Practice

BIOINFORMATICS INSTITUTE

Transcriptomic Insights into *Saccharomyces cerevisiae*: Unraveling the Dynamics of Gene Expression During Anaerobic Fermentation

Popov Ilia

Abstract

This study employed RNA sequencing to investigate the dynamic alterations in the *Saccharomyces cerevisiae* transcriptome during a 30-minute fermentation period. Complementary DNA (cDNA) libraries were constructed from yeast RNA, and high-throughput sequencing provided insights into gene expression changes. Differential expression analysis revealed significant modifications in various biological processes, emphasizing the yeast's adaptive response to anaerobic conditions. Upregulated genes, including those involved in glycerol biosynthesis and glucose transportation, highlighted the crucial role of these processes in yeast survival and fermentation efficiency. Conversely, downregulated genes suggested shifts in protein synthesis, metabolic strategy, and cellular pH homeostasis.

Keywords: Yeasts, Differential expression, Fermentation

Introduction

Saccharomyces cerevisiae, a eukaryotic model organism, plays a pivotal role in various biotechnological processes, particularly industrial fermentation (Parapouli *et al.* 2020). Understanding the dynamic changes in its transcriptome during fermentation is essential for optimizing industrial processes and harnessing yeast's metabolic capabilities. RNA sequencing (RNA-seq) has emerged as a powerful tool for gene expression analysis (Wang *et al.* 2009). In this study, the transcriptomic response of *S. cerevisiae* to a 30-minute fermentation period was investigated, aiming to elucidate the intricate alterations in gene expression under anaerobic conditions. This study contributes to the broader understanding of yeast biology and its adaptive responses to environmental changes, specifically during fermentation.

Materials and methods

Data

Yeast's RNA was used to construct complementary DNA (cDNA) libraries for subsequent high-throughput sequencing. Single-ended reads from two distinct experimental conditions were obtained from the Sequence Read Archive (SRA) FTP. Each condition involved two replicates: SRR941816 and SRR941817 for the pre-fermentation state, and SRR941818 and SRR941819 for the condition 30 minutes post-fermentation onset. Quality assessment was performed using FastQC v. 0.12.0 (noa ????a) and MultiQC v. 1.21 (Ewels *et al.* 2016). The reference genome and its annotation used in this study were sourced from the NCBI FTP, specifically the *Saccharomyces cerevisiae* strain S288c, assembly R64.

RNA-seq Alignment and Quantification

The RNA-seq alignment and quantification protocol followed a pipeline, utilizing tools such as HISAT2 v. 2.2.1 (noa ????d) and featureCounts v. 2.0.1 (Liao *et al.* 2014) from the subread package. Initially, the 'hisat2-build' command was used to construct a reference genome index, which facilitated subsequent alignment of reads in single-ended mode with default settings. The resulting alignments were organized using the samtools v. 1.19.2 utility. For concurrent use of featureCounts, the annotation file was formatted into GTF (Gene Transfer Format) using the gffread tool. Mapped reads were then quantified for genomic features using the featureCounts utility with default parameters.

Differential Expression Analysis

Gene counts obtained through the HISAT2 + featureCounts pipeline were analyzed with the DESeq2 package (noa ????b) for R. The main parameters for determining significance were the adjusted p-value (adjusted with the Benjamini-Hochberg method) and the absolute log fold change, with thresholds set at 0.05 and 2, respectively. Genes with a log fold change greater than 2 or less than -2 were considered upregulated and downregulated, respectively.

Functional Annotation

To determine the primary biological processes involving differentially expressed genes, goSlimMapper (noa ????c) and shinyGO v0.80 (Ge *et al.* 2020) websites were used, which utilize the Gene Ontology (GO) biological process database. For GO annotation KEGG enrichment, the Yeast GOSlim:process Set was specified, which includes all terms. Upregulated and downregulated genes were functionally annotated separately. The significance threshold for GO terms was set at FDR < 0.05.

Results

Alignment and Quantification

An assessment of the sequencing data quality was performed using the FastQC utility, which provided a distribution of quality

and other statistical measures for each sample. The average quality (Phred-score) across all reads and samples was found to be around 35. The count of sequence reads for each sample was depicted in Figure 1. Subsequent alignment of the reads to the reference sequence using HISAT2 revealed an alignment rate of 94-96% (Table 1). The featureCounts analysis showed that about 74% of all alignments were successfully mapped to a specific feature from the annotation (Table 2). Interestingly, the total number of alignments exceeded the total number of reads for each sample, indicating that some reads were aligned to the reference multiple times.

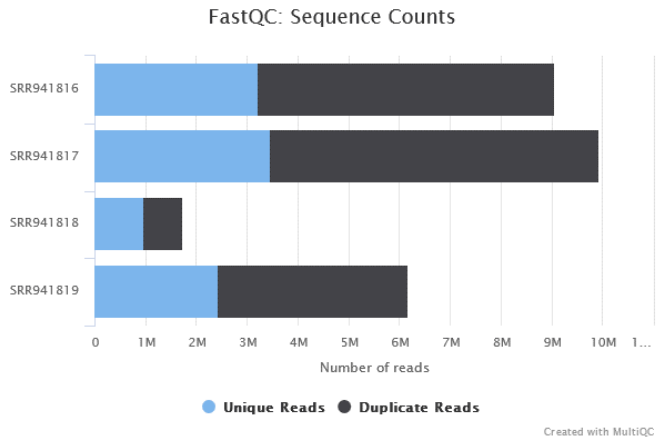

Figure 1 The number of sequences counts for each read.

Table 1 Summary of alignment results using the HISAT2 utility.

Sample	Total reads number	Reads aligned 1 time	Alignment rate
SRR941816	9043877	7930593	94.25
SRR941817	9929568	8645384	94.91
SRR941818	1721675	1508002	96.22
SRR941819	6172452	5368133	96.28

Table 2 Summary of featureCount results.

Sample name	Total alignments	Assigned alignments
SRR941816	9773838	7285693 (74.5
SRR941817	10832704	7986987 (73.7
SRR941818	1885543	1406729 (74.6
SRR941819	6800272	4994723 (73.4

Differential expression

The identification of significantly differentially expressed genes prior to and following fermentation was achieved and graphically represented in a heatmap (Figure 2). Utilizing the HISAT2 + featureCounts + DESeq2 pipeline for analysis, a total of 1905 genes exhibited significant differential expression (with an absolute log2FoldChange > 2 and $padj < 0.05$). Of these, 987 genes were upregulated, and 918 genes were downregulated. The distribution of significantly differentially expressed genes was depicted in a volcano plot (Figure 3), with the top 5 upregulated and downregulated genes highlighted and enumerated in Table 3.

Functional annotation

Gene Ontology (GO) terms associated with upregulated genes were predominantly related to ribosome synthesis (Figure 4). In the context of ribosome biogenesis, upregulated genes were indicated in the visualization of the KEGG pathway (Figure 5). Conversely, the annotations of downregulated genes were primarily associated with various metabolic pathways, with a particular emphasis on carbohydrate and pyruvate metabolism (Figure 6). For the citrate and pyruvate metabolic pathways, downregulated genes were indicated in the visualization of the KEGG pathway (Figure 7). Figures 4b and 6b depict the connection network for annotated GO biological processes for upregulated and downregulated genes, respectively.

Discussion

During this study, the analysis of *Saccharomyces cerevisiae* RNA sequencing results included the evaluation of read quality, alignment to the reference genome, and assignment to genomic features. Subsequently, a differential expression analysis was conducted to identify changes in the yeast transcriptome following a 30-minute fermentation period. The shift to anaerobic conditions led to significant alterations in gene expression across various biological processes, with a substantial number exhibiting low p-values. Table 3 details the top five upregulated and downregulated genes, along with their corresponding products.

The upregulated genes identified are primarily involved in glycerol biosynthesis and glucose transportation, which are crucial for yeast survival and fermentation efficiency under anaerobic conditions.

For instance, the gene YER062C (HOR2), which codes for DL-glycerol-3-phosphate phosphatase, showed the highest upregulation. This enzyme plays a key role in glycerol biosynthesis, a process that is vital for maintaining cellular osmotic balance and redox potential during fermentation (Byrne and Wolfe 2005). Similarly, YDR536W (STL1), a glycerol proton symporter of the plasma membrane (Ferreira *et al.* 2005), was also significantly upregulated, suggesting an increased need for glycerol transport under these conditions.

The gene YHR094C (HXT1), a low-affinity glucose transporter of the major facilitator superfamily (Byrne and Wolfe 2005; Mosley *et al.* 2003), was also significantly upregulated. This indicates an increased need for glucose transport, which is consistent with the shift to fermentation, a process that relies heavily on glucose. The upregulation of YNL065W and YNL112W (DBP2), both involved in RNA processing and transportation (Gbelska *et al.* 2006; Tkach *et al.* 2012; Lai *et al.* 2019; Cloutier *et al.* 2012), further supports the notion of a dynamic transcriptomic response to fermentation.

On the other hand, the downregulated genes seem to be involved in various biological processes. For example, YLR327C, a protein of unknown function that associates with ribosomes (Tkach *et al.* 2012; Byrne and Wolfe 2005), was downregulated. This could indicate a reduced need for protein synthesis machinery under anaerobic conditions. Similarly, YKR097W (PCK1), which codes for phosphoenolpyruvate carboxykinase, an enzyme involved in gluconeogenesis (Haurie *et al.* 2001), was also

downregulated, possibly reflecting a shift in metabolic strategy from glucose production to consumption.

YMR081C (ISF1) and YDR342C (HXT7), both involved in glucose metabolism and transportation (Byrne and Wolfe 2005), were also downregulated, further supporting the idea of a metabolic shift. Lastly, YCR021C (HSP30), a negative regulator of the H(+)-ATPase Pma1p (Wu *et al.* 2000), was downregulated, which could suggest changes in cellular pH homeostasis.

In conclusion, this study has illuminated the dynamic nature of the yeast transcriptome during fermentation, identifying key genes that are upregulated and downregulated in response to anaerobic conditions. The findings enhance our understanding of yeast biology, particularly the metabolic shifts that occur during fermentation. The identified genes and their functions provide a valuable resource for future reference and could have practical implications for improving industrial fermentation processes. This work represents a significant contribution to our knowledge of *Saccharomyces cerevisiae* and its response to fermentation conditions. It underscores the power of RNA sequencing in revealing the intricate changes in gene expression that underpin the yeast's ability to adapt to different environmental conditions.

Supplementary materials

Lab journal can be found in GitHub repository.

Literature cited

????a. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data.

????b. DESeq2.

????c. Gene Ontology Slim Term Mapper | SGD.

????d. HISAT2.

Byrne KP, Wolfe KH. 2005. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Research. 15:1456–1461.

Cloutier SC, Ma WK, Nguyen LT, Tran EJ. 2012. The DEAD-box RNA helicase Dbp2 connects RNA quality control with repression of aberrant transcription. The Journal of Biological Chemistry. 287:26155–26166.

Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 32:3047–3048.

Ferreira C, van Voorst F, Martins A, Neves L, Oliveira R, Kielland-Brandt MC, Lucas C, Brandt A. 2005. A member of the sugar transporter family, Stl1p is the glycerol/H+ symporter in Saccharomyces cerevisiae. Molecular Biology of the Cell. 16:2068–2076.

Gbelska Y, Krijger JJ, Breunig KD. 2006. Evolution of gene families: the multidrug resistance transporter genes in five related yeast species. FEMS yeast research. 6:345–355.

Ge SX, Jung D, Yao R. 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 36:2628–2629.

Haurie V, Perrot M, Mini T, Jenö P, Sagliocco F, Boucherie H. 2001. The transcriptional activator Cat8p provides a major contribution to the reprogramming of carbon metabolism during the diauxic shift in Saccharomyces cerevisiae. The Journal of Biological Chemistry. 276:76–85.

Lai YH, Choudhary K, Cloutier SC, Xing Z, Aviran S, Tran EJ. 2019. Genome-Wide Discovery of DEAD-Box RNA Helicase Targets Reveals RNA Structural Remodeling in Transcription Termination. Genetics. 212:153–174.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics (Oxford, England). 30:923–930.

Mosley AL, Lakshmanan J, Aryal BK, Ozcan S. 2003. Glucose-mediated phosphorylation converts the transcription factor Rgt1 from a repressor to an activator. The Journal of Biological Chemistry. 278:10322–10327.

Parapouli M, Vasileiadis A, Afendra AS, Hatziloukas E. 2020. Saccharomyces cerevisiae and its industrial applications. AIMS Microbiology. 6:1–31.

Tkach JM, Yimit A, Lee AY, Riffle M, Costanzo M, Jaschob D, Hendry JA, Ou J, Moffat J, Boone C *et al*. 2012. Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. Nature Cell Biology. 14:966–976.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews. Genetics. 10:57–63.

Wu K, Dawe JH, Aris JP. 2000. Expression and subcellular localization of a membrane protein related to Hsp30p in Saccharomyces cerevisiae. Biochimica Et Biophysica Acta. 1463:477–482.

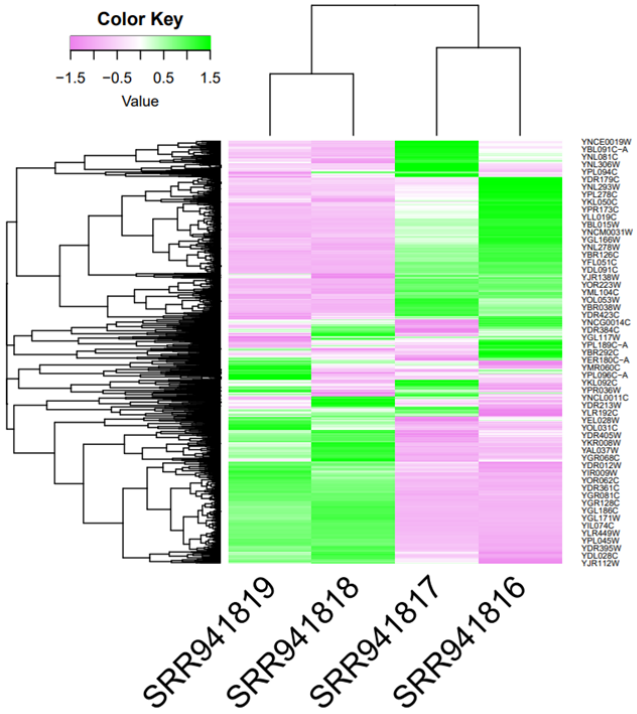

Figure 2 Heatmap for significantly differentially expressed genes before (SRR941816 & SRR941817) and after fermentation (SRR941818 & SRR941819). Gradient from violet to green in cells of heatmap stands for log2FoldChange value (from negative to positive, respectively). Data were visualized using R DESeq2 package.

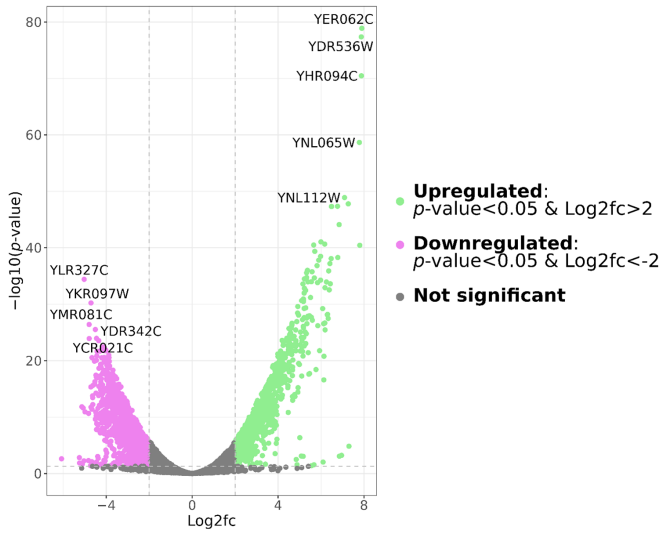

Figure 3 Volcanoplot. Top 5 upregulated and downregulated genes are marked.

.

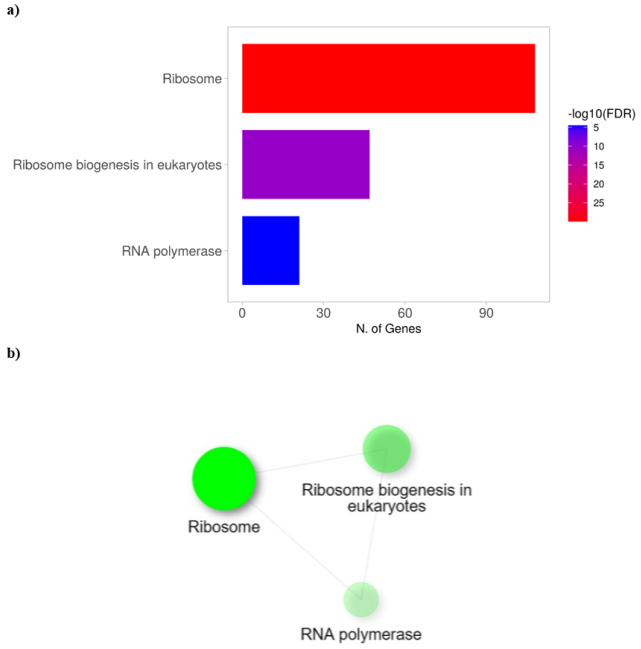

Figure 4 KEGG enrichment analysis (a) and GO network of biological processes (b) for upregulated genes. Visualization were performed on websource shinyGO v0.80

.

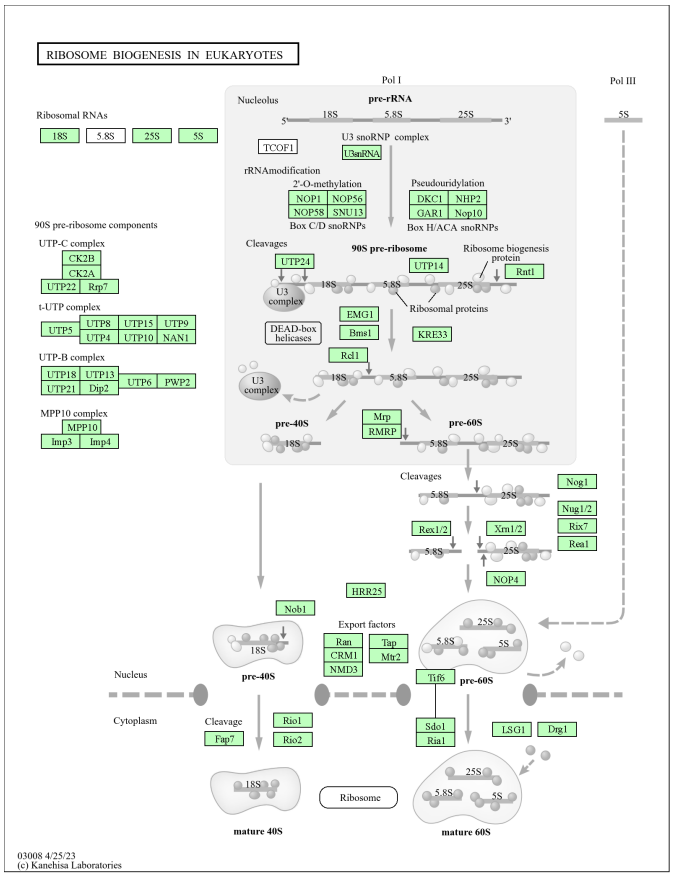

Figure 5 KEGG pathway for ribosome biogenesis in eucaryotes with presentation of upregulated genes in green blocks. Visualization was performed on websource shinyGO v0.80

.

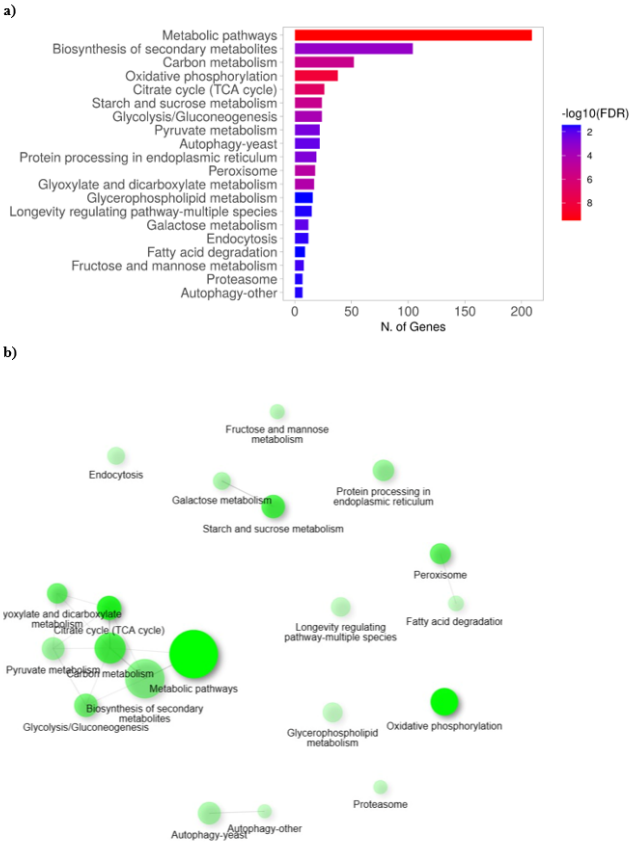

Figure 6 KEGG enrichment analysis (a) and GO network of biological processes (b) for downregulated genes. Visualization was performed on websource shinyGO v0.80

.

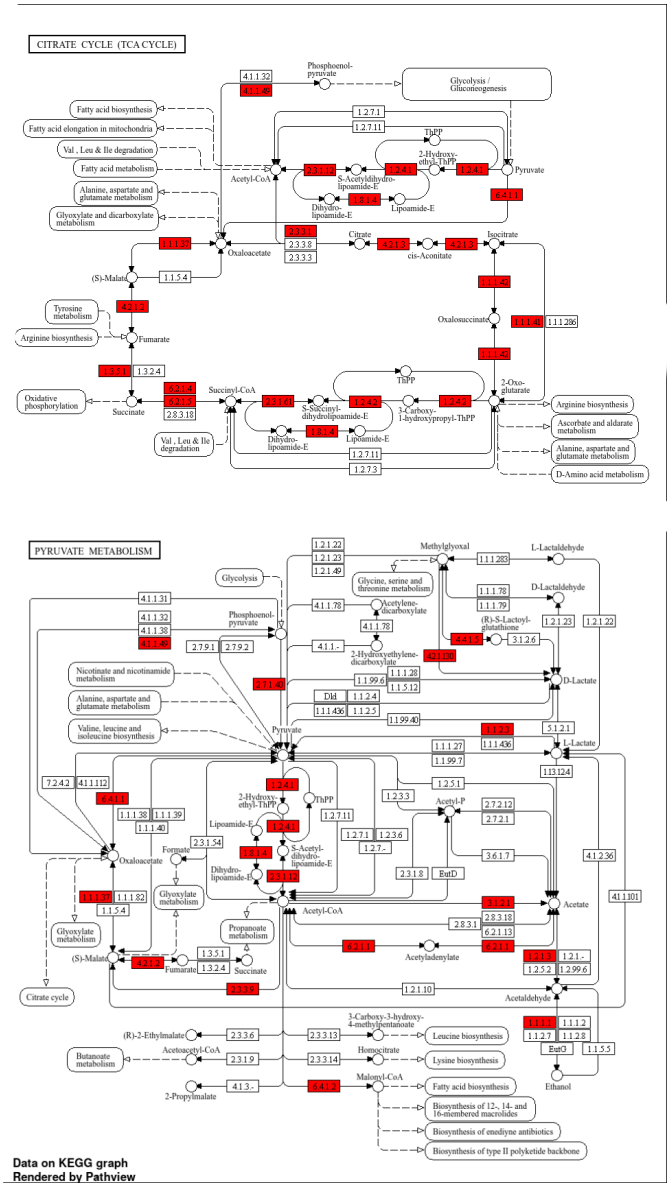

Figure 7 KEGG pathway for citracte cycle and pyruvate metabolism with presentation of downregulated genes in red blocks. Visualization was performed on websource shinyGO v0.80

.

Table 3 Top Five Upregulated and Downregulated Genes at 30 Minutes of Fermentation.

Gene ID	log2FoldChange	p_{adj}	Product
Upregulated genes			
YER062C	7.899115	1.327198e-79	DL-glycerol-3-phosphate phosphatase involved in glycerol biosynthesis
YDR536W	7.873318	4.343371e-78	Glycerol proton symporter of the plasma membrane
YHR094C	7.880683	3.396192e-71	Low-affinity glucose transporter of the major facilitator superfamily
YNL065W	7.784312	2.181703e-59	Plasma membrane transporter of the major facilitator superfamily
YNL112W	7.093562	1.272895e-49	ATP-dependent RNA helicase of the DEAD-box protein family
Downregulated genes			
YLR327C	-5.026001	3.894650e-35	Protein of unknown function that associates with ribosomes; protein abundance increases in response to DNA replication stress
YKR097W	-4.708442	5.900998e-31	Phosphoenolpyruvate carboxykinase
YMR081C	-4.794479	3.852522e-27	Serine-rich, hydrophilic protein; overexpression suppresses growth defects of hap2, hap3, and hap4 mutants; expression is under glucose control
YDR342C	-4.507653	2.983276e-26	High-affinity glucose transporter; member of the major facilitator superfamily
YCR021C	-4.451892	1.118774e-24	Negative regulator of the H(+)-ATPase Pma1p; stress-responsive protein