How does the yeast feel just before baking?

Sobolev A., Kupaeva D.

Abstract

Yeast dough is an integral part of the world's culinary arts. Loosening of the yeast dough occurs due to the release of carbon dioxide, ethanol and other metabolites by the yeast into the dough. When yeast gets into a glucose-rich environment, it switches to active metabolism of simple sugars and inactivates gluconeogenesis. The most pronounced process associated with switching metabolic pathways is a sharp activation of protein biosynthesis and associated processes.

In this project, we focused on changes in differential gene expression in *Saccharomyces cerevisiae* strain S288c upon fermentation. Using raw RNA sequencing data, we aligned it to the reference genome with HISAT2, assigned alignments to the genomic features using featureCounts and conducted a differential gene expression analysis using the Deseq2 package. As a result, it was confirmed that in anaerobic conditions yeast upregulates a group of proteins involved in protein synthesis, oxidative stress response and suppress gluconeogenesis enzymes, switching its metabolism of carbohydrates to anaerobic glycolysis and fermentation.

Introduction

Fermentation of dough with yeast is an important step in the baking process. The fundamental fermentation process of dough has been used by humans for at least 6,000 years. During the fermentation of the dough, yeast releases 2 main metabolites into the dough - carbon dioxide and ethanol[1] . In addition, they release metabolites - organic acids, glycerin and aromatic compounds, which affect the taste, aroma and shelf life of the dough. Currently, a large number of technological nuances are known that affect the taste, aroma and efficiency of dough fermentation. These include salt and sugar content, temperature, freezing, and yeast strain[2]. Fermentation also leads to a dough with less flow and lower extensibility that breaks more easily under stress and strain[3]. Investigation of metabolic changes happening in yeast on fermentation not only can result in better understanding of yeast metabolism, but also in improved organoleptic properties of the bread.

However, yeast physiology and process of fermentation in the absence of water, termed solid-state fermentation (SSF), has not been well-studied yet despite wide usage of this process in cheese and cocoa production in addition to dough fermentation[4]. Despite the existence of a huge number of articles devoted to the peculiarities of yeast metabolism and

fermentation processes, researchers still do not understand the full picture of the complexity of regulatory gene networks[5]. It is known that yeast cells regulate a variety of cellular activities in accordance with the levels of glucose found outside and inside the cell. For instance, the Snf3 / Rgt2 signaling pathway is a sensory cascade present in yeast for detecting extracellular glucose levels[6]. Moreover, on different fermentation steps yeast cells respond to osmotic stress, induce genes involved in amino acids metabolism and activate pathways associated with starvation and stress responses[4]. All of the above information indicates that one should estimate gene expression of the yeast to deeply understand its metabolic processes.

The best tool that can be used to identify genes involved in various metabolic processes is experiments on the analysis of differentially expressed genes based on high-throughput RNA sequencing[7]. This research method allows you to obtain a list of genes that differ at different stages of the life cycle of the studied organisms, as well as to compare quantitatively the level of their expression. The analysis usually involves mapping the readings of the experiment to a reference genome or transcript and subsequent qualitative and quantitative analysis using the appropriate packages[8]. The most widely used and effective mapping tools are Bowtie2[9], TopHat[10], BWA[11], pseudo-alignment tools like salmon[12], and the expression analysis packages are deseq2[13] and limma[14]. The ease of use of the method, combined with the wide possibilities of using the algorithms, make DGE a universal tool of work.

Methods

We studied the changes in RNA expression levels of the yeast before and during fermentation. Raw Illumina HiSeq 2000 data was obtained using SRA FTP server and presented as four samples: two biological replicates of yeast before (replicate 1: SRR941816, replicate 2: SRR941817) and after fermentation (replicate 1: SRR941818, replicate 2: SRR941819). We used R64 assembly of *Saccharomyces cerevisiae* strain S288c genome as a reference and downloaded it from the NCBI FTP server as well as the genome annotation file. Indexing of the genome and alignment of the raw reads to the reference was performed using HISAT2[15] in single-end mode with default parameters. We used gffread to convert GFF annotation to GTF. Assignment of aligned reads to genomic features and quantification was conducted using featureCounts[16] with default parameters. Differential gene expression and corresponding metrics were calculated using Deseq2[13] (we used R script available here). Resulting file containing differentially expressed genes (DEG) sorted by adjusted p-values was filtered to study only top 50 DEGs. Finally, we mapped these genes to corresponding gene ontology (GO) terms using GO Slim Mapper website to investigate their function. Illustration was prepared with the help of packages gplots and EnhancedVolcano.

Results

After exploratory analysis of raw data and alignment to the reference, we summarized all obtained information in Table 1. Decent alignment rate in every case and low percentage of ambiguously aligned reads indicate that all procedures were done correctly.

SRR ID	Number of reads	Reads aligned 1 time	Reads aligned >1 time	Unmapped reads	Alignment rate
SRR941816	9043877	7930595 (87.69%)	600310 (6.64%)	512972 (5.67%)	94.33%
SRR941817	9929568	8645392 (87.07%)	778968 (7.84%)	505208 (5.09%)	94.91%
SRR941818	1721675	1508003 (87.59%)	148595 (8.63%)	65077 (3.78%)	96.22%
SRR941819	6172452	5368136 (86.97%)	574634 (9.31%)	229682 (3.72%)	96.28%

Table 1. Raw data information and HISAT2 alignment results.

Next, we assigned aligned reads to the genomic features using featureCounts. As a result, most of the genes were assigned to the features, with some of the fragments left unassigned due to mapping to multiple positions, absence of annotated features or because they were unmapped at all (Fig. 1). Differences in the counts on the plot are due to the different number of reads in raw data files (See Table 1).

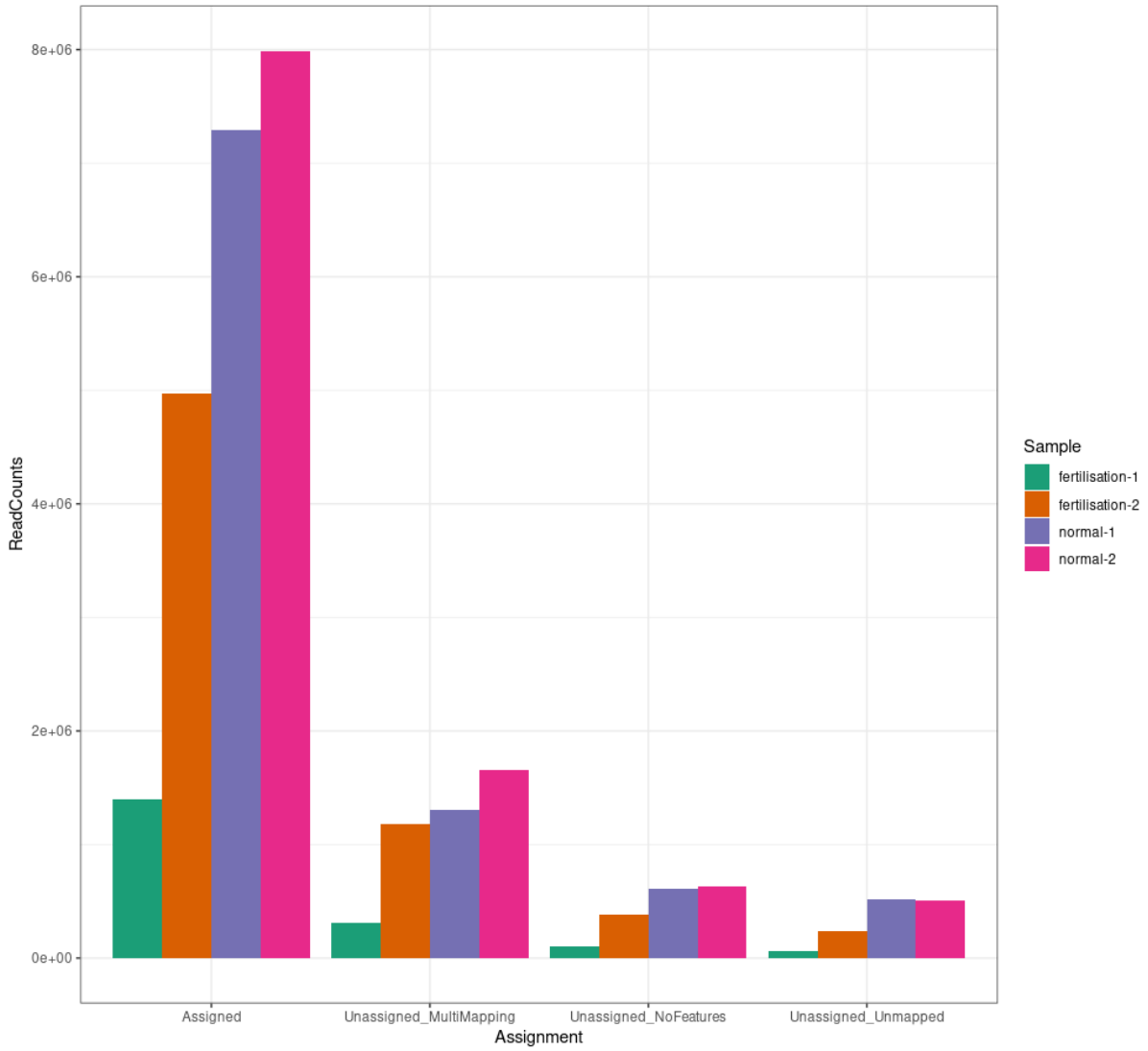

Fig. 1. FeatureCounts results. Different colors refer to the samples.

As a result of gene expression analysis, we found 3450 differentially expressed genes (p-value < 0.05). About half of them are expressed only before the beginning of fermentation of the dough, the other half - only after the start of fermentation.

Using the heatmap (Fig. 2), we can visually analyze the level of differences between samples. Visually, on the map and the accompanying tree, 5 large groups of genes are distinguished, differing in the level of gene expression among themselves. The first group - including YER160C (Retrotransposon TYA Gag) and D (GUC) L1 (aspartate transport RNA) - is expressed unevenly, the level of gene expression differs in different replicates of samples of the same type. The next group - which includes YDL078C (peroxisomal malate dehydrogenase), YER155C (rho GTPase activating protein (RhoGAP)), among others - is expressed under normal conditions and is not expressed after the start of fertilization. The next group - an example of a gene is YER043C (S-adenosyl-L-homocysteine hydrolase) is expressed uniformly. Then there is a large group of genes that are expressed only during fermentation. This group includes YHR070W (tRNA (m (1) G37) methyltransferase), YPL183C (WD40 domain-containing protein involved in endosomal recycling). Finally, the latter group is intensely expressed in only one of 2 yeast replicates under normal conditions.

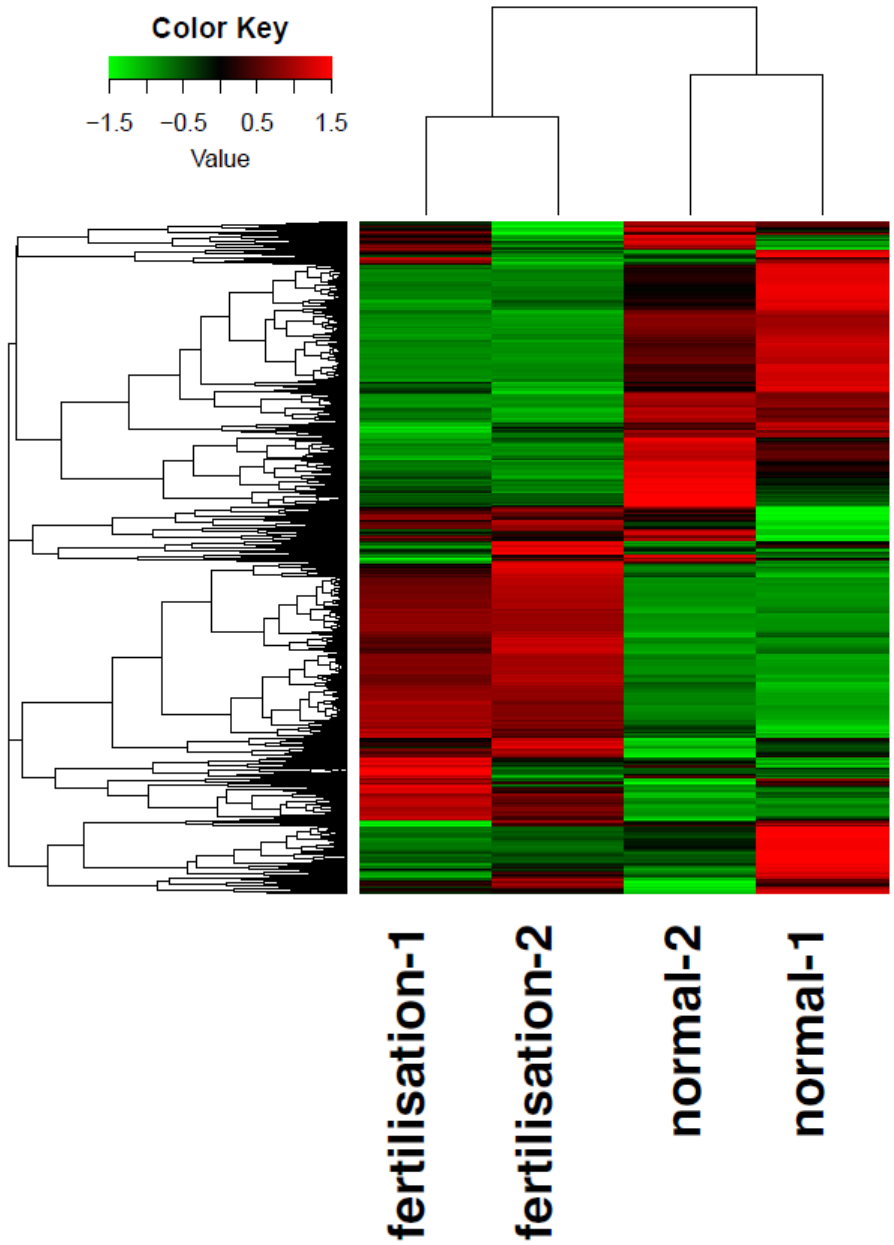

Fig. 2. Heatmap of differential expression genes.

We also built a volcano plot to visually assess the spread in the level of gene expression and the statistical significance of the differences.

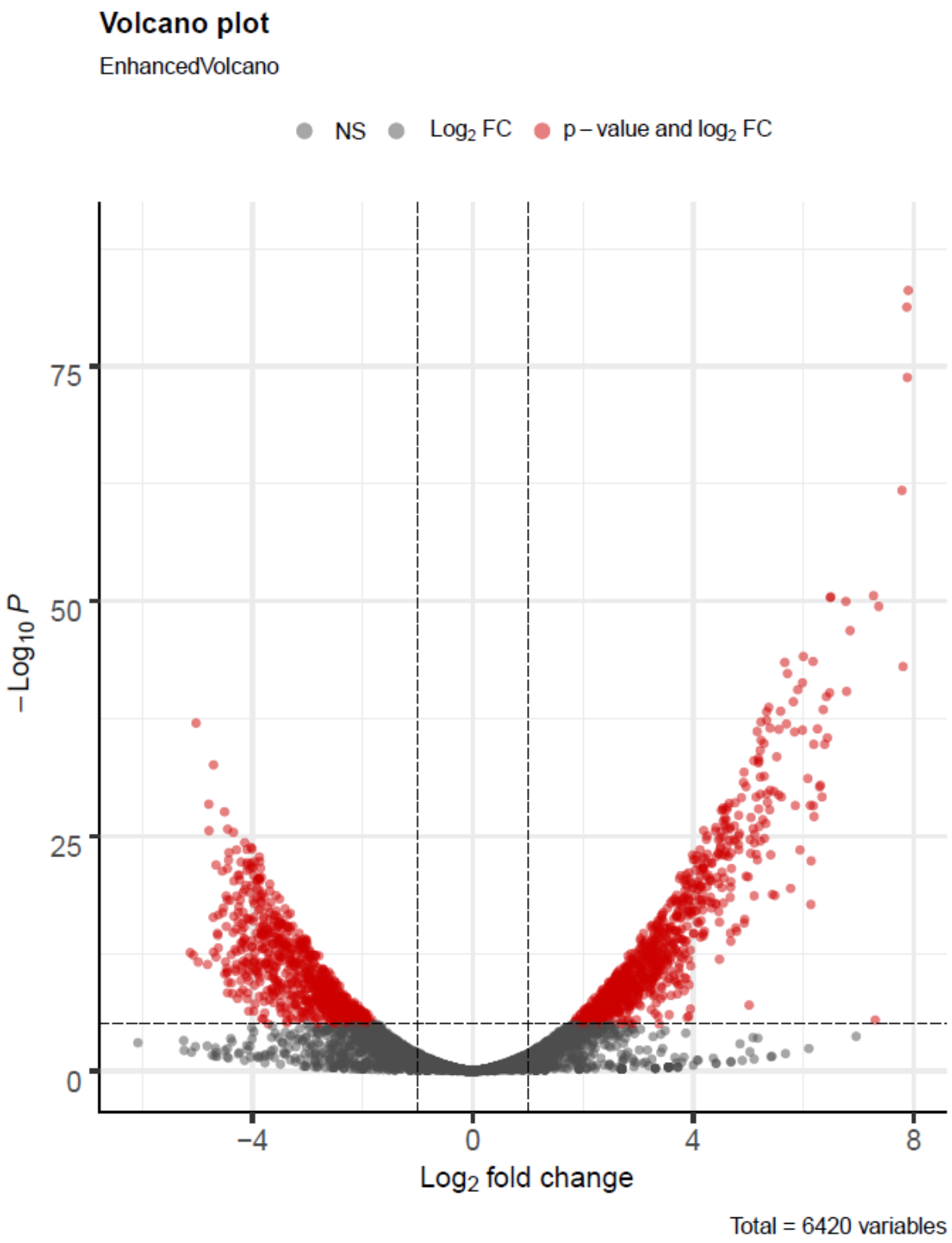

Fig. 3. Volcano plot of all analysed genes

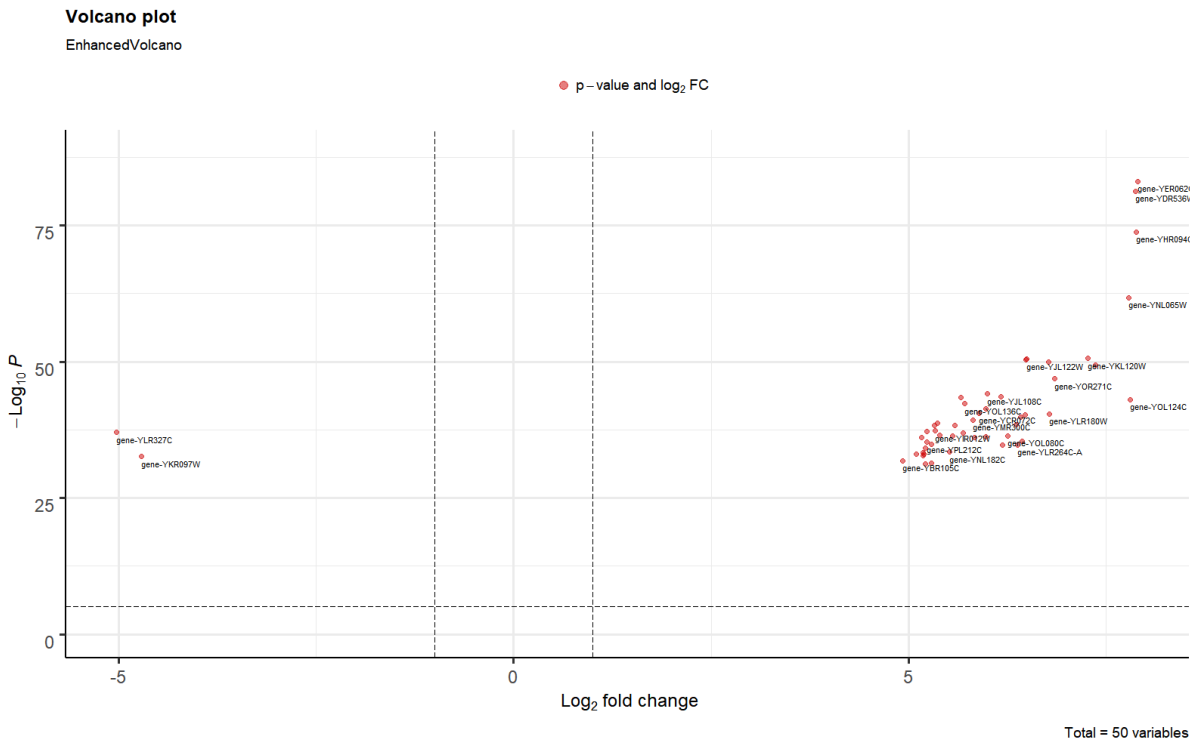

Figure 4. Volcano plot of top-50 most expressed genes.

In Figure 4, we see that the genes that differ the most in terms of the level of expression are those specific for fermentation. Only 2 genes that differ significantly in their level of expression are silenced during fermentation. Annotation of these genes is shown in Table 2.

Table 2. Genes downregulated upon fermentation.

Gene	Function
YKR097W	Phosphoenolpyruvate carboxykinase, key enzyme in gluconeogenesis, catalyzes early reaction in carbohydrate biosynthesis, glucose represses transcription and accelerates mRNA degradation, regulated by Mcm1p and Cat8p, located in the cytosol
YLR327C	Protein of unknown function that associates with ribosomes; protein abundance increases in response to DNA replication stress; TMA10 has a paralog, STF2, that arose from the whole genome duplication

The remaining 48 of the 50 genes, expressed very differently at different stages of the life cycle, belong to the fermentation phase. An annotation and a complete list of genes can be found in the supplementary materials.

These 48 genes belong to the following categories: genes associated with protein biosynthesis (rRNA processing, subunit biogenesis, ribosome assembly, transcription, translation regulation, nucleotide synthesis, and others). It also activates genes for ionic and transmembrane transport, genes associated with oxidative stress (YER062C), genes associated with fructose metabolism, and others.

Discussion

Upon fermentation, yeast undergo changes in both cellular state and metabolism. These adjustments imply changes in gene expression. In this study, we found out that 3450 genes change their expression in *Saccharomyces cerevisiae* in 30 minutes after the start of fermentation. Among them, we picked 50 genes with the least adjusted p-value and described GO terms associated with these genes (see Results). One of the two downregulated genes, YKR097W, also known as PCK1, is involved in carbohydrate metabolism. More specifically, PCK1 encodes phosphoenolpyruvate carboxykinase that serves as one of the main enzymes in gluconeogenesis, catalyzing the formation of phosphoenolpyruvate from oxaloacetate, with the release of carbon dioxide and GDP. Additionally, PCK1 is involved in utilization of excessive amounts of oxaloacetate. Downregulation of this enzyme likely indicates that the yeast cell is supplied with enough glucose and does not need gluconeogenesis to produce additional amounts of energy since glycolysis and gluconeogenesis are reciprocally regulated in the cells[17]. Moreover, decreased expression of PCK1 can be explained by the fact that oxaloacetate is not included in metabolic processing of glucose in fermentation, and, therefore, PCK1 do not need to turn it into phosphoenolpyruvate at all.

The most prominent group that undergoes up-regulation during dough fermentation are genes associated with protein synthesis. Their presence is explained by the fact that after the yeast enters the dough, their metabolism switches to the active use of glucose, and, accordingly, the cells accumulate the necessary enzymes for glucose metabolism. The response of cells to the presence of glucose is confirmed by the expression of genes from the carbohydrate transport (GO: 0008643) group: YHR094C (low affinity glucose transporter), YDR536W (glycerol proton symporter)[18], both induced in the presence of glucose[19]. Also, the expression of genes for the metabolism of fructose - carbohydrate metabolic process (GO: 0005975) - is enhanced in the cell. In particular, YBR105C, YER062C, YOL136C appear in the cell, all of them are associated with the transport or metabolism of fructose. The expression of the metabolic switch protein of the glyoxylate cycle (YLR224W) is sharply increased.

Finally, the synthesis of a protein associated with osmotic or oxidative stress (YER062C[20], YDR536W[18]) is activated in yeast. This may be due to the evolutionary adaptation of baker's yeast to the conditions of dough fermentation: under dough conditions, the yeast is in solution with sugar, salt, and oxygen deficiency.

The upregulation of genes associated with iron metabolism is not entirely clear. The YOR271C gene is expressed more intensively after the start of fermentation. At the same time, it was shown for him that his expression and functioning does not correlate with the presence of oxygen[21]. The YOR271C gene is also potentially associated with iron, the function of which is not precisely known.

Citations

1. Struyf, N. *et al.* Bread Dough and Baker's Yeast: An Uplifting Synergy. *Compr. Rev. Food Sci. Food Saf.* **16**, 850–867 (2017).
2. Cho, I. H. & Peterson, D. G. Chemistry of bread aroma: A review. *Food Science and Biotechnology* vol. 19 575–582 (2010).
3. Rezaei, M. N., Jayaram, V. B., Verstrepen, K. J. & Courtin, C. M. The impact of yeast fermentation on dough matrix properties. *J. Sci. Food Agric.* **96**, 3741–3748 (2016).
4. Aslankoohi, E. *et al.* Dynamics of the Saccharomyces cerevisiae transcriptome during bread dough fermentation. *Appl. Environ. Microbiol.* **79**, 7325–7333 (2013).
5. Dashko, S., Zhou, N., Compagno, C. & Piškur, J. Why, when, and how did yeast evolve alcoholic fermentation? *FEMS Yeast Res.* **14**, 826–832 (2014).
6. Kayikci, Ö. & Nielsen, J. Glucose repression in Saccharomyces cerevisiae. *FEMS Yeast Res.* **15**, (2015).
7. Costa-Silva, J., Domingues, D. & Lopes, F. M. RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One* **12**, e0190152 (2017).
8. Zhang, Z. H. *et al.* A comparative study of techniques for differential expression analysis on RNA-Seq data. *PLoS One* **9**, e103207 (2014).
9. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
10. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
11. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
12. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
13. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
14. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
15. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
16. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
17. Jardón, R., Gancedo, C. & Flores, C.-L. The Gluconeogenic Enzyme Fructose-1,6-Bisphosphatase Is Dispensable for Growth of the Yeast Yarrowia lipolytica in Gluconeogenic Substrates. *Eukaryot. Cell* **7**, 1742 (2008).
18. Ferreira, C. *et al.* A member of the sugar transporter family, Stl1p is the glycerol/H+ symporter in Saccharomyces cerevisiae. *Mol. Biol. Cell* **16**, (2005).
19. Mosley, A. L., Lakshmanan, J., Aryal, B. K. & Ozcan, S. Glucose-mediated phosphorylation converts the transcription factor Rgt1 from a repressor to an activator. *J. Biol. Chem.* **278**, (2003).
20. Hirayama, T., Maeda, T., Saito, H. & Shinozaki, K. Cloning and characterization of seven cDNAs for hyperosmolarity-responsive (HOR) genes of Saccharomyces cerevisiae. *Mol. Gen. Genet.* **249**, (1995).
21. Miotto, G., Tessaro, S., Rotta, G. A. & Bonatto, D. In silico analyses of Fsf1 sequences, a new group of fungal proteins orthologous to the metazoan sideroblastic anemia-related sideroflexin family. *Fungal Genet. Biol.* **44**, (2007).

Supplemental resources

---

1) Laboratory notebook: https://www.notion.so/project6-lab-journal-0aed48dd40694ee9841b4cc367f76115
2) Gene ontology for 50 most expressed genes, which are character for fertilisation stage: https://docs.google.com/spreadsheets/d/1z3yR81_wCNJMMx-q8AF8m6isJ2Y8xl5565iP7fgjnjI/edit?usp=sharing