GO! GO! *Saccharomyces cerevisiae*!

Michael Greenberbg, Polina Vaganova
Bioinformatics Institute, Saint Petersburg, Russia

Abstract.
Saccharomyces cerevisiae, commonly known as baker's yeast, possesses a remarkable ability to efficiently convert sugars into ethanol and carbon dioxide, which played an important role throughout human history. In our study we applied differential transcriptomic analysis methods to investigate changes in RNA expression during respiration and anaerobic fermentation processes in S. cerevisiae. Our results indicate an upregulation in the expression of genes involved in transcription and translation upon transition to fermentation, likely reflecting the restructuring of metabolic networks.

Keywords: Saccharomyces cerevisiae, RNA-seq, Gene Ontology, fermentation.

1 Introduction

Yeast fermentation of various plant carbohydrate sources stands as one of humanity's oldest technologies. Even in modern times, yeasts play indispensable roles in numerous biotechnological processes, including beer and wine fermentations, as well as the baking industry. *Saccharomyces cerevisiae*, commonly referred to as baker's yeast, exhibits a remarkable ability to efficiently convert sugars into ethanol and carbon dioxide, even under anaerobic or, in some cases, aerobic conditions.[1] Still, the complexity behind this process is still far from being completely understood. Transcriptomic analysis combined with Gene Ontology approaches provides a comprehensive method for studying these processes, offering a wider perspective on our understanding.

RNA sequencing (RNA-seq) has become an important method for transcriptome analysis in the past 10 years, with the advantages of larger dynamic range, and high sensitivity.[2] RNA-seq can help to compare transcriptomic profiles of yeast in different environmental conditions, allowing many aspects to be understood and explored, such as the cellular status during specific environments, regulatory circuits that determine different cell behaviors.[3]

In our study, we conducted an analysis of differential RNA expression in Saccharomyces cerevisiae strain S288c under two conditions: during oxygen respiration and 30 minutes after start of the fermentation.

2 Materials and methods

In this study, we utilized two sets of yeast RNA-seq data, each consisting of 2 replicates. The first set was obtained under normal growth conditions (respiration), while the second set was acquired 30 minutes after the onset of fermentation. For our reference, we used the S288c strain genome (R64 assembly). Read quality was examined with FastQC tool (version 0.12.1).[4] We conducted read alignment using the Hisat2 tool (version 2.2.1).[5] Samtools (version 1.19)[6] was used for resulting alignments sorting. For subsequent analysis, the annotation file

was transformed into GTF format using the Gffread tool (version 0.12.7).[7] We then utilized the featureCounts from the Rsubread package (version 2.16.1)[8] to generate count files.

Differential expression analysis utilized DESeq2[9] (version 1.42.0).

To investigate pathway and functional enrichment, we employed online tools shinyGO[10] (version 0.80) and goSlimMapper.

3 Results

All reads were assessed using FastQC, indicating consistently high quality across the dataset. Alignment with Hisat2 demonstrated overall mapping rate ranging from 94% to 96%. Following alignment, featureCounts successfully assigned approximately 87% of all alignments to specific features based on annotation (Table 1).

Table 1: Alignment results

Sample	Total sequences	Alignment rate	Assigned alignments
respiration_rep1	9043877	94.33%	86.1%
respiration_rep2	9929568	94.91%	87.3%
fermentation_rep1	1721675	96.22%	88.4%
fermentation_rep2	6172452	96.28%	87.2%

The generated count table was inputted into Deseq2 for subsequent analysis. Genes were identified as differentially expressed based on their adjusted p-value (≤ 0.05) and absolute log fold change ($\leq$ -2 for downregulated genes and ≥ 2 for upregulated genes). In total of **1,979** differentially expressed genes were identified, with **1,042** showing upregulation and **938** displaying downregulation (fig.1).

Functional annotation was conducted separately on the sets of upregulated and downregulated genes. Upregulated genes were predominantly associated with the regulation and biogenesis of ribosomes (fig. 2A,B), as well as amino acid synthesis pathways. Downregulated genes were primarily linked to various metabolic pathways and the biosynthesis of secondary metabolites (fig. 2C,D).

4 Discussion

Switching from aerobic to anaerobic conditions, makes yeasts to change life strategy and leads to different gene expression pattern. Through analyzing the results of *Saccharomyces cerevisiae* RNA sequencing, we identified changes in gene expression after 30 minutes of anaerobic fermentation.

Among the downregulated genes, a significant portion is associated with various metabolic pathways, including the citrate cycle (TCA), pyruvate metabolism, fatty acid metabolism, and the synthesis of secondary metabolites, among others (fig. 2B). This trend is consistent with one of the most enriched cellular processes at this stage being energy derivation by oxidation of organic compounds (fig. 2A). The downregulation of these pathways aligns with

the transition from respiration to fermentation. Interestingly, some other cellular processes, such as protein folding, are also downregulated, suggesting a broader reprogramming of cellular activities during this transition.

Surprisingly, the upregulated genes are predominantly involved to RNA processing, transcription, ribosome assembly, and nucleotide, and synthesis of non-coding RNA (fig. 2C,D). This suggests that, at the 30-minute stage after the onset of fermentation, yeast cells undergo a reorganization of their biosynthetic machinery to efficiently synthesize proteins needed in an anaerobic environment. This likely includes the production of enzymes required for the fermentation process.

Altogether our findings shed some more light on the peculiarities of yeast metabolism.

References

1 S. Dashko, N. Zhou, C. Compagno, *et al.*, "Why, when, and how did yeast evolve alcoholic fermentation?," *FEMS Yeast Research* **14**, 826–832 (2014).

2 A. Sîrbu, G. Kerr, M. Crane, *et al.*, "Rna-seq vs dual- and single-channel microarray data: Sensitivity analysis for differential expression and clustering," *PLoS ONE* **7**, e50986 (2012).

3 C. M. V. L. Wong, S. Y. Boo, C. L. Y. Voo, *et al.*, "A comparative transcriptomic analysis provides insights into the cold-adaptation mechanisms of a psychrophilic yeast, glaciozyma antarctica pi12," *Polar Biology* **42**, 541–553 (2019).

4 S. Andrews, "Fastqc. a quality control tool for high throughput sequence data," (2010).

5 D. Kim, J. M. Paggi, C. Park, *et al.*, "Graph-based genome alignment and genotyping with hisat2 and hisat-genotype," *Nature Biotechnology* **37**, 907–915 (2019).

6 P. Danecek, J. K. Bonfield, J. Liddle, *et al.*, "Twelve years of SAMtools and BCFtools," *GigaScience* **10** (2021). giab008.

7 G. Pertea and M. Pertea, "Gff utilities: Gffread and gffcompare," *F1000Research* **9**, 304 (2020).

8 W. Shi, "Rsubread," (2017).

9 M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for rna-seq data with deseq2," *Genome Biology* **15** (2014).

10 S. X. Ge, D. Jung, and R. Yao, "Shinygo: a graphical gene-set enrichment tool for animals and plants," *Bioinformatics* **36**, 2628–2629 (2019).

5 Supplementary materials

Data availability

1. fermentation 0 minutes: repl1 repl2

2. fermentation 30 minutes: repl1 repl2

FastQC: google disc
Scripts: github

6 Figures

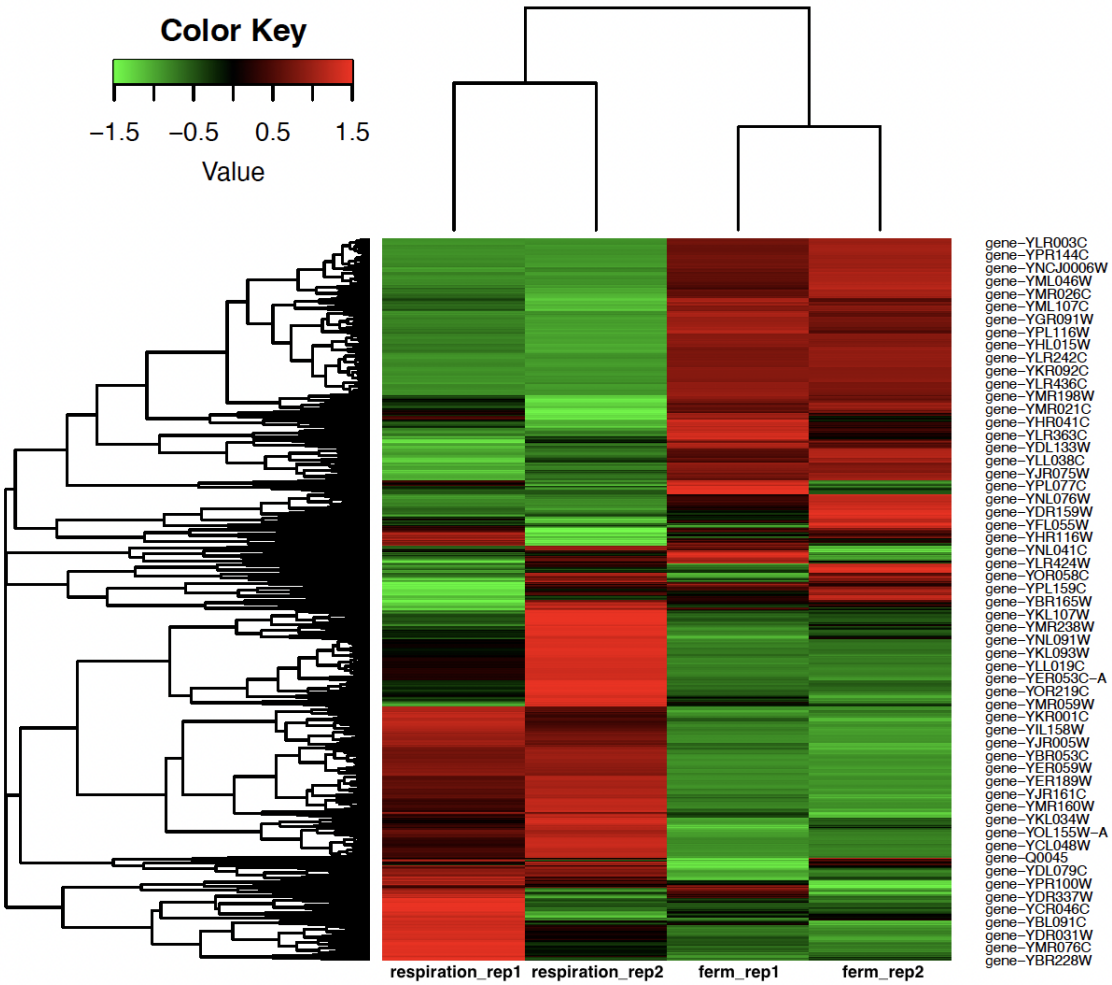

Fig 1: Heatmap of differently expressed genes based on normalised count table. Green stands for downregulation and red for upregulation based on absolute log fold change.

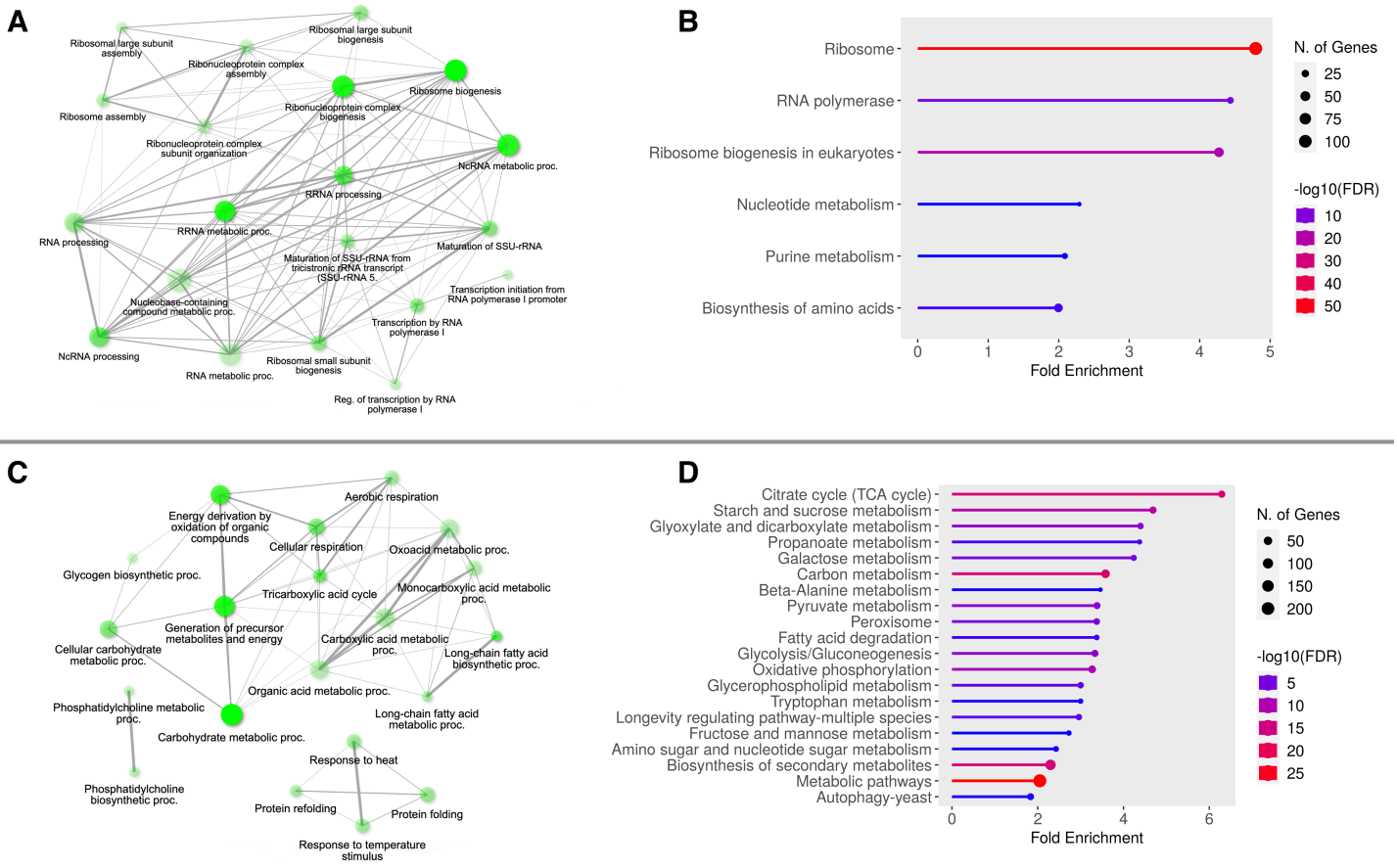

Fig 2: Functional annotation of upregulated (A,B) and downregulated (C,D) genes. A, C - GO cellular process visualisation; B, D - KEGG enrichment. Visualised with shinyGO