**RNA-seq and bread**

Alekseeva L., Kureev N., Ovsyannikova A.

**Abstract**

There are two replicates of RNA-seq data from yeast before and during fermentation, and our goal is to find out if the yeast express different genes during fermentation than they do under normal growth. In this project we analyzed in-detail the changes in RNA expression during the fermentation process and explored how RNA expression levels change as yeast undergo fermentation to make bread rise.

**Introduction**

*Saccharomyces cerevisiae* is a microorganism used both in the classical food and beverage industry and technology as well as in modern biotechnology and scientific research. A feature of this organism is the ability to ferment in solid-phase conditions, particularly in bread. Yeasts uses organic substances to produce energy and carbon. They need oxygen for respiration and oxidation processes. But yeasts in the absence of oxygen can go to the fermentation with produce of alcohols [1]. However, if oxygen becomes available, the yeasts goes back to the breathing process, producing carbon dioxide. Differential expression plays an important role in changing yeast metabolism. When switching to a different type of using substance, yeasts can activate genes that are involved in the fermentation, and were not used before. In this study we are trying to understand which genes are involved in this process?

RNA-seq methods allow to study cell transcriptome and make special libraries. By sequencing RNA from cells of various tissues or from cells at different times of their existence, it is possible to identify differentially expressed (DE) genes [2]. This feature allows cells with the same genome to be part of different tissues of one organism or (like yeasts) to adapt to changing external conditions, such as composition of the environment. DE genes can be detected by analyzing the number of transcripts related to this gene.

**Materials and Methods**

We analyzedyeast reads with several parameters in Table 1.

Table 1 Analyzed Data

|  |  |  |
| --- | --- | --- |
| accession | fermentation | replicate |
| SRR941816 | 0 minutes | 1 |
| SRR941817 | 0 minutes | 2 |
| SRR941818 | 30 minutes | 1 |
| SRR941819 | 30 minutes | 2 |

As a reference genome we used *Saccharomyces cerevisiae* strainS288c and assembly R64.

Aligning was performed using HISAT2 (<https://daehwankimlab.github.io/hisat2/>).

Convert GFF files into GTF was performed using gffread.

Quantifying was performed using featureCounts (<http://subread.sourceforge.net>).

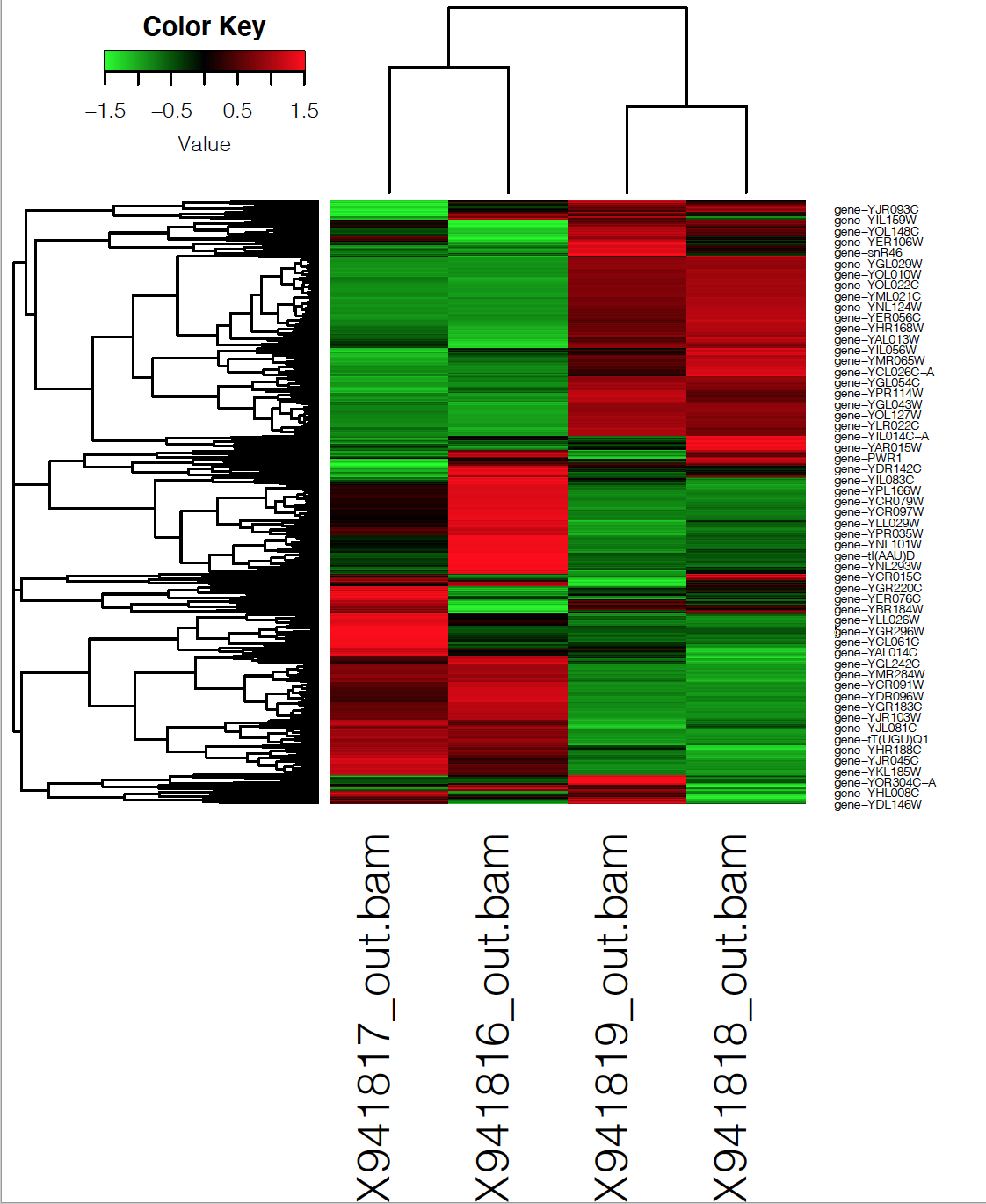
Finding differentially expressed genes was performed using Deseq2 (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>). Parameters and examples of code execution are in our Notebook [2].

**Results**

We analyzed4yeast reads - transcriptional data containing information on gene expression before and during fermentation. Statistics on aligning with HISAT2 is presented in our Notebook [3].

We received a total of 6421 genes after featureCounts. Then for each gene metrics were calculated using Deseq2. Based on normalized counts a heat map was built (Figure 1).

Figure 1. Heatmap based on normalized counts of 6421 genes.



As can be seen in the Figure 1, there is a rough division into 3 groups:

1) Genes present during fermentation and absent before it

2) Genes present before fermentation and absent during it

3) Genes present in both cases.

For our further analysis, we selected the top 50 genes based on adjusted p-values. for each gene, it’s function was identified using gene ontology (GO). As a the result 38 GO were obtained, including the following terms in Table 2.

Table 2. GO for 50 genes

|  |  |  |  |
| --- | --- | --- | --- |
| GOID | TERM | ANNOTATED\_GENES | |
| GO:0006364 | rRNA processing | YDR449C, YEL026W, YER127W, YGR159C, YHR066W, YHR196W, YJL069C, YLR264W, YMR093W, YNL112W, YNL182C, YOL041C, YOL080C | |
| GO:0042273 | ribosomal large subunit biogenesis | YCR072C, YDL063C, YEL026W, YHR066W, YIR012W, YJL122W, YNL182C, YOL041C, YOL080C | |
| GO:0042274 | ribosomal small subunit biogenesis | YDR449C, YEL026W, YER127W, YGR159C, YHR196W, YJL069C, YLR264W, YMR093W | |
| GO:0042255 | ribosome assembly | YCR072C, YGR159C, YHR066W, YIR012W, YLR264W, YNL182C, YOL080C | |
| GO:0006811 | ion transport | YDR536W, YHR094C, YKL120W, YNL065W, YNR060W, YOR271C | |
| GO:0006360 | transcription by RNA polymerase I | YHR196W, YJL148W, YJR063W, YML043C, YMR093W, YNL248C | |
| GO:0055085 | transmembrane transport | YDR536W, YHR094C, YKL120W, YNL065W, YOR271C | |
| GO:0055086 | nucleobase-containing small molecule metabolic process | YBL039C, YMR300C, YNL141W, YOL136C, YOR360C | |
| GO:0005975 | carbohydrate metabolic process | YBR105C, YER062C, YKR097W, YOL136C | |
| GO:0006401 | RNA catabolic process | YER049W, YLR264W, YNL112W, YOR359W | |
| GO:0042221 | response to chemical | YKL120W, YLR224W, YNL065W | |
| GO:0006520 | cellular amino acid metabolic process | YBL039C, YLR180W, YMR300C | |
| GO:0006865 | amino acid transport | YNL065W, YOR271C | |
| GO:0006397 | mRNA processing | YEL026W, YGR159C | |
| GO:0008643 | carbohydrate transport | YDR536W, YHR094C | |
| GO:0015931 | nucleobase-containing compound transport | YKL120W, YLR264W | |
| GO:0006417 | regulation of translation | YER049W, YOR359W | |
| GO:0051052 | regulation of DNA metabolic process | YNL182C, YOR359W | |
| GO:0051603 | proteolysis involved in cellular protein catabolic process | YBR105C, YLR224W | |
| GO:0006353 | DNA-templated transcription, termination | YJR063W, YNL112W | |
| GO:0006629 | lipid metabolic process | YBL039C, YOL151W | |
| GO:0006366 | transcription by RNA polymerase II | YJR063W, YNL112W | |
| GO:0006354 | DNA-templated transcription, elongation | YJL148W, YNL248C | |
| GO:0009451 | RNA modification | YOL124C, YPL212C | |
| GO:0006352 | DNA-templated transcription, initiation | YML043C, YNL248C | |
| GO:0008033 | tRNA processing | YOL124C, YPL212C | |
| GO:0002181 | cytoplasmic translation | YLR264W | |
| GO:0051186 | cofactor metabolic process | YLR180W | |
| GO:0018193 | peptidyl-amino acid modification | YER049W | |
| GO:0032787 | monocarboxylic acid metabolic process | YOL136C | |
| GO:0070647 | protein modification by small protein conjugation or removal | YLR224W | |
| GO:0006091 | generation of precursor metabolites and energy | YOL136C | |
| GO:0006260 | DNA replication | YNL182C | |
| GO:0006418 | tRNA aminoacylation for protein translation | YDR037W | |
| GO:0006970 | response to osmotic stress | YER062C | |
| GO:0008380 | RNA splicing | YEL026W | |
| GO:0006873 | cellular ion homeostasis | YNR060W | |
| GO:0006605 | protein targeting | YBR105C | |

Only two of the 50 genes were expressed in the cell before fermentation. These are the genes responsible for the carbohydrate metabolic process and associated with ribosomes.

The remaining genes were mainly expressed during fermentation.

**Discussion**

In order to obtain deeper insight into the transcriptional response of yeast cells during fermentation, we used the Gene Ontology classification to check for enrichment of specific functionally related gene groups.

The analysis showed several groups of genes presented during the fermentation.

First group contained genes involved in ribosome biogenesis, tRNA processing, and nuclear transport. Second group contained genes involved in responding to osmotic stress. Third group contained genes involved in in amino acid metabolism, translation. And the last group contained genes protein catabolic processes, metabolic processes, amino acid modification, autophagy.

The first group of genes according to published data is expressed in the early stage of fermentation [5]. At the same stage, the expression of genes of the fourth group is suppressed. These processes reflect the transition of nutrient-depleted cells of the stationary phase to active fermentation.

Bread is a solid-phase substance and can cause osmotic stress in yeast cell. In S. cerevisiae, the response to hyperosmotic stress is regulated through a mitogen-activated protein (MAP) kinase pathway, the HOG pathway, which results in accumulation of glycerol - second group [4]. This also shows genes regulated by the high-osmolarity glycerol (HOG) pathway, the major pathway involved in the response to osmotic stress and glycerol homeostasis, are among the most differentially expressed genes at the onset of fermentation. [5]

The middle fermentation phase is associated with increased expression levels of third group genes involved in amino acid and protein biosynthesis. Finally, at the last stage of the fermentation process, cells upregulate genes involved in nutrient metabolism, glycogen biosynthesis, autophagy, and protein catabolic processes.

**References**

1. Perricone M. et al. Yeasts //The Microbiological Quality of Food. – Woodhead Publishing, 2017. – С. 121-131.
2. Oshlack, A., Robinson, M.D. & Young, M.D. From RNA-seq reads to differential expression results. Genome Biol 11, 220 (2010). https://doi.org/10.1186/gb-2010-11-12-220
3. <https://github.com/Lolita05/IB_projects/blob/master/project%207.ipynb>
4. Gustin MC, Albertyn J, Alexander M, Davenport K. 1998. MAP kinase pathways in the yeast Saccharomyces cerevisiae. Microbiol. Mol. Biol. Rev. 62:1264–1300
5. Aslankoohi E. et al. Dynamics of the Saccharomyces cerevisiae transcriptome during bread dough fermentation //Appl. Environ. Microbiol. – 2013. – Т. 79. – №. 23. – С. 7325-7333.