL278C YPR022C JID1 ASA1 YPR091C YPR098C YPR109W YPR117W YPR127W URN1 CUR1 YPR159C-A		
phosphorus metabolic process [GO:0006793]	3.6E-12	COR1 ATP1 PKC1 ATP3 ATP16 MPS1 STE7 INH1 SDH4 ATP5 ATP17 QCR7 RIP1 COX13 QCR9 QCR10 COX6 KIC1 PBS2 QCR8 ATP2 ATP7 MDH1 SDH3 SDH1 SDH2 COX12 RIM11 COX7 FPK1 CYT1 ATP4 ATP20 QCR2	34	48
mitochondrion degradation [GO:0000422]	2.8E-10	ATG8 ATG12 PTC6 ATG20 ATG9 CIS1 ATG1 ATG7 ATG32 SNX4 FCJ1 ATG10 ATG33 ATG17 ATG16 ATG4 ATG2 YOR019W ATG21 ATG29 ICY2 ATG11 ATG13	23	29
nucleotide metabolic process [GO:0009117]	1.2E-09	PET9 COR1 ATP1 ATP3 ATP16 COX9 INH1 SDH4 RAV2 ATP5 ATP17 URC2 QCR7 APA2 RIP1 PRS2 QCR6 PMC1 COX4 COX13 QCR9 QCR10 COX6 QCR8 RAV1 ATP2 ATP7 MDH1 SDH3 SDH1 SDH2 COX12 ATP14 COX8 ATP18 STV1 COX7 COX5A CYT1 ATP4 ATP15 ATP20 QCR2	43	78
cofactor metabolic process [GO:0051186]	1.8E-09	COR1 ATP3 COX9 INH1 SDH4 ATP5 ATP17 QCR7 COX4 BIO2 QCR8 ATP7 MDH1 SDH2 CIT1 ISU1 QCR2	17	19
sporulation resulting in formation of a cellular spore [GO:0030435]	6.9E-09	OAF1 SEF1 GIP1 REG2 MUM2 SPO23 UBX7 SNT1 BDF2 FMP45 ADY3 NTH1 REG1 DOA4 SPO71 SAC6 RMD5 DON1 ZIP1 XRS2 DIT1 RMD6 GPA2 MEI4 GIP2 SHC1 SPR6 DMC1 PES4 RMD8 RIM8 ATG1 IME4 HOS2 MDS3 SEC9 GSC2 SPR3 AMA1 SPO11 SPS100 SSP1 PIG2 SPO22 SGA1 SLM1 HOS4 ATG32 BBC1 LOH1 UBX6 GSM1 YAK1 TPK1 CDC16 SPO14 TGL4 SPO75 KNS1 OSW2 CRR1 GSY2 CDC25 BDF1 SMA2 MSC1 SPO20 RIM11 FKS3 SPS18 SHE4 MPC54 GAC1 SPR1 SMA1 LGE1 CSM4 CSR2	78	185
piecemeal microautophagy of nucleus [GO:0034727]	4.1E-07	ATG8 ATG14 ATG12 ATG15 ATG9 CIS1 ATG18 ATG1 VAM7 ATG7 SNX4 ATG10 ATG17 ATG16 ATG4 ATG2 VAM3 ATG21 ATG29 ATG11 ATG13	21	32
ion transport [GO:0006811]	4.5E-06	PET9 COR1 ATP1 ATP3 YRO2 FTH1 YBR241C PCA1 HSP30 ATP16 MCH1 COX9 INH1 YDL206W ENA5 SDH4 ATP5 TIM11 ATP17 YDR506C QCR7 FIT1 RIP1 FET5 QCR6 PMC1 PMR1 COX4 COX13 QCR9 DUR3 ARN1 QCR10 COX6 CTR2 TOK1 KHA1 TRK1 QCR8 ATP2 ATP7 MDH1 SDH3 SDH1 ZRT3 SDH2 FRE6 COX12 FRE8 NHA1 ATP14 COX8 ATP18 STV1 COX7 COX5A ATO2 MAM3 ATP19 CYT1 FRE3 FIT3 FRE5 PMA2 YPL060W ATP4 ATP15 ATP20 QCR2	69	181
electron transport chain [GO:0022900]	1.5E-05	COR1 GRX1 TRX3 SDH4 ARH1 GRX2 QCR7 RIP1 QCR6 OLE1 QCR9 QCR10 YJL045W QCR8 SDH3 SDH1 SDH2 FRE6 FRE8 YLR164W ERO1 YMR118C CYT1 FRE3 FRE5 QCR2	26	51
ATP biosynthetic process [GO:0006754]	3.3E-05	ATP1 ATP3 PCA1 ATP16 ENA5 ATP5 TIM11 ATP17 PMC1 PMR1 ATP2 ATP7 ATP14 ATP18 ATP19 PMA2 ATP15 ATP20	18	31
CVT pathway [GO:0032258]	7.2E-05	ATG8 ATG14 ATG12 ATG20 ATG9 ATG18 ATG1 ATG7 SNX4 TAX4 IRS4 ATG10 ATG16 ATG4 ATG2 ATG19 ATG21 ATG11 ATG13	19	35
regulation of transcription, DNA-dependent [GO:0006355]	4.2E-04	OAF1 HAP3 SEF1 EDS1 ERT1 THI2 MAL33 SRD1 HMRA2 MBP1 ARO80 CAD1 URC2 SWI4 YER184C GAT1 RIM15 PDR1 HSF1 TOS8 CUP2 HOS2 HAP2 MGA1 STB5 SKN7 CST6 XBP1 MET28 YAP5 GSM1 YJL206C RGT1 HAP4 YKL222C GAT3 RFX1 HAP1 YAP1 ARG81 MAC1 GAT2 CAT8 MET4 HAL9 CIN5 AZF1 SFL1 YRR1 YRM1 SAS5 PIP2 RDR1 ECM23 RDS2 SPP1 HAA1 SUA7	58	166
energy reserve metabolic process [GO:0006112]	5.0E-04	TPS1 NTH1 REG1 GLC3 GIP2 PCL10 PIG2 PCL7 YAK1 TPK1 GLG1 GSY2 PGM2 GAC1 GPH1	15	28
negative regulation of gluconeogenesis [GO:0045721]	8.6E-04	VID24 RMD5 UBC8 VID30 VID28 FYV10 GID8	7	9

aerobic respiration [GO:0009060]	1.1E-03	PET9 COR1 ETR1 DLD1 PET100 COX20 QCR7 RIP1 RPO41 QCR6 COX4 COX13 SHY1 QCR9 QCR10 QCR8 CBP1 MDH1 MBR1 PET10 HAP1 RSF1 AAC1 ISF1 PAH1 NCA2 QCR2	27	66
fatty acid metabolic process [GO:0006631]	1.4E-03	OAF1 AGP2 FAA2 YAT2 POX1 POT1 MGA2 SPT23 FOX2 ECI1 CAT2 IZH2 CRC1 DCI1 FAA1 PIP2	16	33
ubiquitin-dependent endocytosis [GO:0070086]	1.7E-03	RSP5 ROG3 RIM8 ART5 ALY2 ALY1 LDB19 CSR2	8	12
response to toxin [GO:0009636]	2.0E-03	PRX1 AAD3 AAD4 HSP31 GTO1 AAD10 CYT2 ECM4 SUL2 GTO3 AAD14 IZH2 GRE2 PDR10 FRE3	15	31
cristae formation [GO:0042407]	3.3E-03	TIM11 FCJ1 UPS2 ATP14 ATP20	5	6
signal transduction [GO:0007165]	4.6E-03	PSK1 GPB2 PKC1 IRA1 COS111 ERT1 GPR1 SNF3 MTH1 SAC7 GPA2 STE2 CMK1 GPG1 SIP2 GPA1 PKP1 CYR1 TOR1 TUS1 PSK2 IRA2 RTS1 BAG7 GSP2 GPB1 RDS2 MKK2 PLC1 TIP41	30	82
autophagic vacuole assembly [GO:0000045]	5.0E-03	ATG8 ATG12 ATG9 ATG1 ATG17 ATG4 ATG2	7	11
macroautophagy [GO:0016236]	6.0E-03	VPS15 ATG14 ATG12 PTC6 ATG20 ATG18 ATG7 ATG10 ATG16 ATG21 ATG29 ATG13	12	25
DNA metabolic process [GO:0006259]	6.6E-03	MMS4 DMC1 SPO11 CST6 TOP3 MEC3 TOP2 RAD1	8	14
cellular ion homeostasis [GO:0006873]	6.6E-03	FYV5 CRD1 MDM31 PTK2 GIS4 SKY1 MDM32 HRK1	8	14
lipid metabolic process [GO:0006629]	7.0E-03	ATG15 YDL109C YDR444W FAA2 YAT2 OLE1 POX1 POT1 TAX4 FOX2 IRS4 TGL4 ECI1 CAT2 TGL3 IZH2 DCI1 DGA1 FAA1 GDE1 PGC1 PLC1	22	57
negative regulation of Ras protein signal transduction [GO:0046580]	9.4E-03	GPB2 IRA1 AVO1 IRA2 GPB1	5	7

Table S9 Fold changes of genes involved in fermentation and amino acid catabolism for the SOK2 overexpressing strain versus control.

Systematic name	Gene name	Functional Description	SOK vs Control
YJR155w	AAD10	Putative aryl-alcohol dehydrogenase	-4.17
YNL331c	AAD14	Putative aryl-alcohol dehydrogenase	-2.33
YCR107w	AAD3	Putative aryl-alcohol dehydrogenase	-1.57
YDL243c	AAD4	Putative aryl-alcohol dehydrogenase	-1.60
YGL256w	ADH4	alcohol dehydrogenase IV	3.10
YBR145w	ADH5	alcohol dehydrogenase V	1.57
YMR170c	ALD2	aldehyde dehydrogenase 2 (NAD+)	-1.82
YMR169c	ALD3	stress inducible aldehyde dehydrogenase	-11.49
YPL061w	ALD6	aldehyde dehydrogenase, cytosolic	1.69
YGL148w	ARO2	chorismate synthase	2.31
YBR249c	ARO4	2-dehydro-3-deoxyphosphoheptonate aldolase, catalyzes the first step in aromatic amino acid biosynthesis	3.35
YGL202w	ARO8	Aromatic aminotransferase, expression is regulated by general control of amino acid biosynthesis	2.18
YHR137w	ARO9	Aromatic aminotransferase, catalyzes the first step of tryptophan, phenylalanine, and tyrosine catabolism	-2.05
YDR380w	ARO10	Phenylpyruvate decarboxylase, catalyzes the first specific step in the Ehrlich pathway	-3.02
YGR177C	ATF2	Alcohol acetyltransferase, forms volatile esters during fermentation	4.38
YCL064c	CHA1	L-serine/L-threonine deaminase, catalyzes the degradation of both L-serine and L-threonine	3.33
YDL174c	DLD1	D-lactate dehydrogenase, oxidizes D-lactate to pyruvate	-2.52
YEL071w	DLD3	D-lactate dehydrogenase	3.06
YEL066W	HPA3	D-Amino acid N-acetyltransferase, catalyzes N-acetylation of D-amino acids	2.29
YER086w	ILV1	Threonine deaminase, catalyzes the first step in isoleucine biosynthesis	1.97
YMR108w	ILV2	Acetolactate synthase, catalyses the first common step in isoleucine and valine biosynthesis	1.50
YLR355c	ILV5	Acetohydroxyacid reductoisomerase, mitochondrial protein involved in branched-chain amino acid biosynthesis	1.58
YCL009c	ILV6	Regulatory subunit of acetolactate synthase	1.62
YOR108w	LEU9	2-Isopropylmalate synthase, catalyzes the first step in the leucine biosynthesis pathway	2.32
YDR081c	PDC2	pyruvate decarboxylase regulatory protein	-1.51

YLR134w	PDC5	pyruvate decarboxylase, isozyme 2	-2.21
YGR087c	PDC6	pyruvate decarboxylase 3	-5.09
YPR026w	ATH1	acid trehalase, vacuolar	-3.40
YDR001c	NTH1	neutral trehalase (alpha,alpha-trehalase)	-2.20
YBR001c	NTH2	alpha,alpha-trehalase	-3.92
YKL127w	PGM1	phosphoglucomutase, minor isoform	3.03
YMR105c	PGM2	phosphoglucomutase, major isoform	-1.85
YBR126c	TPS1	alpha,alpha-trehalose-phosphate synthase, 56 KD subunit	-1.80
YDR074w	TPS2	alpha,alpha-trehalose-phosphate synthase, 102 KD subunit	-2.45
YMR261c	TPS3	alpha,alpha-trehalose-phosphate synthase, 115 KD subunit	-1.52
YML100w	TSL1	alpha,alpha-trehalose-phosphate synthase, 123 KD subunit	-2.77
YAL054c	ACS1	acetyl-CoA synthetase	-2.20
YLR153c	ACS2	acetyl-coenzyme A synthetase	2.17
YPL028W	ERG10	Acetyl-CoA C-acetyltransferase; involved in the first step in mevalonate biosynthesis	1.77
YML126C	ERG13	3-hydroxy-3-methylglutaryl-CoA synthase, involved in the second step in mevalonate biosynthesis	2.09
YIL160C	POT1	3-ketoacyl-CoA thiolase involved in beta-oxidation of fatty acids	-4.82