American Journal of Biochemistry and Biotechnology

Original Research Paper

Transcriptomic Analysis of the Erythritol High-Yielding Mutant Strain *Yarrowia lipolytica* Y44

[1,2]Xiaojie Ren, [1]Heng Ban, [1]Chao Huang, [1]Baoyue Liu, [1]Yuanda Song, [1,3,4]Xinhe Zhao and [5]Jiangsen Pei

[1]Colin Ratledge Center for Microbial Lipids, School of Agriculture Engineering and Food Science, Shandong University of Technology, Zibo, China
[2]Baolingbao Biology Co., Ltd., Dezhou, China
[3]State Key Laboratory of Microbiology, Shandong University, Qingdao, Shandong, China
[4]Chongqing Academy of Science and Technology, Chongqing, China
[5]China National Research Institute of Food and Fermentation Industry Co., Ltd. Beijing, China

Article history
Received: 15-09-2021
Revised: 06-11-2021
Accepted: 12-11-2021

Corresponding Author:
Xinhe Zhao
Colin Ratledge Center for Microbial Lipids, School of Agriculture Engineering and Food Science, Shandong University of Technology, Zibo, China
E-mail: zhaoxinhe2018@163.com

Jiangsen Pei
China National Research Institute of Food and Fermentation Industry Co., Ltd. Beijing, China
E-mail: pei21cn@163.com

Abstract: RNA-seq technique was used to analyze the transcriptomics of *Yarrowia lipolytica* control strain Po1g, an erythritol-producing natural isolate *Yarrowia lipolytica* Y 22 and an erythritol high-yielding mutant strain Y44 derived from Y 22. Functional annotation of genes and classification of metabolic pathways were carried out to identify Differentially Expressed Genes (DEGs). These DEGs were classified into related metabolic pathways to explain the molecular mechanism of high yield of erythritol production. The results showed that the upregulated genes in erythritol-producing natural isolate Y 22 and erythritol-producing mutant strain Y 44 were mainly involved in pentose phosphate pathway, Tricarboxylic Acid Cycle (TCA) and malic acid cycle. The downregulated genes were mainly involved in amino acid synthesis and oxycarboxylic acid metabolism. The synergistic regulation of the above metabolic pathways can increase the input and reduce the output of erythrose-4-P (E-4-P), which is the precursor of erythritol, promoting erythritol production. In addition, compared with the natural isolate Y 22, the genes related to cell wall synthesis were down-regulated and the expression of transmembrane transporter protein was up-regulated in high-yield mutant strain Y 44. In this way, the permeability of yeast cells is enhanced and the synthesized erythritol can be quickly transported to the outside of the cell, which reduces the decomposition of erythritol and further promotes its yield.

Keywords: *Yarrowia lipolytica*, Transcriptomic Analysis, GO Clustering, KEGG Clustering

Introduction

Erythritol, a four-carbon sugar alcohol, is widely found in algae, fungi and fruits as an energy storage substance. It is also widely found in fermented foods and beverages (Moon *et al*., 2010; Rice *et al*., 2020). The relative sweetness of erythritol is 60-70% that of sucrose. As "zero-calorie sweetener", erythritol cannot be used by bacterial fermentation and will not lead to dental caries and islet reaction. Therefore, it is widely used in functional food and health product industry (Dominique and Jean-Louis, 2017; Regnat *et al*., 2018). Due to the fact that erythritol is the only sugar alcohol produced by fermentation on a large scale and its demand in food, medicine, cosmetics and feed industry is increasing, the production of erythritol by microbial fermentation has attracted more and more attention.

Many yeasts, such as *Zygosaccharomyces*, *Debaryomyces*, *Hansenula*, *Pichia*, can grow under high osmotic pressure (high sugar content or salinity) (Brown, 1978). These hyperosmolar yeasts can stimulate an osmotic response mechanism in a hyperosmolar environment, that is, accumulate one or more compatible solutes for resistance to hyperosmolar environment (Rzechonek *et al*., 2020; Libo, 2015). Glycerol is the most common osmoprotective agent in yeast, but sugar alcohols such as erythritol, D-arabitol and mannitol also have similar functions (Rzechonek *et al*., 2020; Libo, 2015). In eukaryotes, the main function of pentose phosphate pathway is to produce reductive NADPH for cellular reactions and to produce a variety of precursors for the synthesis of nucleotides and amino acids (Horecker and Mehler, 1955). While, PPP is also the main pathway for yeast and other fungi to synthesize erythritol (Xinhe *et al*., 2020).

For a long time, strategies for improving the yield and production rate of target products in the fermentation industry mainly include the breeding of yeast strains with high fermentation performance and the exploration of new fermentation processes (Koh *et al*., 2003; Musial *et al*., 2011). In the early work of our lab, a yeast strain Y-22 which can ferment glucose to erythritol was isolated from nature. After identification, it belongs to *Yarrowia lipolytica*. The ability of the *Yarrowia lipolytica* Y-22 to produce erythritol under high glucose conditions was much higher than that of many isogeneic, heterologous and unidentified impermeable yeasts obtained from domestic storage institutions and isolated from nature. Fermentation tests on 5 and 50 m^3 fermenters showed that the conversion rate of glucose could reach 47% and erythritol yield could reach more than 150 g/L within 96 h (Pei *et al*., 2015). In the preliminary work, the strain *Y. lipolytica* Y 44 with higher yield of erythritol was obtained by physical and chemical mutagenesis on *Y. lipolytica* Y-22. The yield of erythritol could reach 200 g/L on 210 m^3 industrial fermenter and the conversion rate of glucose could reach 64%. However, the mutation mechanism of high production is still unclear.

Transcriptomics has been widely used in scientific research and has been used in the study of *Yarrowia lipolytica*. Xueliang (2020), aiming at improving the heat tolerance of *Yarrowia lipolytica*, obtained heat-resistant strains through adaptive evolution. They applied transcriptomic analysis to explore the key pathway genes for heat resistance. Based on the analysis results, a high yield strain of erythritol was obtained by metabolic engineering. Libo (2015) studied the transcriptional level of *Y. lipolytica* in response to hypertonic mechanism and found that hypertonic can inhibit the transcriptional level of EMP pathway genes. It can induce the transcription of gluconeogenesis pathway genes and TCA cycle genes, thus enhancing the substrate utilization efficiency and energy supply under hyperosmosis. Hyperosmosis can induce the transcription of key genes for erythritol synthesis in the HMP pathway and greatly upregulate the transcription level of oxidative stress response protein genes, thus promoting massive synthesis of erythritol as compatible solute. In this study, the transcriptome of the standard strain *Yarrowia lipolytica* Po1g that does not produce erythritol, the wild erythritol producing strain *Y. lipolytica* Y 22 and the high erythritol producing mutant strain *Y. lipolytica* Y 44 were compared to explore the differentially expressed genes in erythritol metabolism and reveal the metabolic mechanism of high production.

Materials and Methods

Species and Experimental Design

Standard strain *Yarrowia lipolytica* Po1g (W 29 series, does not produce erythritol) was obtained from YEASTERN Biotech Co. Ltd (Taipei, Taiwan), wild erythritol production strain *Yarrowia lipolytica* Y22 was isolated from nature in our lab (Pei *et al*., 2015) and high-yield erythritol mutant strain *Yarrowia lipolytica* Y44 was obtained by physical (UV) and chemical (nitrosoguanidine) mutagenesis of *Y. lipolytica* Y-22. In the nitrosoguanidine mutagenesis method, acetone is used as a solvent to dissolve nitrosoguanidine to a concentration of 200 ppm. Take 2 mL of the mutagenesis solution and add it into 2 g of cells. Keep the mutation time to 20s, to ensure the lethality of the bacteria around 90%. UV mutagenesis is to irradiate the cells with UV rays for 20 min to keep the lethality rate around 99%. Spread the mutagenized cells on a YPD plate to obtain a mutagenized strain.

The *Y. lipolytica* strains were cultured in yeast Extract Peptone Dextrose (YPD) medium with 10 g/L yeast extract, 20 g/L tryptone and 200 g/L glucose. The experiment design was shown in Fig. 1. The concentrations of glucose and erythritol were analyzed according the methods described in (Pei *et al*., 2015). In brief, 1 mL of culture medium was filtered (0.2 μm, Millipore) and analyzed by HPLC equipped with a C18 column, the mobile phase is water and the elution rate is 1 mL/min. The concentrations were calculated by internal standard methods. The conversion rate of glucose = (glucose concentration in the beginning of fermentation - glucose concentration in the end of fermentation)/glucose concentration in the beginning of fermentation ×100%.

RNA Extraction and Sequencing

Cells of *Y. lipolytica* were collected by centrifuging 4 mL of 60 h cultured *Y. lipolytica*strains (Po1g, Y22, Y44) under 10000 r/min, for 3 mins respectively. Total RNA was extracted using UNIQ-10 column Trizol kit and enriched, segmented and reversely transcribed to produce cDNA. cDNAs were ligated to Illumina sequencing adapters and sequenced using Illumina Novaseq 6000 by Gene Denovo Biotechnology Co. (Guangzhou, China).

Differentially Expressed Genes (DEGs) Analysis

The RNAs differential expression analysis was compared by the DESeq2 (Love *et al*., 2014) software between two different samples. The gene/transcript False Discovery Rate (FDR) parameter below 0.05 and the absolute fold change (value after the logarithm at bottom 10) ≥2 is the gene/transcript considered Differentially Expressed (DEGs).

All DEGs were mapped to the Gene Ontology database, used for GO enrichment analysis and pathway significant enrichment (Van der Auwera *et al*., 2013) was used for further KEGG enrichment analysis of differentially expressed genes. We considered genes with a P value ≤0.05 to be differentially expressed. Finally, we selected genes with at least two-fold change in the expression level.

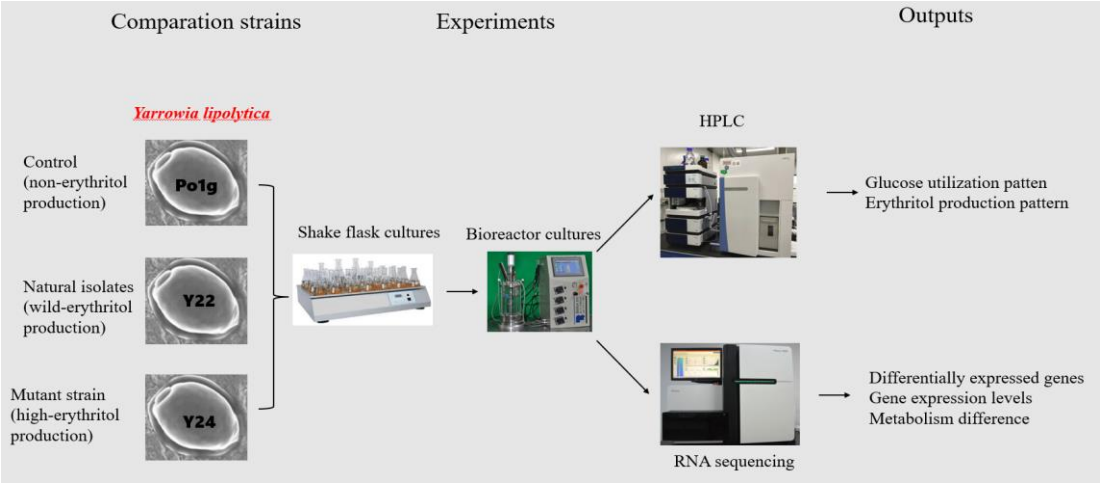


Fig. 1: Experimental design. The fermentation and transcriptomic characteristics of the three *Y. lipolytica* strains were compared based on this experimental process

Results and Discussion

Studying the fermentation process of glucose to erythritol between the three strains in YPD medium shows a large difference (see HPLC map of Fig. 2), where Po1g basically did not accumulate erythritol, while the conversion rate of glucose was 46.3% and erythritol yield was 139 g/L in *Y. lipolytica* Y22. And the conversion rate of glucose was 62.1% and the yield of erythritol was 190 g/L in *Y. lipolytica* Y44. Therefore, the conversion rate and erythritol yield of Y22 and Y44 were greatly improved.

The differential expression analysis was done to explore the mechanisms of erythritol production in the wild producing and high-yield mutant strains. GO is an international classification system (Ashburner *et al*., 2000), which includes Biological Process (BP), Cellular Component (CC) and Molecular Function (MF). These were used to describe the functions of different genes. GO annotation carries out in-depth analysis of the products encoded by genes to explore the biological processes, molecular functions and cellular environment they participate in respectively. After gene GO annotation, genes successfully annotated were classified according to the next level. Based on the significant difference analysis, Fig. 3 shows the gene function categories of the top 20 classes with significant differences between standard strain Po1g and wild producing strain Y22. As can be seen from the figure, the differences between the two in GO classification mainly focus on biological processes and molecular functions, while the differences in cell composition are not obvious. Among these genes classified into biological processes, a large number of genes are involved in organic acid metabolism, oxoacid metabolism, carboxylic acid metabolism, amino acid metabolism. The genes with different molecular functions mainly focus on catalytic activity, oxidoreductase activity and lyase activity.

(X-axis is z-score, positive value indicates up regulated genes, negative value indicates down regulated genes; Y-axis is -log10 (p value), which is the negative logarithm of p value after multiple corrections. The larger the value is, the less obvious the significance is. The circle size shows the number of genes under GO Term. The following figures are the same)

Figure 4 shows the functional categories of genes in the top 20 classes differentially expressed by standard strain Po1g versus mutant strain Y44. The significant differences between the two are also mainly focused on biological processes and molecular functions. Similar to Y22, a large number of genes classified into biological processes which include the metabolism of organic acids, oxoacids, carboxylic acids and amino acids. Genes with different molecular function mainly focus on catalytic activity, oxidoreductase activity and carbon-oxygen lyase activity.

The erythritol metabolism process is mainly generated mainly through the PPP pathway under a series of Transketolase/Transketolase (TKL/TAL) and Erythritose Reductase (ER) (Fig. 5), Po1g is a standard strain of *Yarrowia lipolytica* without erythritol producing ability and its similar differences between Y22 and Y44 are indicating some important gene classes associated with erythritol synthesis, which include the metabolism of organic acids, oxoacids, carboxylic acids and amino acids. In addition, due to the intensive reductive reaction among aldose, ketose and sugar alcohol in the metabolic process, which is accompanied by carbon-oxygen bond cracking (Janek *et al*., 2017; Cheng *et al*., 2018), the activities of oxidoreductase and carbon-oxygen lyase are significantly differentially expressed. Compared with the standard strain that did not produce erythritol, the genes specifically expressed in erythritol synthesis were fully demonstrated.

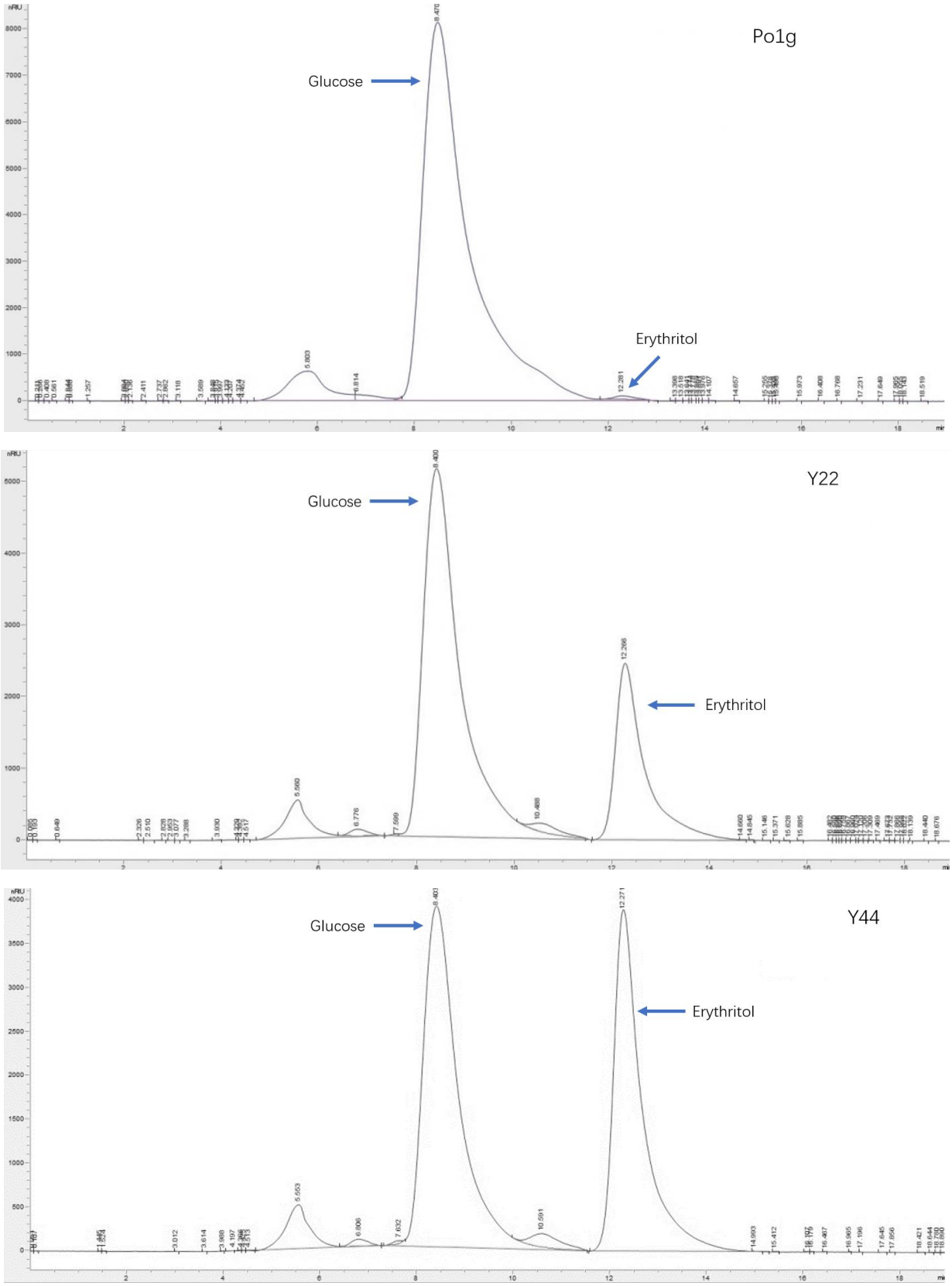
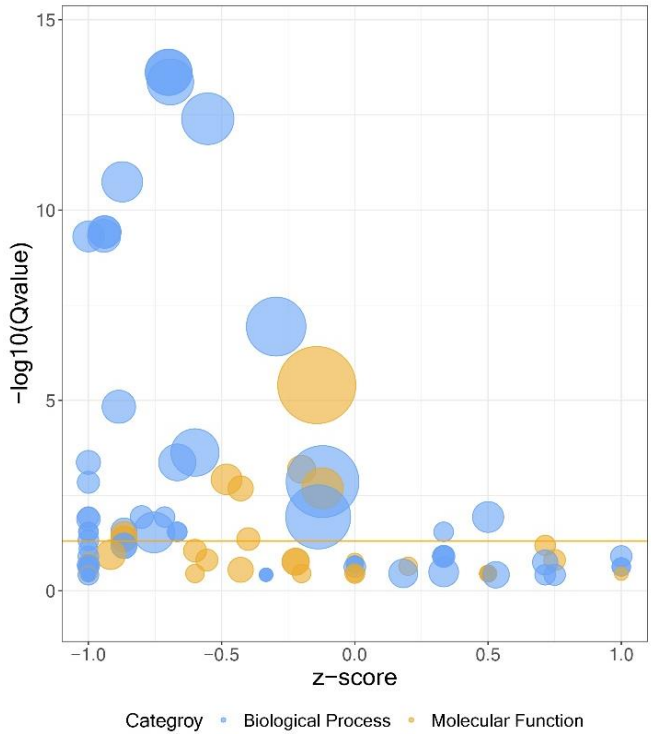


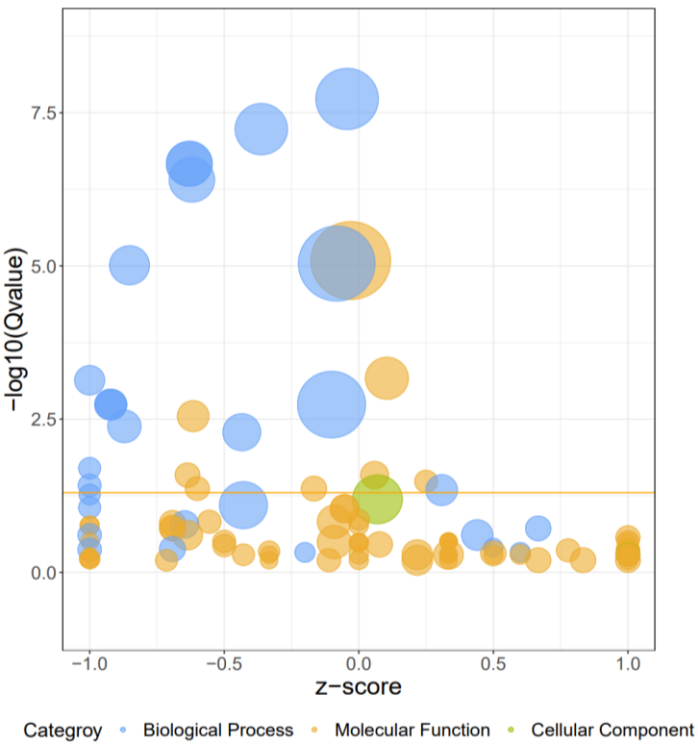
Fig. 2: HPLC analyses of glucose conversion and erythritol production in *Y. lipolytica* strains



GO term Top 20

ID	Descrption
GO:0006082	organic acid metabolic process
GO:0043436	oxoacid metabolic process
GO:0019752	carboxylic acid metabolic process
GO:0044281	small molecule metabolic process
GO:0006520	cellular amino acid metabolic process
GO:0016053	organic acid biosynthetic process
GO:0046394	carboxylic acid biosynthetic process
GO:0044283	small molecule biosynthetic process
GO:0008652	cellular amino acid biosynthetic process
GO:0044710	single−organism metabolic process
GO:1901605	alpha−amino acid metabolic process
GO:1901564	organonitrogen compound metabolic process
GO:0044711	single−organism biosynthetic process
GO:0009066	aspartate family amino acid metabolic process
GO:0044699	single−organism process
GO:0009067	aspartate family amino acid biosynthetic process
GO:0003824	catalytic activity
GO:0016614	oxidoreductase activity, acting on CH−OH group of donors
GO:0016829	lyase activity
GO:0016491	oxidoreductase activity

Fig. 3: GO differential clustering analysis between Po1g-vs-Y22



GO term Top 20

ID	Descrption
GO:0044710	single−organism metabolic process
GO:0044281	small molecule metabolic process
GO:0006082	organic acid metabolic process
GO:0043436	oxoacid metabolic process
GO:0019752	carboxylic acid metabolic process
GO:0044699	single−organism process
GO:0006520	cellular amino acid metabolic process
GO:0008652	cellular amino acid biosynthetic process
GO:0044763	single−organism cellular process
GO:0044283	small molecule biosynthetic process
GO:0016053	organic acid biosynthetic process
GO:0046394	carboxylic acid biosynthetic process
GO:1901605	alpha−amino acid metabolic process
GO:0044711	single−organism biosynthetic process
GO:0006553	lysine metabolic process
GO:0003824	catalytic activity
GO:0016491	oxidoreductase activity
GO:0016829	lyase activity
GO:0016835	carbon−oxygen lyase activity
GO:0016614	oxidoreductase activity, acting on CH−OH group of donors

Fig. 4: GO differential clustering analysis between Po1g-vs-Y44

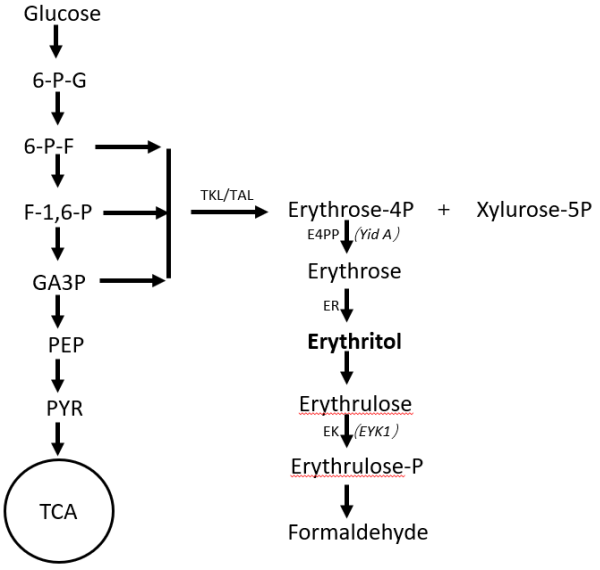


Fig. 5: Synthesis and decomposition pathway of erythritol

Figure 6 shows the top 20 classes of differentially expressed gene categories between wild erythritol producing strain Y22 versus the mutant high-producing strain Y44. As we can see from the figure, difference between Y22 and Y44 was not significant and the -log10 (q value) were all below 1 (Fig. 6). The first two categories of genes with more differences between them belong to anionic transport function in biological process cluster and membrane composition difference in cell composition cluster respectively. All other differences are concentrated in clusters of molecular function, with multiple secondary classifications of transmembrane transporter activity. It can be seen that the high-yielding strain Y44 has great changes with Y 22 in transmembrane transport function in terms of biological process, cell composition and molecular function. By changing the composition of cell membrane, it affects transmembrane molecular function and transmembrane biological process and finally achieves high yield of erythritol.

Vandermies *et al*. (2017) reported that the gene *eyk1*, encodes an erythrulosekinaseerythrulose, the direct downstream product of erythritol to generate erythrulose-P (Fig. 5). Then erythrulose-P can continue to generate formaldehyde and Dihydroxyacetone Phosphate (DHAP) which direct the metabolism into glycolytic pathway or TCA pathway. The catabolism of erythritol had a negative effect on the accumulation of erythritol in *Yarrowia lipolytica*. Therefore, it is inferred that the mutant Y44 with high yield erythritol may accelerate the transport of erythritol outside the cell, therefore reduces its accumulation and decomposition in the cell through the mutation of membrane structure. In addition, the high-yielding mutant strain Y44 also showed higher expression level of transerring aldehyde or ketonic groups, which are key enzymes catalyzing the synthesis of erythritol by PPP pathway. Therefore, Y44 showed a higher ability of erythritol production from increasing the synthesis of erythritol to decreasing the decomposition of erythritol in the cell.

However, GO classification was still not detailed enough, so we further analyzed the KEGG database. KEGG database is a database that systematically analyzes the functions of gene products and which metabolic pathways they are involved in cells (Zhang *et al*., 2016). After KEGG Ortholog (KO) annotation of genes, they can be divided into the following categories according to the KEGG metabolic pathway they participate in: Cellular processes (A), environmental information processing (B), genetic information processing (C), genetic information processing, metabolism (D) and organismal systems (E). Moreover, there is a more detailed secondary classification under the primary classification, thus enabling a more accurate analysis of differential information on gene expression.

It can be seen from Fig. 7-9 that the specific differences between strains producing (Y22, Y44) and strains not producing (Po1g) erythritol mainly include 2-oxycarboxylic acid metabolism, lysine biosynthesis, histidine metabolism, phenylalanine, tyrosine and tryptophan biosynthesis. There was no significant difference between high yield (Y 44) and low yield (Y 22) strains. It shows that the synthesis of these amino acids has a direct effect on the synthesis of erythritol.

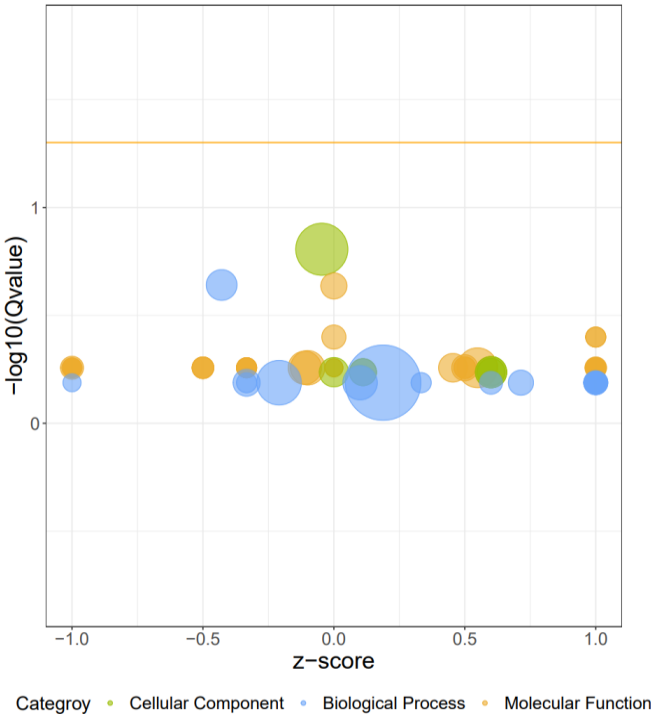
In addition, there were also widespread differences between strains producing erythritol (Y 22, Y 44) and the strain not producing erythritol (Po1g), as well as between wild producing strain (Y 22) and high-producing mutant strain (Y 44). These differences involved in metabolic pathways, degradation of valine, leucine and isoleucine; biosynthesis of secondary metabolites; sulfur metabolism; quinones and other terpenoids, quinone biosynthesis; pyruvate metabolism; glyoxylic acid and dicarboxylic acid metabolism; These differences indicated that these categories involved in promotion of the erythritol yield.

In an organism, many genes coordinate with each other and work together to perform a certain physiological function. Pathway significant enrichment can clarify what biochemical metabolic pathways and signal transduction pathways these differentially expressed genes specifically participate in Bayerlová *et al*. (2015). Metabolic pathway analysis showed that lysine and histidine metabolism, phenylalanine, tyrosine and tryptophan biosynthesis were down-regulated in Y 44 compared with Y 22 (Supplementary Fig. 1). The synthesis of these amino acids starts from the pentose phosphate pathway via PRPP or erythrose-4-P, so the synthesis of these amino acids will consume erythrose-4-P (the precursor of erythritol synthesis) (Hove-Jensen *et al*., 2017). Therefore, down-regulation of these amino acids' synthesis ensures sufficient precursor for erythritol synthesis. In addition, oxocarboxylic is down-regulated across the board in Y44 compared with

Y22 (Supplementary Fig. 2). Oxocarboxylic is a precursor of the valine, leucine and isoleucine synthesis pathways (KEGG Pathway database) and its down-regulation directly affects downstream amino acid metabolism (Supplementary Fig. 1). In the same time, aerobic acid metabolism and organic acid metabolism, such as TCA cycle reactions were generally upregulated (E1.1.1.37, E 4.2.1.2, E1.3.5.1, E6.2.1.4, E6.2.1.5, E1.1.1.41). TCA cycle feeds back pyruvate through the malic acid cycle pathway, thereby upregulating pyruvate anabolic metabolism. In addition, because the carboxylic acid cycle and oxycarboxylic acid share a common precursor of pyruvate, the synergic regulation of oxycarboxylic acid (down-regulation) and oxycarboxylic acid cycle (up-regulation) results in a large number of carbon metabolism flow interception at the pyruvate node, which fully ensures sufficient substrate for upstream carbon metabolism.

In addition, pentose phosphate pathway is generally upregulated in carbon metabolism. In this study, glucose was used as substrate to generate erythritol. Therefore, it can be seen from the metabolic diagram of PPP pathway (Fig. 10) that up-regulated genes from glucose to 4-P-erythritol include transketolase (E2.2.1.1), fructose-bisphosphate aldolase (E4.1.2.13), UDP-glucose pyrophosphorylase (E2.7.7.9); xylose reductase (E1.1.1.307); xylose-reductase (E1.1.1.9); arabinitol 2-dehydrogenase (E1.1.1.250). Transketolase catalyzes sedoheptulose-7-phosphate (S-7-P) and d-Glyceraldehyde-3-Phosphate (GAPD) to eythrose-4-P. Fructose bisphosphate aldolase cleaves 1, 6 - bisphosphate - fructose to glycerol phosphate and d-Glyceraldehyde 3-Phosphate (GAPD). GAPD can also enter the PPP pathway to produce eythrose-4-P. The up-regulation of these genes promotes the synthesis of erythritol precursors and accelerates the accumulation of erythritol. In addition, phosphoribulose epimerase (E5.1.3.1) expression was down-regulated, which reduced the conversion of D-ribulose-5-phosphate (R-5-P) to D-xylulose 5-phosphate (X-5-P) and enabled the PPP pathway to go forward, leading to the synthesis of eythrose-4-P.

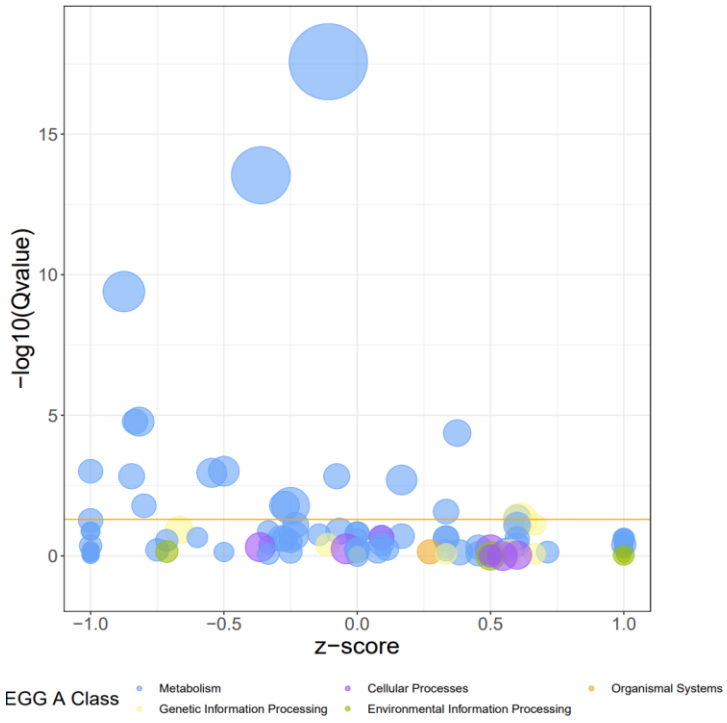
Genes closely related to erythritol synthesis were screened out from the above differential genes and metabolic pathway analysis to make metabolic map. Meanwhile the changing folds of these differentially expressed genes were listed in supplementary Table 1. As shown in Fig. 11, glucose entered pentose phosphate pathway through gluconate-6-P (G-6-P) and synthesized erythrose-4-P (E-4-P) under the action of a series of transketolase/transketolase (TKL/TAL). Erythritol is then synthesized by phospho-phosphatase (*yidA*) and erythritol reductase (ER). The overall up-regulation of pentose phosphate pathway ensured the accumulation of E-4-P, the precursor of erythritol synthesis, in the strains producing erythritol (Y22, Y44) compared with the strain not producing erythritol. At the same time, the amino acid anabolism (TRY, TYR, PHE, HIS) shunted from E-4-P was down-regulated, further ensuring that the direction of carbon metabolism changed from amino acid anabolism to erythritol anabolism.



GO term Top 20

ID	Descrption
GO:0006820	anion transport
GO:0016020	membrane
GO:0008509	anion transmembrane transporter activity
GO:0008514	organic anion transmembrane transporter activity
GO:0043565	sequence−specific DNA binding
GO:0043566	structure−specific DNA binding
GO:0003677	DNA binding
GO:0005337	nucleoside transmembrane transporter activity
GO:0016744	transferase activity, transferring aldehyde or ketonic groups
GO:0005342	organic acid transmembrane transporter activity
GO:0016776	phosphotransferase activity, phosphate group as acceptor
GO:0046943	carboxylic acid transmembrane transporter activity
GO:0004497	monooxygenase activity
GO:0016859	cis−trans isomerase activity
GO:1901505	carbohydrate derivative transporter activity
GO:0001071	nucleic acid binding transcription factor activity
GO:0000988	transcription factor activity, protein binding
GO:0000989	transcription factor activity, transcription factor binding
GO:0015932	nucleobase−containing compound transmembrane transporter activity
GO:0032553	ribonucleotide binding

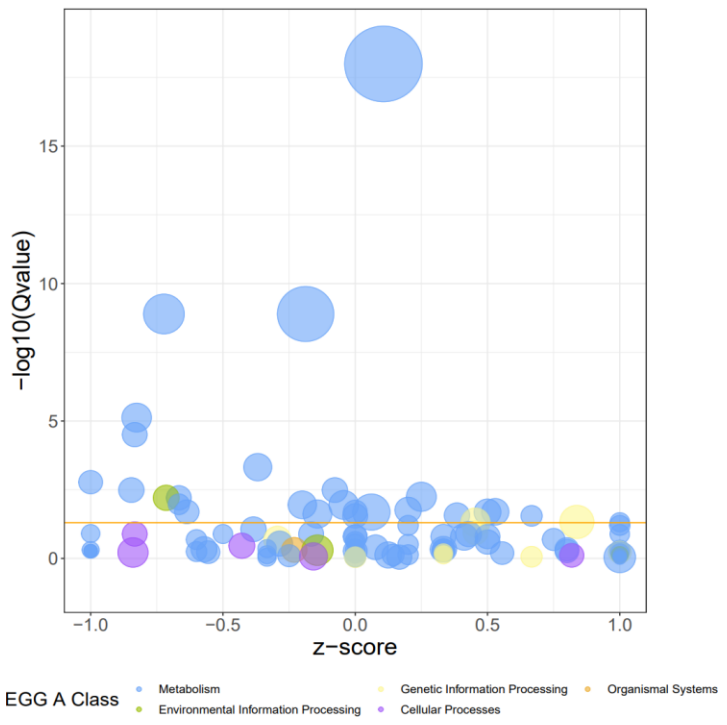
Fig. 6: GO differential clustering analysis between Y22-vs-Y44



Pathway Top 20

ID	Descrption
ko01100	Metabolic pathways
ko01110	Biosynthesis of secondary metabolites
ko01230	Biosynthesis of amino acids
ko01210	2−Oxocarboxylic acid metabolism
ko00290	Valine, leucine and isoleucine biosynthesis
ko00360	Phenylalanine metabolism
ko00300	Lysine biosynthesis
ko00270	Cysteine and methionine metabolism
ko00260	Glycine, serine and threonine metabolism
ko00340	Histidine metabolism
ko00400	Phenylalanine, tyrosine and tryptophan biosynthesis
ko00620	Pyruvate metabolism
ko00630	Glyoxylate and dicarboxylate metabolism
ko00920	Sulfur metabolism
ko01200	Carbon metabolism
ko00350	Tyrosine metabolism
ko00410	beta−Alanine metabolism
ko00770	Pantothenate and CoA biosynthesis
ko04141	Protein processing in endoplasmic reticulum
ko03450	Non−homologous end−joining

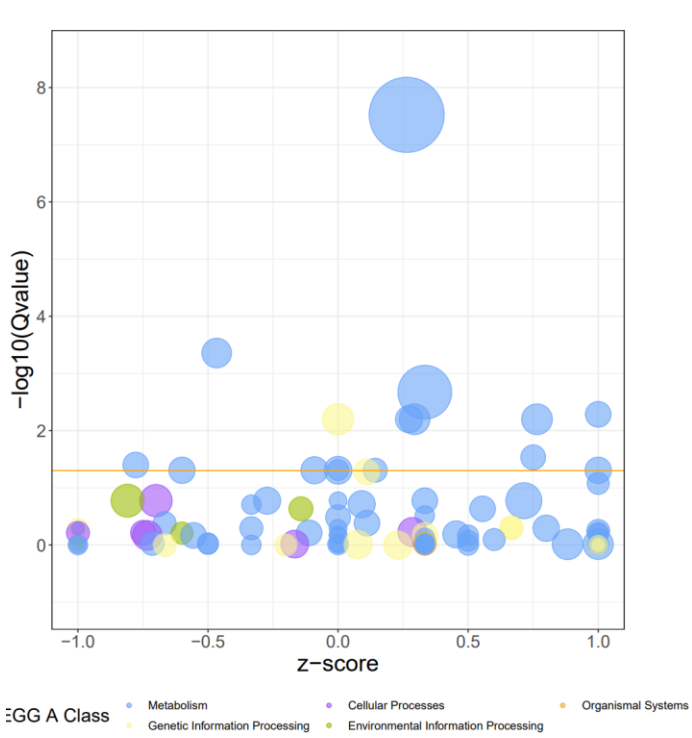
Fig. 7: KEGG differential clustering analysis between Po1g-vs-Y22



Pathway Top 20

ID	Descrption
ko01100	Metabolic pathways
ko01110	Biosynthesis of secondary metabolites
ko01230	Biosynthesis of amino acids
ko01210	2−Oxocarboxylic acid metabolism
ko00290	Valine, leucine and isoleucine biosynthesis
ko00280	Valine, leucine and isoleucine degradation
ko00300	Lysine biosynthesis
ko00340	Histidine metabolism
ko00400	Phenylalanine, tyrosine and tryptophan biosynthesis
ko00620	Pyruvate metabolism
ko00770	Pantothenate and CoA biosynthesis
ko00910	Nitrogen metabolism
ko00640	Propanoate metabolism
ko00630	Glyoxylate and dicarboxylate metabolism
ko00680	Methane metabolism
ko00220	Arginine biosynthesis
ko00510	N−Glycan biosynthesis
ko00650	Butanoate metabolism
ko01200	Carbon metabolism
ko02010	ABC transporters

Fig. 8: KEGG differential clustering analysis between Po1g-vs-Y44



Pathway Top 20

ID	Descrption
ko01100	Metabolic pathways
ko00280	Valine, leucine and isoleucine degradation
ko01110	Biosynthesis of secondary metabolites
ko00920	Sulfur metabolism
ko00130	Ubiquinone and other terpenoid−quinone biosynthesis
ko00620	Pyruvate metabolism
ko00630	Glyoxylate and dicarboxylate metabolism
ko00100	Steroid biosynthesis
ko00360	Phenylalanine metabolism
ko00072	Synthesis and degradation of ketone bodies
ko00410	beta−Alanine metabolism
ko00330	Arginine and proline metabolism
ko00600	Sphingolipid metabolism
ko00680	Methane metabolism
ko00650	Butanoate metabolism
ko00450	Selenocompound metabolism
ko01200	Carbon metabolism
ko04138	Autophagy − yeast
ko03030	DNA replication
ko03440	Homologous recombination

Fig. 9: KEGG differential clustering analysis between Y22-vs-Y44

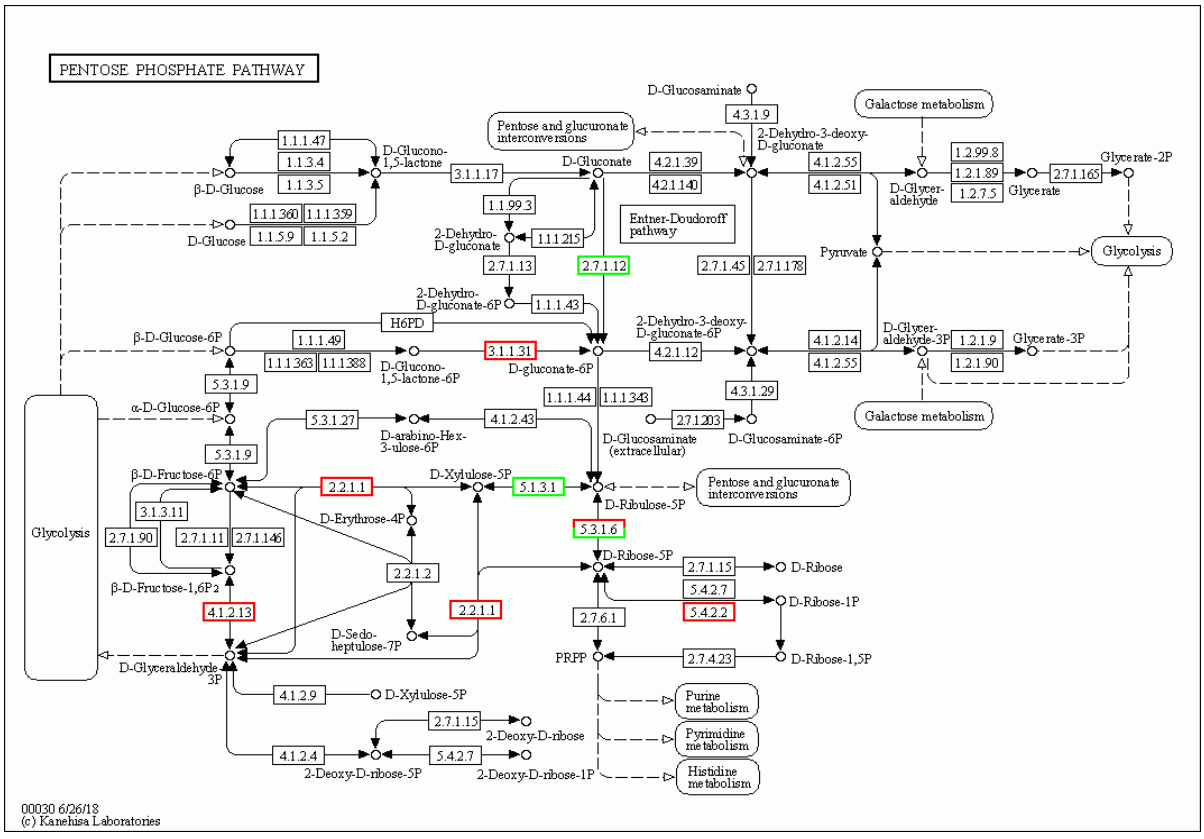


Fig. 10: PPP Pathway difference analysis between Y22-Y44

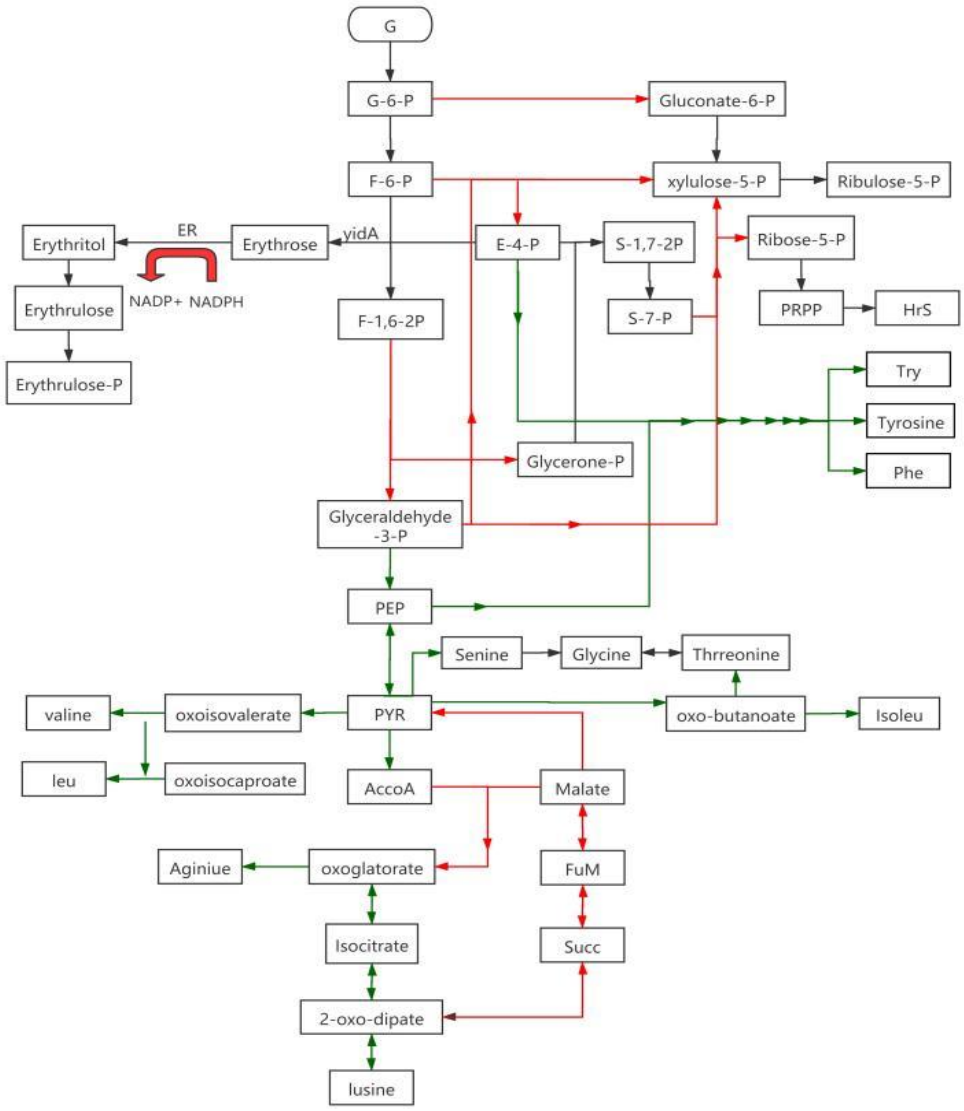


Fig. 11: Metabolic diagram of differential genes related to erythritol synthesis

In addition, oxycarboxylic acid was further down-regulated in high yield mutant strain Y44 compared with wild strain Y22, resulting in a decrease in the synthesis of valine, leucine and isoleucine. At the same time, the tricarboxylic acid cycle is upregulated and the malic acid pathway feeds back pyruvate, which makes the carbon metabolism flow in pyruvate node intercepted, reduces the synthesis level of amino acids in the downstream and fully ensures the substrate of upstream carbon metabolism and further enhances the synthesis of upstream erythritol. In addition, there were significant differences in the expression of chitin synthase, 1,3-β -glucan synthase and NADPH dehydrogenase between Y22 and Y44. Chitin is a characteristic component of fungal cell wall. It is n-acetylglucosamine connected by β-(1,4). Chitin synthase is a key enzyme for biosynthesis of chitin (Li *et al*., 2019). The 1,3-β -glucan synthase, also known as 1,3-β -d-glucan -UDP glucosyltransferase, catalyzes the polymerization and elongation of glucose chains, thus forming cellulose (Aimanianda *et al*., 2017). Cellulose is also a component of fungal cell walls. Therefore, in Y44, the down-regulated expression of these two genes leads to enhanced cell wall permeability. Coupled with the up-regulated transcription level of membrane transporter, the synthesis of erythritol can be rapidly transferred out of the cell, thus reducing its consumption caused by downstream metabolic decomposition. NADPH dehydrogenase is the reducing power provider of PPP pathway and erythritol reductase and the up-regulated expression of this enzyme in Y44 compared with Y22 also promotes the synthesis of erythritol.

Conclusion

Gene transcriptional differences between non-production strains, wild production strains and high-yield mutant strains erythritol were compared by transcriptome analysis, revealing metabolic pathways closely related to erythritol synthesis and high production. The results show that the metabolism of amino acids is an important pathway to influence erythritol synthesis. The synthesis of numerous amino acids compete with erythritol for the precursor substance erythritose-4-phosphate. At the same time, many amino acids with pyruvate and TCA cycle intermediate metabolites as their precursors shunt carbon sources through oxoarboxylic acid metabolism. Transcriptional downregulation of amino acid synthesis pathway genes directly prompted the substantial accumulation of erythritol. In addition, the transcription of erythritol synthetic transketolase/transketolase (TKL/TAL), phospho-phosphatase (*yidA*), Erythritol Reductase (ER) and NADPH dehydrogenase are all substantially upregulated, promoting the erythritol synthesis. In addition to the regulation related to erythritol metabolism, the downregulation of cell wall synthesis-related genes and transporter genes were also found in high-yield strains and the regulation of cell composition also caused enhanced cell permeability and capacity to transport erythritol, thus reducing the degradation of erythritol in the cell. This study elucidates the molecular mechanism of the increased production of erythritol mutant strains and provides theoretical support for further genetic engineering modification.

Acknowledgement

The authors are grateful to the funding support from Major Scientific and Technological Innovation Projects of Shandong Province Key Research and Development Program (2019JZZY011005). Shandong Provincial Natural Science Foundation, China (ZR2019BC099 and ZR2020MC201). National Natural Science Foundation of China (31972851 and 31670064). Taishan Industrial Experts Program (tscy20160101). China Post-doctoral Science Foundation (2019M650167 and 2019M662362). Chongqing technological innovation and application development key projects (cstc2019jscx-gksbX0113).

Author's Contributions

Xiaojie Ren: Participated in the whole experiment process and also contributed to the interpretation of the results.

Heng Ban, ChaoHuang, Baoyue Liu: Contributed to the manuscript preparation.

Jiangsen Pei: Contributed to the guidance of experiment design and ameliorated the manuscript.

Yuanda Song and Xinhe Zhao: Ameliorated the manuscript and provided funding support.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Aimanianda, V., Simenel, C., Garnaud, C., Clavaud, C., Tada, R., Barbin, L., ... & Latge, J. P. (2017). The dual activity responsible for the elongation and branching of β-(1, 3)-glucan in the fungal cell wall. MBio, 8(3), e00619-17. https://journals.asm.org/doi/full/10.1128/mBio.00619-17

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... & Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. Nature genetics, 25(1), 25-29. https://www.nature.com/articles/ng0500_25

Bayerlová, M., Jung, K., Kramer, F., Klemm, F., Bleckmann, A., & Beißbarth, T. (2015). Comparative study on gene set and pathway topology-based enrichment methods. BMC bioinformatics, 16(1), 1-15. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-015-0751-5

Brown, A. D. (1978). Compatible solutes and extreme water stress in eukaryotic micro-organisms. Advances in microbial physiology, 17, 181-242. doi.org/10.1016/S0065-2911(08)60058-2

Cheng, H., Wang, S., Bilal, M., Ge, X., Zhang, C., Fickers, P., & Cheng, H. (2018). Identification, characterization of two NADPH-dependent erythrose reductases in the yeast Yarrowia lipolytica and improvement of erythritol productivity using metabolic engineering. Microbial cell factories, 17(1), 1-12. https://link.springer.com/article/10.1186/s12934-018-0982-z

Dominique, T., & Jean-Louis, B. (2017). EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA); Sugar-free hard confectionery with at least 90% erythritol and reduction of dental plaque which reduces the risk of dental caries: Evaluation of a health claim pursuant to Article 14 of Regulation (EC) No 1924/2006. EFSA Journal. 2017, 15(7), 4923. doi.org/10.2903/j.efsa.2017.4923

Horecker, B. L., & Mehler, A. H. (1955). Carbohydrate metabolism. Annual review of biochemistry, 24(1), 207-274. https://www.annualreviews.org/doi/abs/10.1146/annurev.bi.24.070155.001231?journalCode=biochem

Hove-Jensen, B. andersen, K. R., Kilstrup, M., Martinussen, J., Switzer, R. L., & Willemoës, M. (2017). Phosphoribosyl diphosphate (PRPP): Biosynthesis, enzymology, utilization and metabolic significance. Microbiology and Molecular Biology Reviews, 81(1), e00040-16. https://journals.asm.org/doi/full/10.1128/MMBR.00040-16

Janek, T., Dobrowolski, A., Biegalska, A., & Mirończuk, A. M. (2017). Characterization of erythrose reductase from Yarrowia lipolytica and its influence on erythritol synthesis. Microbial cell factories, 16(1), 1-13. https://link.springer.com/article/10.1186/s12934-017-0733-6

Koh, E. S., Lee, T. H., Lee, D. Y., Kim, H. J., Ryu, Y. W., & Seo, J. H. (2003). Scale-up of erythritol production by an osmophilic mutant of Candida magnoliae. Biotechnology letters, 25(24), 2103-2105. https://link.springer.com/article/10.1023/B:BILE.0000007076.64338.ce

Li, Y., Sun, H., Zhu, X., Bian, C., Wang, Y., & Si, S. (2019). Identification of new antifungal agents targeting chitin synthesis by a chemical-genetic method. Molecules, 24(17), 3155. doi.org/10.3390/molecules24173155

Libo, Y. (2015). Study on the hyperosmotic fermentation of erythritol by*Yarrowialipolytica* and mechanisms of high osmotic pressure response. Doctoral thesis. Jiangnan University. 2015. doi.org/10.4014/jmb.1412.12026

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology, 15(12), 1-21. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8

Moon, H. J., Jeya, M., Kim, I. W., & Lee, J. K. (2010). Biotechnological production of erythritol and its applications. Applied microbiology and biotechnology, 86(4), 1017-1025. https://link.springer.com/article/10.1007/s00253-010-2496-4

Musial, I., Marcinkiewicz, M., Cibis, E. (2011). Optimization of erythritol production from glycerol by *Yarrowialipolytica*Wratislavia K1 [J]. Current Opinion Biotechnology, 2011, 22, 95. doi.org/10.1016/j.copbio.2011.05.296

Pei, J., Huang, L., Zhang, L., & Shubin, T. (2015). Erythritol production with *Yarrowia lipolitica*. Food and Fermentation Industries, 41, 28-33. doi.org/10.13995/j.cnki.11-1802/ts.201504006

Regnat, K., Mach, R. L., & Mach-Aigner, A. R. (2018). Erythritol as sweetener-wherefrom and whereto?. Applied microbiology and biotechnology, 102(2), 587-595. https://link.springer.com/article/10.1007/s00253-017-8654-1

Rice, T., Zannini, E., K. Arendt, E., & Coffey, A. (2020). A review of polyols–biotechnological production, food applications, regulation, labeling and health effects. Critical reviews in food science and nutrition, 60(12), 2034-2051. doi.org/10.1080/10408398.2019.1625859

Rzechonek, D. A., Szczepańczyk, M., Wang, G., Borodina, I., & Mirończuk, A. M. (2020). HOG-Independent Osmoprotection by Erythritol in Yeast Yarrowia lipolytica. Genes, 11(12), 1424. doi.org/10.3390/genes11121424

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., ... & DePristo, M. A. (2013). From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Current protocols in bioinformatics, 43(1), 11-10. doi.org/10.1002/0471250953.bi1110s43

Vandermies, M., Denies, O., Nicaud, J. M., & Fickers, P. (2017). EYK1 encoding erythrulose kinase as a catabolic selectable marker for genome editing in the non-conventional yeast Yarrowia lipolytica. Journal of microbiological methods, 139, 161-164. doi.org/10.1016/j.mimet.2017.05.012

Xinhe, Z., Tianlong, H., Yaxin, Z., Xiaojie, R., Yuanda, S., & Bei, Guo. (2020). Knock Down of Erythrulose Kinase (eyk1) Leads to the Enhancement of Erythritol Production in *Yarrowialipolytica* Mutant Strain YE4-2. American Journal of Biochemistry and Biotechnology 2020, 16 (3), 332.344.

Xueliang, Qiu. (2020). Thermotolerance mechanism analysis and combination strategy transformation of erythritol producing *Yarrowialipolytica*. Doctoral thesis. Jiangnan University. 2020.

Zhang, Y. H., Chu, C., Wang, S., Chen, L., Lu, J., Kong, X., ... & Cai, Y. D. (2016). The use of gene ontology term and KEGG pathway enrichment for analysis of drug half-life. PloS one, 11(10), e0165496. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0165496

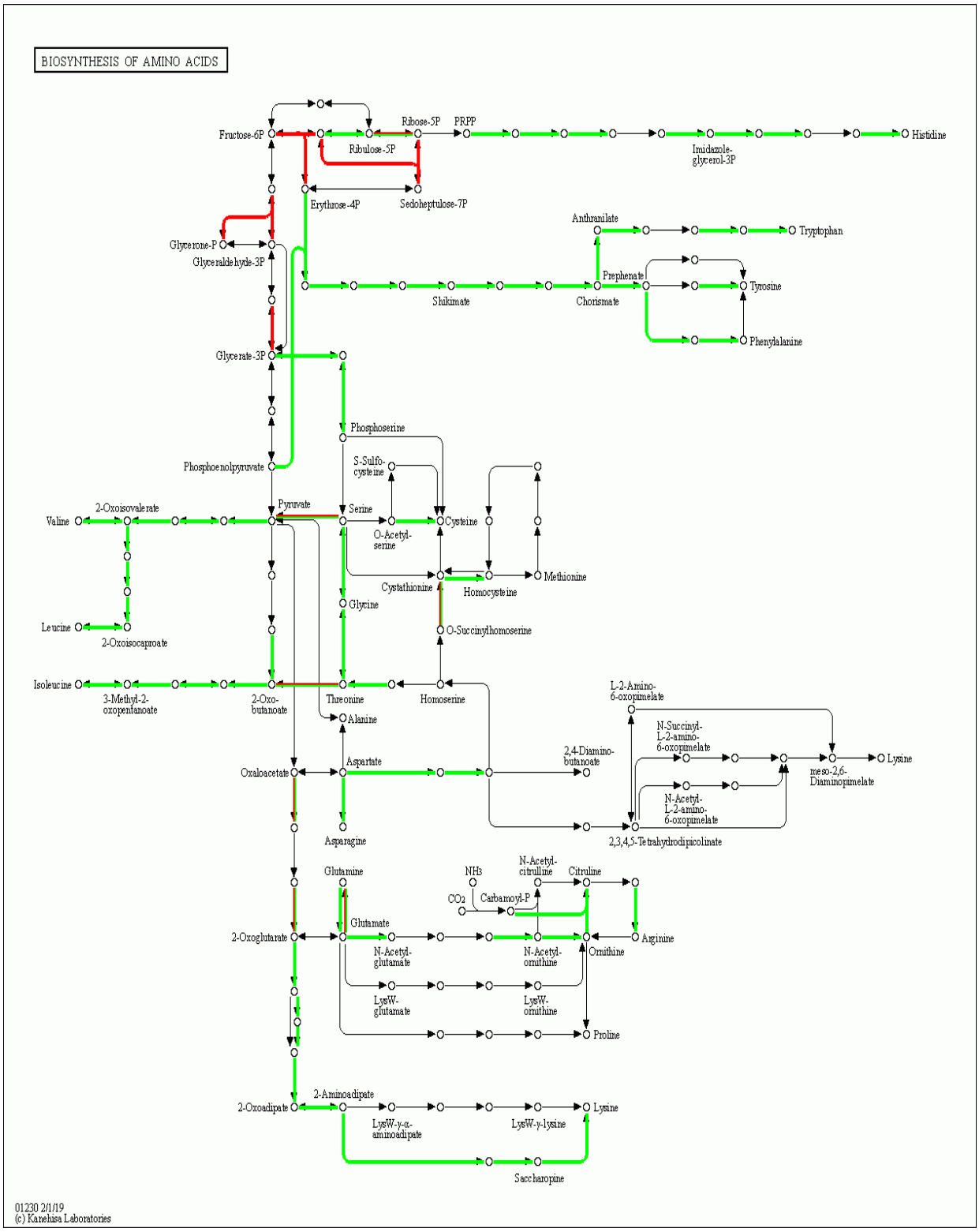


Fig. 1: Differentially expressed genes of Y22-vs-Y44 in amino acid synthesis pathway (red represents up-regulated reactions, green represents down-regulated reactions)

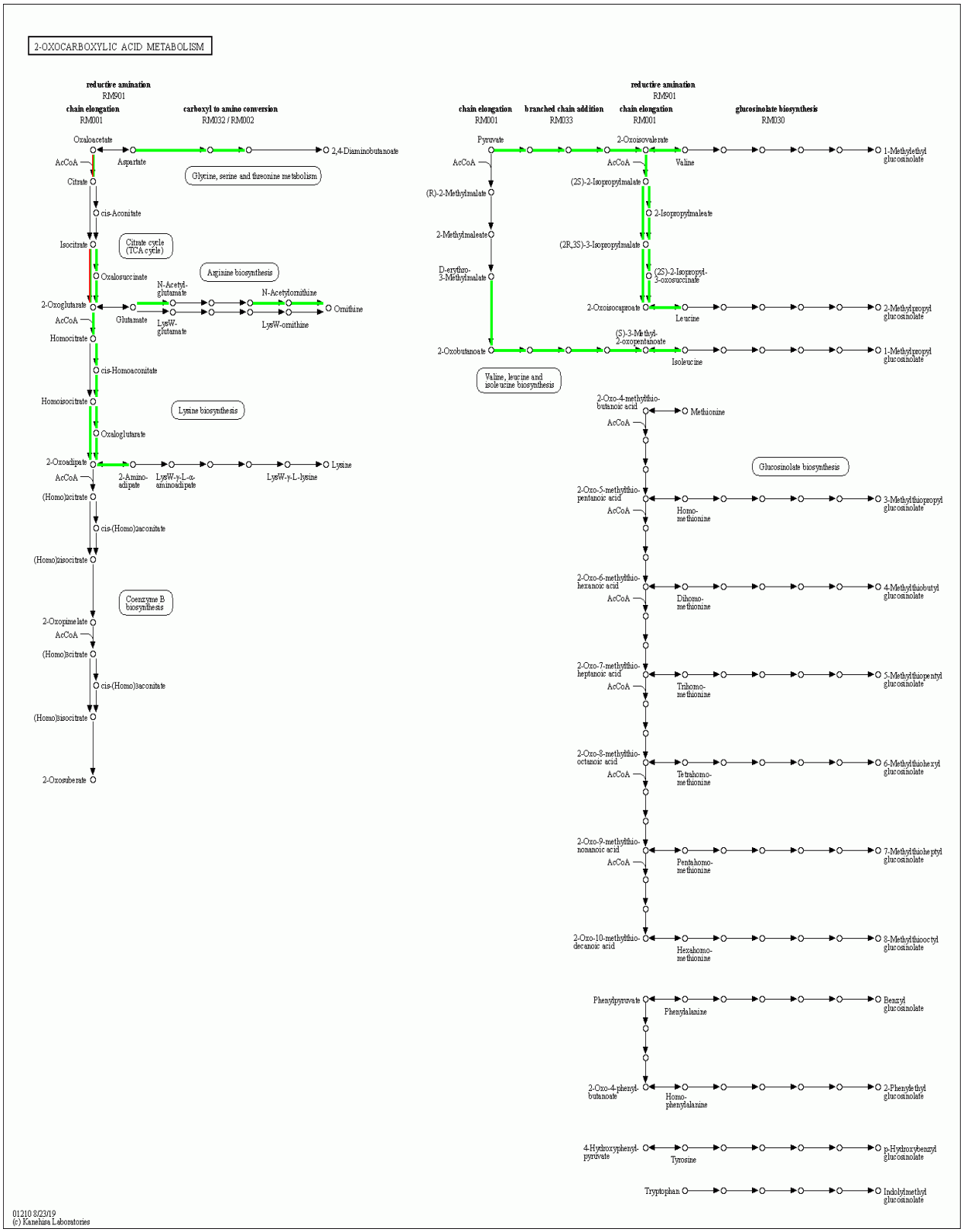


Fig. 2: Differentially expressed genes of Y22-vs-Y44 in ox carboxylic acid synthesis pathway (red represents up-regulated reactions, green represents down-regulated reactions)

Table 1: Differentially expressed genes in mutant strain Y44 compared with non-erythritol producing strain Po1g

Code of enzymes and related metabolic pathways	Description	Fold change
Oxocarboxylic metabolism		8.7038
2.3.3.1	citrate oxaloacetate-lyase	4.1956
2.3.3.13	oxobutanoate-lyase	8.7925
4.4.1.11	L-methionine methanethiol-lyase (deaminating; 2-oxobutanoate-forming)	5.3148
4.1.1.112	oxaloacetate decarboxylase	8.7925
1.4.7.1	glutamate synthase (ferredoxin); converted 2-oxoglutarate to glutamate	8.7038
1.1.1.286	isocitrate---homoisocitrate dehydrogenase; producing 2-oxoglutarate	3.6919
1.1.1.42	oxalosuccinate decarboxylase	5.2136
1.1.1.86	ketol-acid reductoisomerase; producing 3-oxobutanoate	4.0529
Amino acid metabolism		
1.4.1.2	L-glutamate:NAD+ oxidoreductase	5.2319
2.3.1.1	amino-acid N-acetyltransferase	4.1956
2.3.1.5	arylamine N-acetyltransferase	4.7832
2.1.3.3	ornithine carbamoyltransferase	5.3028
4.3.2.1	argininosuccinate lyase	6.3201
6.3.5.4	asparagine synthase	5.9485
6.3.5.5	carbamoyl-phosphate synthase (glutamine-hydrolysing)	3.6919
2.4.2.14	amidophosphoribosyltransferase	5.0496
1.2.1.11	aspartate-semialdehyde dehydrogenase	4.8509
2.6.1.42	branched-chain-amino-acid transaminase	3.5938
1.2.4.4	3-methyl-2-oxobutanoate dehydrogenase (2-methylpropanoyl-transferring)	4.3948
2.3.1.168	dihydrolipoyllysine-residue (2-methylpropanoyl) transferase	6.3095
6.4.1.4	methylcrotonoyl-CoA carboxylase	3.9484
4.1.3.4	hydroxymethylglutaryl-CoA lyase	8.3744
1.1.1.31	3-hydroxyisobutyrate dehydrogenase	6.2842
2.3.1.16	acetyl-CoA C-acyltransferase; producing L-cysteine;	5.2984
Tricarboxylic acid cycle		
1.1.1.37	malate dehydrogenase	6.2173
4.2.1.2	fumarate hydratase	4.7602
1.3.5.1	succinate dehydrogenase	5.2824
6.2.1.4	succinate---CoA ligase	5.0184
6.2.1.5	succinate---CoA ligase	7.2231
1.1.1.41	isocitrate dehydrogenase	6.3091
Pentose phosphate pathway		
2.2.1.1	transketolase	8.3209
4.1.2.13	fructose-bisphosphate aldolase	4.3984
2.7.7.9	UTP---glucose-1-phosphate uridylyltransferase	5.3122
1.1.1.307	Xylose reductase	6.3982
1.1.1.9	D-xylulose reductase	4.9572
1.1.1.250	arabinitol 2-dehydrogenase	3.8452

Note: Green represented down-regulated; red represented up-regulated