

RESEARCH

Open Access



# Comparative transcriptome and metabolome analyses reveal the methanol dissimilation pathway of *Pichia pastoris*

Yi-fan Yu, Jiashuo Yang, Fengguang Zhao, Ying Lin and Shuangyan Han\*

## Abstract

**Background:** *Pichia pastoris* (*Komagataella phaffii*) is a model organism widely used for the recombinant expression of eukaryotic proteins, and it can metabolize methanol as its sole carbon and energy source. Methanol is oxidized to formaldehyde by alcohol oxidase (AOX). In the dissimilation pathway, formaldehyde is oxidized to CO<sub>2</sub> by formaldehyde dehydrogenase (FLD), S-hydroxymethyl glutathione hydrolase (FGH) and formate dehydrogenase (FDH).

**Results:** The transcriptome and metabolome of *P. pastoris* were determined under methanol cultivation when its dissimilation pathway cut off. Firstly,  $\Delta fld$  and  $\Delta fgh$  were significantly different compared to the wild type (GS115), with a 60.98% and 23.66% reduction in biomass, respectively. The differential metabolites between GS115 and  $\Delta fld$  were mainly enriched in ABC transporters, amino acid biosynthesis, and protein digestion and absorption. Secondly, comparative transcriptome between knockout and wild type strains showed that oxidative phosphorylation, glycolysis and the TCA cycle were downregulated, while alcohol metabolism, proteasomes, autophagy and peroxisomes were upregulated. Interestingly, the down-regulation of the oxidative phosphorylation pathway was positively correlated with the gene order of dissimilation pathway knockdown. In addition, there were significant differences in amino acid metabolism and glutathione redox cycling that raised our concerns about formaldehyde sorption in cells.

**Conclusions:** This is the first time that integrity of dissimilation pathway analysis based on transcriptomics and metabolomics was carried out in *Pichia pastoris*. The blockage of dissimilation pathway significantly down-regulates the level of oxidative phosphorylation and weakens the methanol assimilation pathway to the point where deficiencies in energy supply and carbon fixation result in inefficient biomass accumulation and genetic replication. In addition, transcriptional upregulation of the proteasome and autophagy may be a stress response to resolve formaldehyde-induced DNA–protein crosslinking.

**Keywords:** Dissimilation pathway, Formaldehyde, Oxidative phosphorylation, Methanol metabolism, Glutathione

## Background

The methylotrophic yeast *Pichia pastoris* (*Komagataella phaffii*) is one of most commonly used expression systems. It can metabolize methanol as its sole carbon

and energy source [1–3]. Methanol is oxidized to formaldehyde by alcohol oxidase (AOX) in peroxisomes. Formaldehyde is further metabolized by assimilation or dissimilation. During assimilation, formaldehyde is directly fixed with dihydroxyacetone synthase (DAS) [4]. The entire methanol assimilation pathway is localized to peroxisomes. In the dissimilation pathway, formaldehyde is oxidized to S-formylglutathione by the NAD<sup>+</sup>-dependent enzyme formaldehyde dehydrogenase (FLD) in the first step, and S-formylglutathione

\*Correspondence: syhan@scut.edu.cn

Guangdong Provincial Key Laboratory of Fermentation and Enzyme Engineering, School of Biology and Biological Engineering, South China University of Technology, Guangzhou 510006, China



then reacts spontaneously with glutathione to form S-hydroxymethyl glutathione. Then, S-formylglutathione hydrolase (FGH) hydrolase hydrolyses this compound into formate and glutathione. Next, formate is oxidized to CO<sub>2</sub> by NAD<sup>+</sup>-dependent formate dehydrogenase (FDH) and released in vitro [5–8].

The dissimilation pathway metabolizes 2 molecules of NADH to the extent that it is considered to be the main source of power and energy. A surplus of FLD activity was shown to result in higher theoretical NADH formation rates and finally also in significantly improved butanediol production rates when a *P. pastoris* strain overexpressing FLD and butanediol dehydrogenase was applied for whole-cell biotransformation [9]. Additionally, by deleting the genes coding for dihydroxyacetone synthase isoforms 1 and 2 (DAS1 and DAS2), NADH regeneration via methanol oxidation (dissimilation) was increased significantly, which led to an increase in ATP and higher S-adenosylmethionine production [8, 10]. Therefore, an enhanced dissimilation pathway optimizes the energy distributions of methanol metabolism, guaranteeing more efficient NADH/product coupling and exploitation of both NADH steps for cofactor regeneration in whole-cell biotransformations [9].

Unfortunately, a carbon mass balance analysis revealed that 70–80% of the methanol metabolized is converted into carbon dioxide in the methanol fed-batch phase, when the methanol feed is in excess of the metabolic requirements in general [11–13]. This involves a huge loss of carbon atoms. Previous studies have reported the prevention of the conversion of methanol to CO<sub>2</sub> via formaldehyde and formate by the knockout of genes for enzymes related to the dissimilation pathway. However, the *Δfld* strain suffered severe growth defects upon the addition of formaldehyde [14]. Similarly, whether in the methanol induction stage or the subsequent fermentation stage, the *Δfdh* strain showed a lower biomass and a slightly slower methanol consumption rate too [15]. Knockout studies on FLD and FDH of methylotrophic yeast demonstrated that the FLD knockout phenotype is more severe than the FDH knockout phenotype, which was explained by the higher toxicity of formaldehyde compared to formate [14]. Simply knocking out or weakening one or more genes in the dissimilation pathway without making up for the loss of NADH and ATP, not only results in the accumulation of toxic intermediates but also causes energy imbalance. It is instructive to determine the role of the dissimilation pathway before balancing carbon flow distribution and energy supply in the production and application of engineered yeast [15, 16].

The growth defect of methylotrophic yeast on high methanol medium is not caused directly by methanol

toxicity but rather by formaldehyde, which is a key toxic intermediate of methanol metabolism [17]. As a strongly polarized reactive carbonyl compound, formaldehyde exists in various parts of the organism and plays an important role in cognitive ability and memory formation [18]. The toxicity of formaldehyde leads to partial synthetic lethality and growth defects in various cell lines utilizing methanol as a carbon source. Similarly, mice deficient in formaldehyde-metabolizing enzymes were found to develop partial synthetic lethality and mortality shortly after birth. This phenotype may be due to the accumulation of endogenous formaldehyde [19]. In fungi, plants or animals, the induction of DNA–protein crosslinking (DPC) formaldehyde is a topic of serious interest. Considering that DNA repair and RNA degradation pathways are evolutionarily conserved from yeast to humans, mechanisms of formaldehyde toxicity identified in yeast may be relevant to human disease and genetic susceptibility [20]. Formaldehyde dehydrogenase is critical to minimize the DPC issue, and gene-knockout organisms are commonly used to study the intracellular pathological changes of formaldehyde [19, 21, 22]. DPCs mainly consist of ribosomes and outer membrane proteins. Malfunction of these proteins may cause cell death because of outer membrane porin-induced programmed cell death or metabolic flux imbalance [23]. Membrane structure is a major target of methanol toxicity, while proteins are major targets of formaldehyde toxicity [24].

The dissimilation pathway is speculated to have two physiological functions: formaldehyde detoxification and energy production [4, 25, 26]. Based on these two points, we constructed three dissimilation pathway single gene-knockout strains by CRISPR/Cas9: a formaldehyde dehydrogenase-deficient strain (*Δfld*), an S-(hydroxymethyl) glutathione dehydrogenase-deficient strain (*Δfgh*) and a formate dehydrogenase-deficient strain (*Δfdh*). Then, we analysed key nodes and metabolic pathways by examining differentially expressed genes (DEGs) to discover the function of the dissimilation pathway and explore the stress response of yeast by comparing the transcriptome and metabolome of dissimilation pathway-defective strains and wild-type strains. Performing a whole-transcriptome analysis of dissimilation pathway knockout can promote the pathological study of formaldehyde metabolism and the development of industrial production strains.

## Results and discussion

### Phenotypic difference analysis

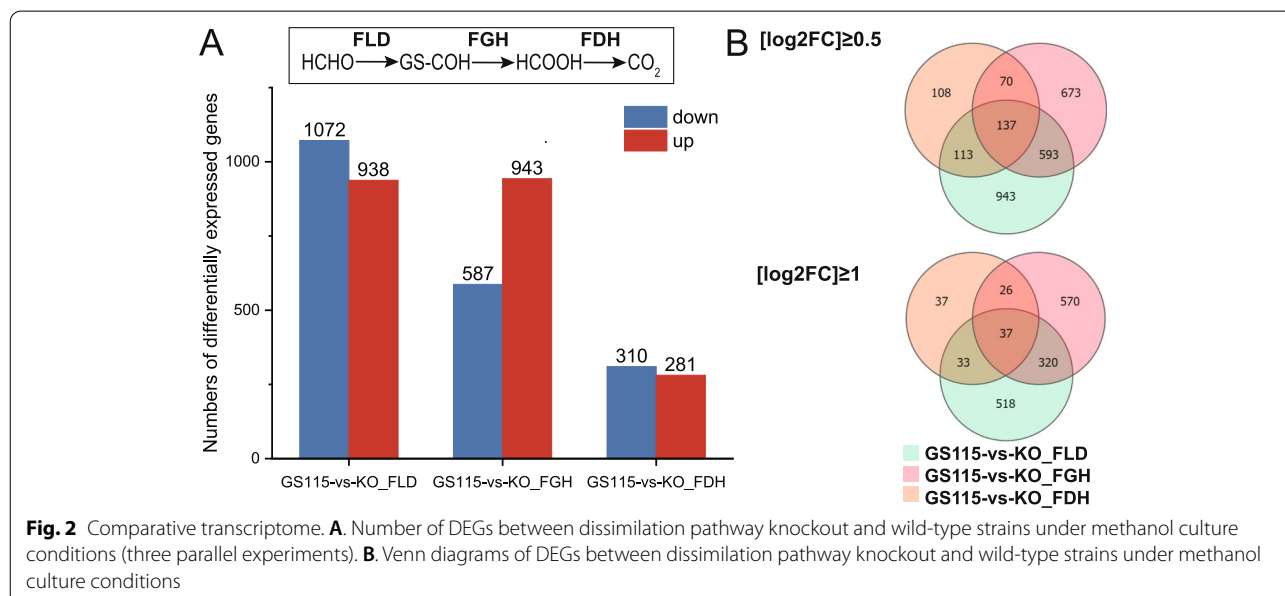
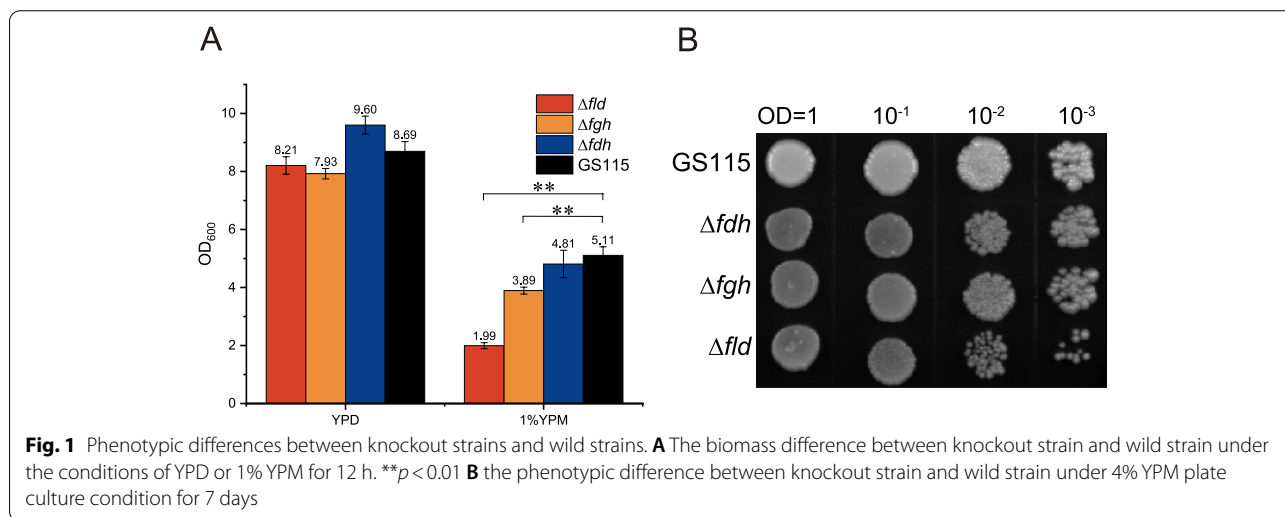
We knocked out enzyme genes of the dissimilation pathway (FLD (PAS\_chr3\_1028), FGH (PAS\_chr3\_0867) and FDH (PAS\_chr3\_0932)) separately in GS115 by CRISPR/Cas9 and found no homologous

sequences in the genome through BLAST [27]. With glucose as the main carbon source, there was no significant growth difference between dissimilation pathway knockout strains and GS115. However, with methanol as the main carbon source, the biomass of dissimilation pathway knockout strains ( $\Delta fld$ ,  $\Delta fgh$ ,  $\Delta fdh$ ) was lower (60.98%, 23.66%, 5.69%) than that of wild-type strain (GS115) (Fig. 1. A). GS115 showed concordance with  $\Delta fdh$  and significant differences with  $\Delta fld$  and  $\Delta fgh$  at 0.01 level ( $p < 0.01$ ) under YPM culture conditions. Therefore, the growth of GS115 was significantly higher than that of  $\Delta fld$  and  $\Delta fgh$  under methanol culture conditions. In addition, the 4% YPM plate used to monitor methanol tolerance between strains showed poor

growth of dissimilation pathway knockout strains, particularly  $\Delta fld$  (Fig. 1. B).

**Comparative transcriptomic**

Comparative transcriptomic were designed to reflect the effects of induced by gene knockout and methanol perturbations on metabolic pathways. The transcriptional profiles of  $\Delta fld$ ,  $\Delta fgh$ ,  $\Delta fdh$  and GS115 incubated in methanol for 12 h were developed and designated KO\_FLD, KO\_FGH, KO\_FDHD and GS115 respectively. GS115 and KO\_FDHD have the smallest sample variation. Then we compared the differences between GS115 versus KO\_FLD (GL), GS115 versus KO\_FGH (GG) and GS115 versus KO\_FDHD (GD). The



upward and downward DEG numbers for GL, GG, and GD were 938 and 1072, 943 and 587, and 281 and 310, respectively (Fig. 2. A). The change in the number of down-regulated genes was positively correlated with the knockout order, which is consistent with the results of the biological phenotype, demonstrating that FLD is a key enzyme gene in the dissimilation pathway of methanol metabolism. Interestingly, single knockouts did not show significant differences in transcript levels of other genes in the dissimilation pathway. The DEGs of GL and GG were analyzed by transcription factor (TF) ranking (Supplemental Table 2). Among the up-regulated genes, the zinc cluster transcriptional activator (CAY71800, CAY69410) was enriched ( $p < 0.05$ ); while among the down-regulated genes, carbon source-responsive zinc-finger transcription factor (CAY71743) and proposed transcriptional activator, member of the Gal4p family of zinc cluster proteins (CAY71429) was significantly enriched among the down-regulated genes ( $p < 0.01$ ).

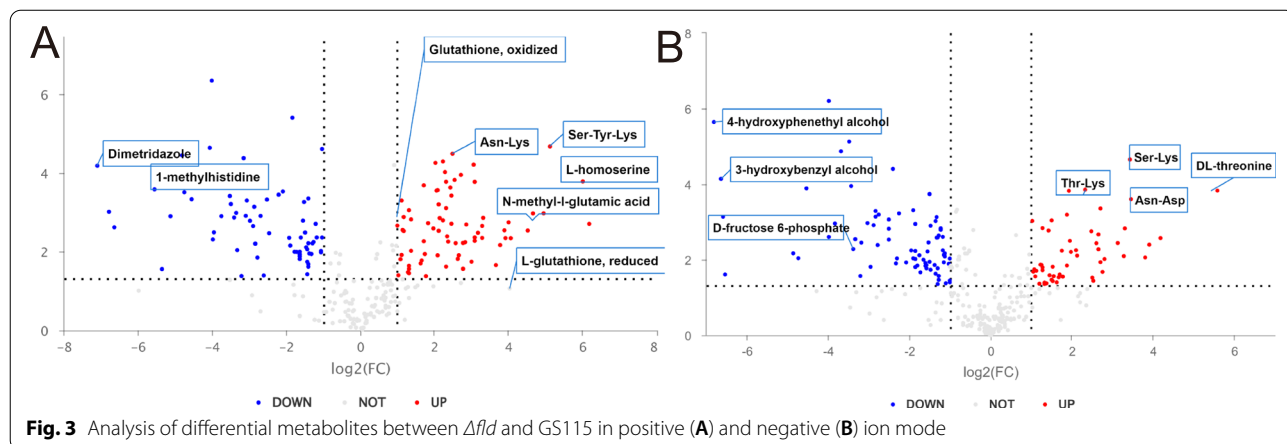
To understand the common impact of knockouts in the dissimilation pathway, we used Venn diagrams to show genes in different comparison groups. A total of 137 DEGs among the three groups were enriched in KEGG metabolic pathways ( $\log_2FC \geq 0.5$ ,  $q \leq 0.05$ ) (Fig. 2. B): glycolysis, TCA cycle, pentose and glucuronate interconversions, pyruvate metabolism biosynthesis of antibiotics, metabolism of various amino acids (tyrosine, phenylalanine, tryptophan, glycine, serine and threonine). A total of 357 DEGs among GL and GG were enriched in KEGG metabolic pathways ( $\log_2FC \geq 1$ ,  $q \leq 0.05$ ) (Fig. 2. B): most genes in the peroxisome pathway were transcriptionally upregulated, whereas a large proportion of genes in oxidative phosphorylation, glycolysis and TCA cycle were significantly downregulated. Furthermore, we were interested in genes involved in cysteine metabolism

( $p \leq 0.05$ ) and glutathione metabolism, related to formaldehyde binding.

**Comparative transcriptomics and metabolomics of  $\Delta fld$  and wild-type strain**

Formaldehyde dehydrogenase is the first enzyme in the dissimilation pathway, and its knockout can severely compromise strain robustness. We used BMM medium with methanol as the sole carbon source for 24 h to evaluate the metabolomic differences between  $\Delta fld$  and wild-type strain. All identified metabolites were classified: organic acids and derivatives (32.561%), lipids and lipid-like molecules (19.562%), and organoheterocyclic compounds (12.227%). The first three significantly abundant differential metabolic pathways were ABC transporters, amino acid biosynthesis, and protein digestion and absorption (Additional Fig. 1). Relatively in DEGs of GL, organonitrogen compound biosynthetic process, ribosome, cellular biosynthetic process, ATP metabolic process, peptide and oxoacid metabolic process were enriched for GO term; ribosome, Biosynthesis of antibiotics, glycolysis, biosynthesis of amino acids, oxidative phosphorylation, pentose and glucuronate interconversions were enriched for KEGG pathway.

Based on univariate analysis, the differences among all metabolites (including unidentified metabolites) detected in positive and negative ion modes were analysed. In the positive ion mode (Fig. 3. A), oxidized glutathione (GSSG) ( $\log_2FC \geq 1.17$ ) and reduced glutathione (GSH) ( $\log_2FC \geq 4.0$ ) were upregulated; metabolites associated with glutathione availability were generally upregulated, such as N-methyl-L-glutamate ( $\log_2FC \geq 4.6$ ), homoserine ( $\log_2FC \geq 6.0$ ), His-Glu ( $\log_2FC \geq 2.8$ ), and L-methionine ( $\log_2FC \geq 1.7$ ). Although the  $\Delta fld$  strain was unable to metabolise formalglutathione, the glutathione redox cycle was accelerated and may have reduced intracellular formaldehyde



**Fig. 3** Analysis of differential metabolites between  $\Delta fld$  and GS115 in positive (A) and negative (B) ion mode

levels. Among the differential metabolites, other different species of amino acids were up-regulated, such as Ser-Tyr-Lys ( $\log_2FC \geq 5.12$ ), His-Asp ( $\log_2FC \geq 2.0$ ), Ile-Lys ( $\log_2FC \geq 2.3$ ), Val-Lys ( $\log_2FC \geq 2.2$ ), Asp-Leu ( $\log_2FC \geq 4.0$ ), Pro pro ( $\log_2FC \geq 2.7$ ), etc. And Ile-Pro was downregulated ( $\log_2FC \leq -8.6$ ). Among the differential metabolites, other different species of amino acids were up-regulated, such as Ser-Tyr-Lys ( $\log_2FC \geq 5.12$ ), His-Asp ( $\log_2FC \geq 2.0$ ), Ile-Lys ( $\log_2FC \geq 2.3$ ), Val-Lys ( $\log_2FC \geq 2.2$ ), Asp-Leu ( $\log_2FC \geq 4.0$ ), Pro-pro ( $\log_2FC \geq 2.7$ ), etc. In addition, acetyl coenzyme A was down-regulated ( $\log_2FC \leq -1.4$ ) and its branch metabolite N-acetyl histidine was significantly down-regulated ( $\log_2FC \leq -5.1$ ). However, diacetylchitosan was upregulated ( $\log_2FC \geq 2.4$ ).

In the negative ion mode (Fig. 3. B), the methanol assimilation pathway metabolites, D-glucose-6-phosphate ( $\log_2FC \leq -1.3$ ) and glyceraldehyde ( $\log_2FC \leq -1.2$ ) were down-regulated. Meanwhile, D-fructose ( $\log_2FC \leq -1.5$ ), D-gluconate ( $\log_2FC \leq -1.7$ ), glucitol ( $\log_2FC \leq -1.8$ ), and malate ( $\log_2FC \leq -0.5$ ) were down-regulated. D-ribose ( $\log_2FC \leq -2.7$ ) and thymine ( $\log_2FC \leq -3.4$ ) were down-regulated, indicating that *Δfld* is restricted during genetic replication, which may severely affect normal cell growth.

Through the interaction of DEGs of amino acid metabolism, aldehyde dehydrogenase (NAD<sup>+</sup>) (PAS\_chr4\_0043), glutamate dehydrogenase (NADP<sup>+</sup>) (PAS\_chr1-1\_0107), ornithine decarboxylase (PAS\_chr3\_0417) were screened out as important genes. The assimilation pathway and central carbon metabolites were down-regulated in *Δfld*, which may affect intracellular organic carbon fixation. The up-regulation of chitinose may be due to the resistance of chitin to oxidative stress. In contrast, most species of amino acids in the differential metabolites are up-regulated when synthesized from scratch. Therefore, knockdown of FLD, although not efficient in carbon utilization, facilitates amino acid metabolism and product synthesis.

### Oxidative phosphorylation

The 60 DEGs from the three groups in the oxidative phosphorylation pathways were clustered. The degree of oxidative phosphorylation of GD was similar, while GL and GG were significantly down-regulated (Fig. 4).

F-type H<sup>+</sup>-transporting ATPase subunit beta (ATP1 and ATP2, PAS\_chr3\_0576 and PAS\_chr2-2\_0165) were conspicuous downregulated in GL and GG ( $\log_2FC \leq -1.5$ ,  $q \leq 0.05$ ). The difference in fold change between RNA-seq and qRT-PCR of Oxidative phosphorylation is demonstrated in Fig. 5. ATP1 and ATP2 are proton-transporting ATP synthase complexes with nucleoside phosphatase activity that bind to purine

nucleotides and act in nucleotide biosynthetic processes and nitrogen compound metabolism. In comparative metabolomics, differences in purine metabolism and pyrimidine metabolism of nucleotides may be associated with transcriptional down-regulation of mitochondria-related enzyme subunits. This leaves ATP in limited supply.

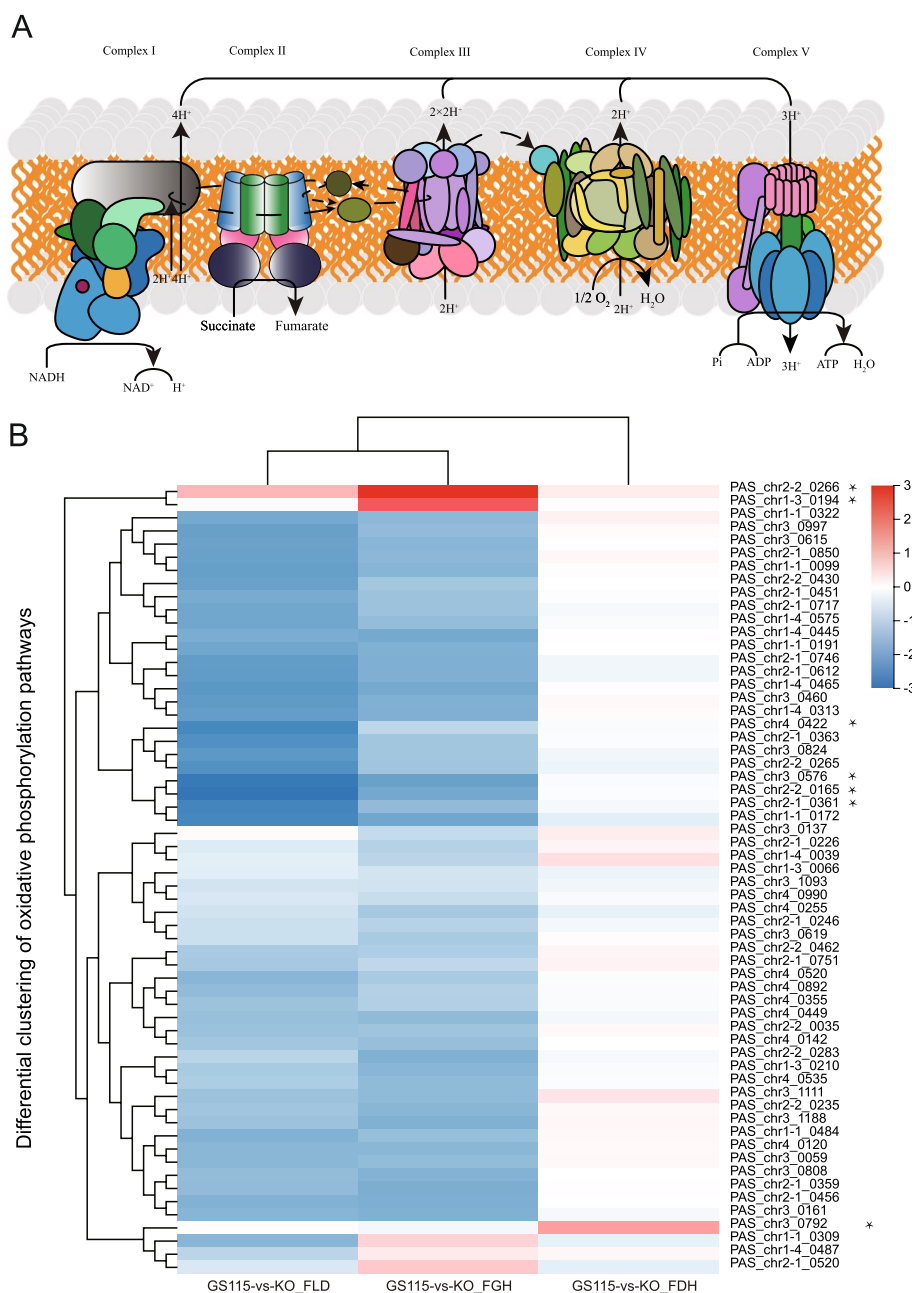
Cytochrome c oxidase subunit (COX5B and COX6B, PAS\_chr2-1\_0361 and PAS\_chr4\_0422) were down-regulated in GL ( $\log_2FC \leq -2.5$ ,  $q \leq 0.05$ ) (Fig. 5). This severely diminished the mitochondrial-cytoplasmic proton transmembrane transporter protein activity. The transcript levels of the ubiquitin-cytochrome c reductase core subunit (e.g. UQCRC2, PAS\_chr2-2\_0430), which has catalytic activity for binding cations and metal ions, were down-regulated ( $\log_2FC \leq -1.9$ ,  $q \leq 0.01$ ), reducing the catalytic efficiency of the enzyme during oxidative phosphorylation. Transcript levels of NADH dehydrogenase complex subunits (e.g. NDUFB7, PAS\_chr1-1\_0172) were significantly down-regulated in GL ( $\log_2FC \leq -2.4$ ) and GG ( $\log_2FC \leq -1.8$ ,  $Q \leq 0.02$ ). It uses NAD(P)H, quinone or similar compounds as receptors and has oxidoreductase activity. Down-regulated genes in the oxidative phosphorylation pathway, such as the NADH dehydrogenase complex subunit, are associated with metabolic pathways in human Alzheimer's disease, cardiovascular disease and non-alcoholic fatty liver disease. The dissection of the mechanisms underlying strains defective in the dissimilatory pathway provides a theoretical basis for human pathology.

Screened for highly expressed genes ( $\log_2FC \geq 2$ ,  $q \leq 0.05$ ) in GL, revealed the upregulation of some dehydrogenase transcripts after FLD knockout, such as pyridoxine 4-dehydrogenase (PAS\_chr4\_0550), alcohol dehydrogenase (NADP<sup>+</sup>) (PAS\_chr3\_0006), NADPH2 dehydrogenase (PAS\_chr3\_1184), D-arabinose 1-dehydrogenase (PAS\_chr2-1\_0775), which may imply that other dehydrogenases compensate for the function of formaldehyde dehydrogenase (not shown in the picture).

In GG, cytochrome c oxidase subunit 7c (COX7C, PAS\_chr2-2\_0266) ( $\log_2FC \geq 2.9$ ,  $q \leq 0.01$ ) and haem o synthase (COX10, PAS\_chr1-3\_0194) ( $\log_2FC \geq 2.0$ ,  $q \leq 0.01$ ) were significantly upregulated. Others were downregulated. In GD, only NADH dehydrogenase (NDH, PAS\_chr3\_0792), a mitochondrial external NADH dehydrogenase or type II NAD(P)H: quinone oxidoreductase, was upregulated ( $\log_2FC \geq 1.2$ ,  $q \leq 0.01$ ). It might compensate for the absence of FDH.

### Methanol metabolism pathway

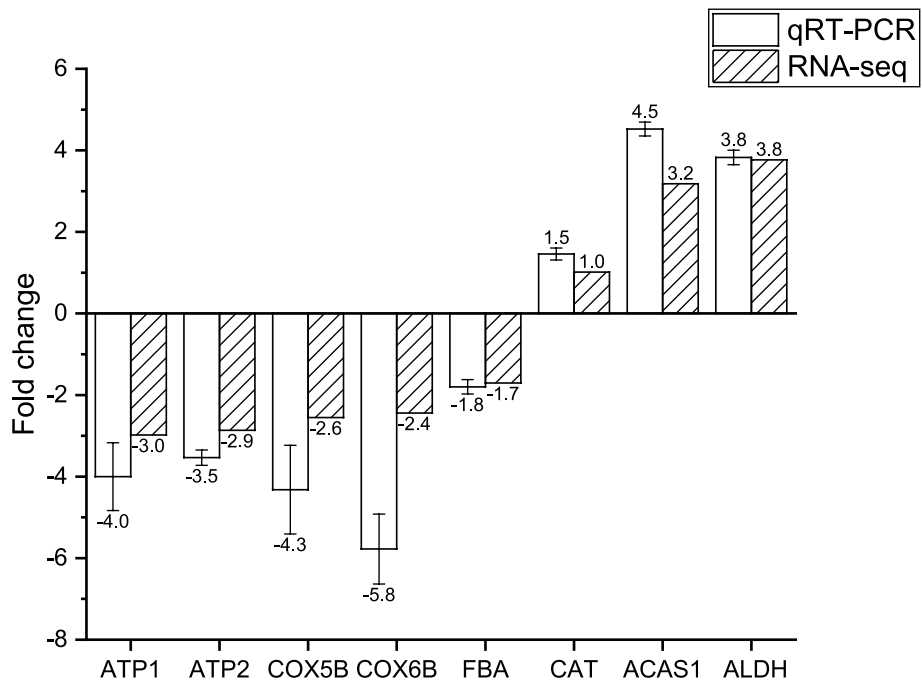
The main carbon metabolism pathways in *P. pastoris* include methanol metabolism, glycolysis, the TCA cycle, the pentose phosphate pathway and ethanol metabolism.



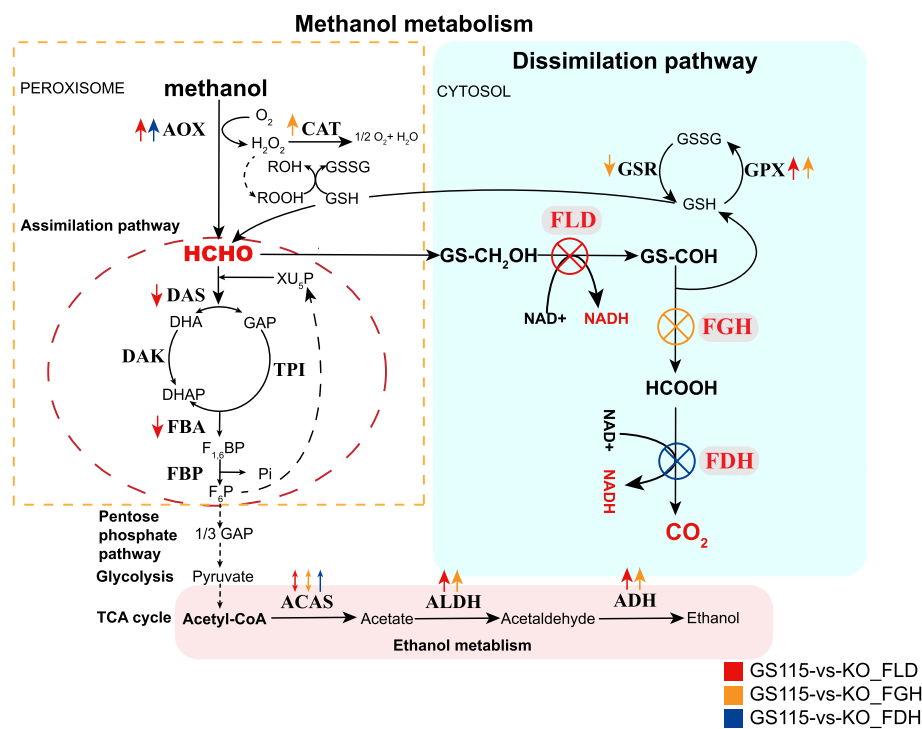
**Fig. 4** Oxidative phosphorylation pathway under methanol culture conditions. **A** Dialogue with oxidative phosphorylation in *Pichia pastoris* [28]. **B** DEGs of oxidative phosphorylation between dissimilation pathway knockout and wild-type strains

By comparing transcriptomes, we attempted to explain the variation in the methanol metabolism pathway after knockout of the dissimilation pathway genes (Fig. 6). In peroxisomes, alcohol oxidase 2 (AOX2, PAS\_chr4\_0152) was upregulated in GL ( $\log_2FC \geq 1.3$ ,  $q \leq 4.96E-12$ ) and GD ( $\log_2FC \geq 0.9$ ,  $q \leq 1E-04$ ). AOX2 is induced by methanol, which may lead to an increase in formaldehyde content in the peroxisome.

Knockout of the dissimilation pathway gene down-regulates the assimilation pathway of formaldehyde. The dihydroxyacetone synthase (DAS, PAS\_chr3\_0832 and PAS\_chr3\_0834,  $\log_2FC \leq -0.8$ ) and fructose-bisphosphate aldolase (FBA, PAS\_chr1-1\_0072,  $\log_2FC \leq -1.7$ ) of the assimilation pathway were significantly down-regulated when FLD was knocked out, which may be one of the reasons for the low biomass of  $\Delta fld$  (Fig. 5).



**Fig. 5** Fold change between the RNA-seq and qRT-PCR of DEGs



**Fig. 6** Methanol metabolism pathway and comparison of transcription levels in *Pichia pastoris*. The arrow indicates the comparative transcription level of the dissimilation pathway knockout and wild-type strains ( $q \leq 0.05$ )

Correspondingly, D-glucose 6-phosphate was downregulated ( $\log_2FC \leq -1.3, VIP \geq 1.8$ ). This again validates the prominence of FLD in the dissimilation pathway.

When FGH was knocked out, catalase (CAT, PAS\_chr2-2\_0131,  $\log_2FC \geq 1$ ) was upregulated, which means reactive oxygen species (ROS) induced a pronounced oxidative stress response (Fig. 5). Superoxide dismutase [Cu-Zn] (PAS\_chr4\_0786), which destroys radicals normally produced within cells that are toxic to biological systems, was upregulated significantly in GL ( $\log_2FC \geq 1.0, q \leq 0.01$ ) and GG ( $\log_2FC \geq 2.5, q \leq 0.01$ ).

Previous studies have shown that under methanol culture conditions, genes encoding methanol metabolism are upregulated, and glycolysis and TCA cycle transcription are downregulated [29]. Significant transcriptional downregulation of genes involved in glycolysis and the TCA cycle was observed in GL and GG. Interestingly, malate dehydrogenase (MDH2, PAS\_chr4\_0815) was transcriptionally upregulated in GL, GG ( $\log_2FC \geq 2.7, q \leq 7.93E-06$ ) and GD. Genes involved in alcohol metabolism were significantly transcriptionally upregulated ( $q \leq 0.05$ ). Among them, transcription of acetyl-coenzyme A synthetase (ACAS1, PAS\_chr3\_0403) and alcohol dehydrogenase (ADH, PAS\_chr1-3\_0153) were upregulated ( $\log_2FC \geq 1, q \leq 0.05$ ) (Fig. 5). Aldehyde dehydrogenase (NAD<sup>+</sup>) (ALDH, PAS\_chr3\_0987) ( $\log_2FC \geq 3.7, q \leq 6.61E-07$ ) was significantly upregulated in GG (Fig. 5). However, ACAS2 (PAS\_chr2-1\_0767) was downregulated in GL and GG ( $\log_2FC \leq -1.4, q \leq 6.61E-07$ ). This may mean that the two ACASs are responsible for different metabolic pathways in *Pichia pastoris* and ACAS2 is more affected by methanol induction.

#### **Glutathione redox cycling and amino acid metabolism**

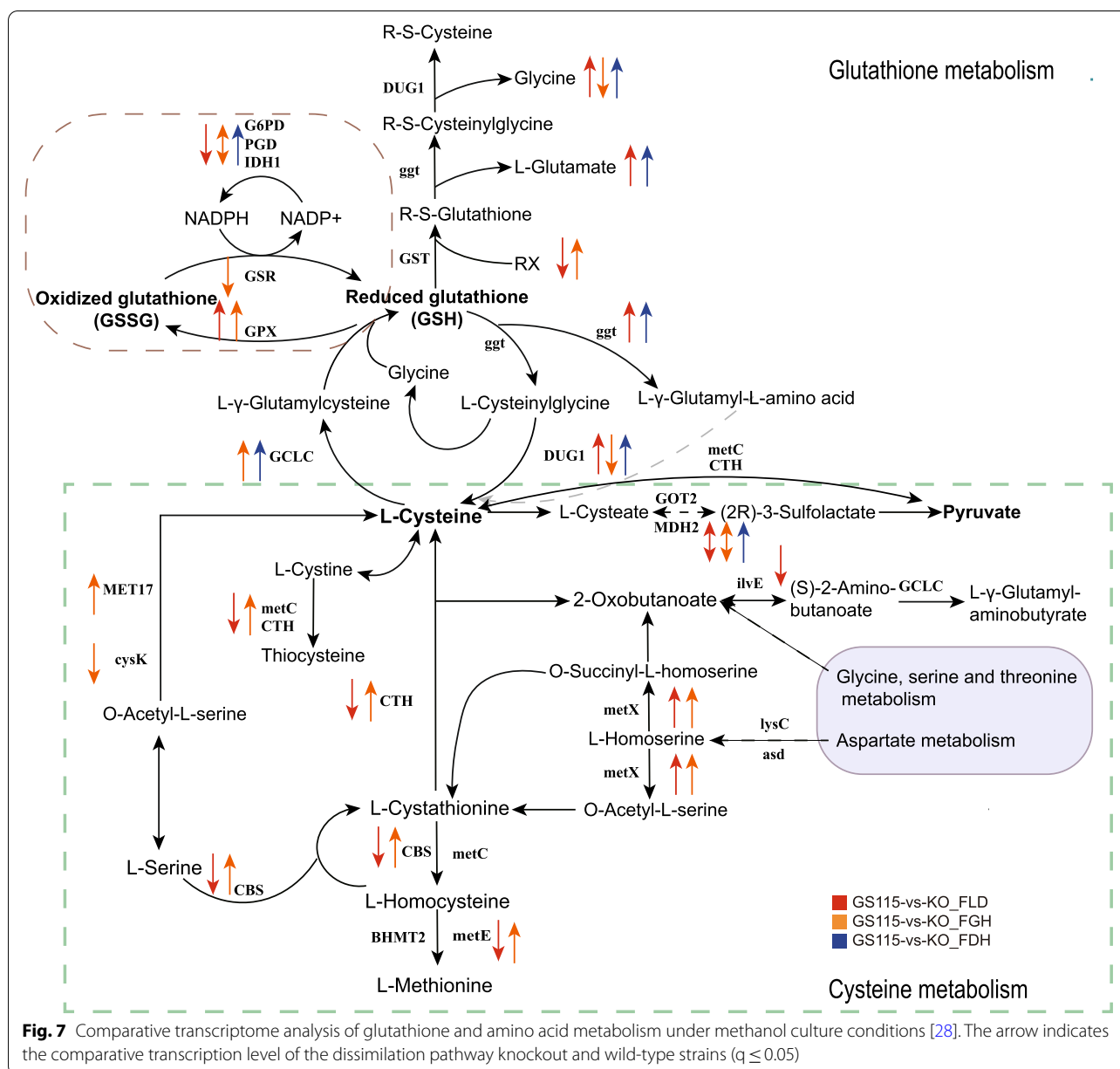
Glutathione (GSH, L- $\gamma$ -glutamyl-L-cysteinylglycine) is the main sulfur compound and appears as the major non-protein thiol compound in yeasts. In cells, glutathione mainly exists in the reduced form GSH, as oxidized glutathione (GSSG) is converted rapidly by glutathione reductase. In the process of glutathione reduction and oxidation (Fig. 7), glutathione peroxidase (GPX, PAS\_chr2-2\_0382) was upregulated ( $\log_2FC \geq 0.6, q \leq 0.05$ ), while glutathione reductase (NADPH) (GSR, PAS\_chr3\_1011) was downregulated ( $\log_2FC \leq -0.6, q \leq 0.05$ ) in GG. In the positive ion mode, GSH and GSSG were upregulated ( $\log_2FC \geq 4.0$  and  $\log_2FC \geq 1.1, VIP \geq 1$ ) when FLD was knocked out. Cystathionine gamma-lyase (CTH, PAS\_chr1-4\_0489) had zero expression in GS115\_M. CTH, an enzyme involved in sulfur compound metabolism and cysteine metabolism, showed significantly upregulated transcription in GG ( $\log_2FC \geq 10, q \leq 2.7E-54$ ).

We clustered the 149 DEGs in the amino acid metabolism. The knockout of each dissimilation gene produced a comparable up and down regulation trend, which may reveal the effect on the amino acid pathway such as arginine, proline, valine, leucine, isoleucine, lysine, glycine, serine and threonine metabolism. 5-Methyltetrahydropteroyltriglutamate-homocysteine methyltransferase (metE, PAS\_chr2-1\_0160) was upregulated in GG ( $\log_2FC \geq 1.6, q \leq 0.05$ ). Proline dehydrogenase (PRODH, PAS\_chr1-3\_0269) is upregulated in order of knockout by sequence ( $\log_2FC \geq 1.4, 2.8, 3.4, q \leq 0.01$ ), which facilitates the process of proline catabolism to glutamate. The glutamate-cysteine ligase catalytic subunit (GCLC, PAS\_chr1-1\_0184) is transcriptionally upregulated in GG and GD ( $\log_2FC \geq 0.5, q \leq 0.01$ ). Glutathione S-transferase (GST, PAS\_chr2-1\_0490) was significantly upregulated in GG ( $\log_2FC \geq 1.6, q \leq 0.01$ ) and downregulated in GL ( $\log_2FC \leq -1.3, q \leq 0.05$ ). Cys-Gly metallopeptidase (DUG1, PAS\_chr3\_0353) was significantly upregulated in GL ( $\log_2FC \geq 2.1, q \leq 1.5E-07$ ). After FLD knocked out, L-homoserine was significantly upregulated ( $\log_2FC \geq 6.0, VIP \geq 4.9$ ). Some amino acids, like Ser-Tyr-Lys, His-Asp, His-Glu, Ile-Lys, Val-Lys were upregulated ( $VIP \geq 1$ ). And 1-methylhistidine was downregulated ( $\log_2FC \leq -5.5, VIP \geq 1$ ). GO enrichment analysis revealed that in GL and GG, genes involved in the metabolic and biosynthetic processes of organic acids and carboxylic acids were transcriptionally downregulated, whereas in GD, they were transcriptionally upregulated ( $q \leq 0.01$ ). Another difference is that primary amine oxidase (AOC3, PAS\_chr1-4\_0441, PAS\_chr2-1\_0307 and PAS\_chr4\_0621) and amidase (amiE, PAS\_chr3\_0283) are upregulated in GL and GD. The knockout of two dehydrogenases in the dissimilation pathway may have affected the metabolism of organic nitrogen compounds.

#### **Upregulation of proteasomes, peroxisomes and autophagy**

The knockout of genes in the dissimilation pathway caused upregulation of peroxisomes and autophagy. The transcription level of the DNA-dependent metalloprotease WSS1 (PAS\_chr3\_0200) increased significantly ( $\log_2FC \geq 1.1$  (GG),  $\log_2FC \geq 1.9$  (GL),  $q \leq 0.01$ ). The protein component of DPCs is targeted for repair by proteases of the Wss1/SPRTN family. This indicated that protease digestion of DPC was a stress response to formaldehyde. In GG, ubiquitin C (PAS\_chr4\_0762), AN1-type domain-containing protein (PAS\_chr4\_0567), 20S proteasome subunit alpha 2 (PAS\_chr1-1\_0433), which is in ubiquitin-proteasome pathway, was significantly upregulated; the difference was not significant in GD.





**Discussion**

**Formaldehyde dehydrogenase is the key enzyme in the dissimilation pathway**

Regardless of phenotypic differences or transcriptome differences, the results after FLD knockout are the most significant. FLD coordinates the formaldehyde level in methylotrophic strains according to the methanol concentration during growth [30]. After FLD is knocked out, the relevant enzyme genes of the assimilation pathway are downregulated, and the content of D-glucose 6-phosphate is decreasing. The assimilation and

dissimilation of formaldehyde have a synergistic effect, and the down-regulation of the assimilation reduced biomass. Moreover, the difference in organic nitrogen metabolism also proves that FLD is a key enzyme that uses methanol and/or methylamine as carbon and/or nitrogen sources [16].

**Driving energy is the main function of the dissimilation pathway**

Our transcripts validate that the level of oxidative phosphorylation is significantly down-regulated after FLD and FGH knocked out. Although knockout of FLD or FDH

reduced NADP supply, the physiological role of FDH was revealed to be mainly detoxification of formate rather than stimulated energy generation [31]. The majority of methanol is metabolized via an energy-generating dissimilation pathway, leading to a corresponding increase in mitochondrial size and number [12]. FLD knockout significantly affected ATP production. FLD activity was identified as the main bottleneck of effective recovery of NADH through methanol dissimilation. In the engineered strain modification, butanediol productivity was improved by increasing FLD activity [9]. The energy driving effect of FLD and FGH on methanol metabolism is important for yeast biosynthesis.

#### **Adsorption of formaldehyde by glutathione and amino acids**

GSH scavenges cytotoxic  $H_2O_2$  and maintains a redox balance in the cellular compartments in the response of yeasts to different nutritional and oxidative stresses [32, 33]. GSH/GSSG ratio rose, suggesting stronger protection against oxidative stress, and was also correlated with high GSR activity [34, 35]. FGH is distributed between peroxisome and cytoplasm to release GSH in the dissimilation pathway [36]. An important reaction by formaldehyde is the formation of compounds with the tripeptide glutathione; between 50 and 80% of endogenous formaldehyde occurs in the form of compounds that include glutathione [16, 37]. Knockout of FLD increased the amount of GSSG and GSH compared with GS115. However down-regulation of GSH/GSSG ratio may increase intracellular oxidative stress. Interestingly, the up-regulation of CAT transcription confirms this when FGH was knockout so that GSH could not be restored. GPX transcription are upregulated in GL and GG. Knockout of FLD and FGH may accelerate the GSH redox cycle. This may mean that accelerated combination of glutathione and formaldehyde still cannot compensate for the impact of knocking out the dissimilation pathway.

The formation of S-[1-(N<sup>2</sup>-deoxyguanosinyl) methyl] glutathione between glutathione and DNA was induced by formaldehyde [32]. The involvement of the cysteine residue of GSH in coupling suggests that other thiols may participate in the formation of this type of DNA damage from formaldehyde [32]. The reactions of formaldehyde with cysteine and histidine are alternative routes of formaldehyde detoxification metabolism [33]. The up-regulated amino acid metabolites were detected after FLD knockout. In the context of the upregulation of many amino acid biosynthesis genes or proteins, this suggests an increased flux towards amino acid and protein synthesis [38]. Knockout of the dissimilation pathway caused differences in amino acid metabolism, but the internal

associations and regular pattern need to be explored. The exploration of amino acid metabolism and methanol metabolism or formaldehyde dehydrogenase knockout samples is of great significance for revealing the toxicity of formaldehyde and physiological mechanism.

#### **DPC caused by formaldehyde during methanol metabolism**

Formaldehyde is the key intermediate of methanol detoxification metabolism and the origin metabolite of the dissimilation pathway. We speculate that the upregulation of peroxisomes accelerates the exchange of formaldehyde, what is more, upregulation of proteasome and autophagy transcription may be to solve the DPC caused by formaldehyde.

In eukaryotes, the autophagy–lysosome system and the ubiquitin–proteasome system (UPS) are the two major quality control pathways responsible for maintaining proteome homeostasis and directing recycling to meet nutrient demand. Formaldehyde exposure triggers widespread ubiquitylation events in cells [39, 40]. The polymerization of ubiquitin, a key molecule known to work in concert with the proteasome, served as a degradation signal for numerous target proteins [41, 42]. Proteasome cleaved the protein moiety of DPC to remove the lesion. Additionally, when FLD is knocked out, differential metabolites are enriched in purine and pyrimidine metabolism. Nucleotide excision repair and homologous recombination are involved in repairing DPCs [43].

In contrast, autophagy can eliminate larger protein complexes, insoluble protein aggregates, and even entire organelles and pathogens in toto due to the sheer size of the engulfing autophagic vesicles [41]. Methanol oxidation stress results in the accumulation of hydrogen peroxide and organic injury [44, 45]. Excessive ROS are formed from peroxisome metabolism when methanol-grown wild-type cells are exposed to excess methanol [46]. The cut-off of the dissimilation pathway may not be able to carry out effective methanol metabolism and thus aggravate the accumulation of ROS. The upregulation of the autophagy pathway may be related to the degradation of damaged peroxisomes or other components of eukaryotic cells [29, 47].

Mutations in the gene encoding the putative human homologue of a yeast DPC protease cause premature ageing and cancer predisposition syndrome in humans [48]. In *Saccharomyces cerevisiae*, several genes involved in DNA repair (eight RAD genes) that have been identified as specific for methanol toxicity were previously reported as determinants of tolerance for formaldehyde [49]. The up-regulation of DNA repair genes may be caused by the accumulation of formaldehyde by knocking out FLD and FGH. It is also of interest that genes related to DPCs in *P.*

*pastoris* could be mining. However, it is also a challenge to determine the concentration of formaldehyde in peroxisomes of yeast.

**Conclusions**

The main function of the dissimilation pathway may be the supply of energy rather than formaldehyde detoxification. Knockout of FLD and FGH significantly downregulated oxidative phosphorylation, leaving the strains with inadequate energy supply and low biomass. Knockdown of the dissimilation pathway leads to downregulation of gene expression in glycolysis and the TCA cycle. In particular, the assimilation pathway was downregulated. Differences in amino acid and glutathione metabolism were apparent, probably due to the adsorption of amino acid and glutathione to formaldehyde to resolve formaldehyde accumulation and oxidative stress responses. In addition, knockout of the dissimilation pathway enhanced the stress response to formaldehyde. Upregulation of the proteasome, peroxisome and autophagy may be the solution of DPC induced by formaldehyde.

**Methods**

**CRISPR/Cas9**

*P. pastoris* GS115 was used as the host for CRISPR/Cas9-based genome editing studies. Efficient CRISPR/Cas9-mediated genome editing with a type III promoter (i.e., SER promoter) to drive the expression of the single guide RNA (sgRNA) was achieved [50]. Using CRISPRdirect (<http://crispr.dbcls.jp/>) to design a targeting sequence (Table 1), a CRISPR plasmid vector (Cas9-III promoter-targeting sequence-gRNA) was successfully constructed [51]. The donor was synthesized with the upper and lower 1000 bp of the target gene (Table 2). Donor and Cas9 plasmids were transformed into *Pichia pastoris* for expression (100 µg/mL ampicillin supplemented when necessary). The successful knockout strains were screened and cultured in YPD for a period of time to remove the plasmids. Single colonies were picked and crushed in a metal bath at 80 °C in yeast crushing solution. After a short centrifugation, 1 µL of supernatant was taken as PCR template and PCR identification was performed using

**Table 1** PAM site of dissimilation pathway knockout gene

Target genes	PAM (5'-3')
FLD	CGGAAACGAGAAGTCCCGTC
FGH	AACCCCACTAAAGCCCCATG
FDH	GAGGTTACGGTGGTGACGTC

**Table 2** Donor primers for PCR

Donor primers	sequence(5'-3')
FLD-DONOR-F	GCTGTCAGTTCTGCGTCAACATGGGC
FLD-donor-1-R	GTACGACGTATGATGAATGAATGAGTTATGTAAGGCC
FLD-donor-2-F	CTCATTTCATTCATACATCGTACCTTCCCTCCCA
FLD-DONOR-R	GCCGCAACTTTAATTATCCGTGGATTGGGAGCT
FGH-DONOR-F	CACCATACAGGTCTCCCTTCGATACCAGTGCAAAGT
FGH-donor-1-R	TGAATGTAGAAAGATGGAATGAACGCTACAGCGAGGAAG
FGH-donor-2-F	TCCATCTTTCTACATTCACAGACGTTTCATGCTGCTC
FGH-DONOR-R	CTTCTAGTCCCCCATTTGTGGCTTACGTAA
FDH-DONOR-F	AAATGGCAGAAGGATCAGCCTGGACGAAGCAACCAG
FDH-donor-1-R	TACCGTTCATGTTTAAAGTGGTGATGTTGGAGG
FDH-donor-2-F	CACTTAAACATGAACGGTAAGTACAAGACCAAGGC
FDH-donor-R	GCCTCAACAATTGGCAGCTCTTCTACGGT

xxx-DONOR-F and xxx-DONOR-R. The PCR products of wild strain GS115 were 3085 bp (FDH), 2866 bp (FGH) and 3007 bp (FLD) in length. The successful knockout resulted in product fragment lengths of 2040 bp (*Δfdh*), 2019 bp (*Δfgh*) and 1806 bp (*Δfld*).

**Strains and cultivation**

*P. pastoris* GS115 (Invitrogen, USA), a histidine auxotrophic strain, was used as the host strain for the construction of recombinant strains [52]. Methylotrophic yeast grows well in a liquid medium containing 1% (vol/vol) methanol as the carbon source to induce the production of heterologous proteins. Culture conditions for transcriptome analysis: the strains were grown on YPD agar medium containing 2% agar, 1% yeast extract, 2% peptone and 2% glucose to obtain the seed liquid. In YP liquid medium (1% yeast extract, 2% peptone) containing 1% (vol/vol) methanol (YPM medium) or 2% glucose (YPD medium), the cells were cultured for 12 h to the exponential logarithmic phase of growth [53]. Culture conditions for metabolome analysis were as follows: in BMM liquid medium containing 100 mM potassium phosphate (pH 6.0), 1.34% YNB, 4 × 10<sup>-5</sup>% biotin and 0.5% methanol, the cells were cultured for 24 h.

**Methanol tolerance plate test**

The strains were cultured in YPD for 24 h. The seed liquid was diluted to OD<sub>600</sub> of 1 with YP liquid medium and diluted to different times in 96-well plates. 3 µl of diluent is used for spotting in 4% YPM plate and incubate for 7 days.

**Transcriptome and metabolome sample testing**

The cultured cells were immediately frozen in a liquid nitrogen tank, and the samples were collected and

sent to BGI Genomics Co.Ltd (Wuhan, China) for RNA extraction and transcriptome sequencing, and Shanghai Applied Protein Technology Co.Ltd (Shanghai, China) for metabolite group detection. Data are presented as the mean of three replicates  $\pm$  standard deviation.

### Transcriptome data analysis

The data was tested on the DNBSEQ platform. The sequencing data was filtered with SOAPnuke (v1.5.2) by (1) Removing reads containing sequencing adapter; (2) Removing reads whose low-quality base ratio (base quality less than or equal to 5) is more than 20%; (3) Removing reads whose unknown base (N' base) ratio is more than 5%, afterwards clean reads were obtained and stored in FASTQ format [54]. The clean reads were mapped to the reference genome using HISAT2 (v2.0.4) [55]. Bowtie2 (v2.2.5) was applied to align the clean reads to the reference coding gene set, then expression level of gene was calculated by RSEM (v1.2.12) [56]. The heatmap was drawn by pheatmap (v1.0.8) according to the gene expression in different samples. Essentially, differential expression analysis was performed using the DESeq2 (v1.4.5) with Q value  $\leq 0.05$  [57, 58]. To take insight to the change of phenotype, GO (<http://www.geneontology.org>) and KEGG (<https://www.kegg.jp/>) enrichment analysis of annotated different expressed gene was performed by Phyper ([https://en.wikipedia.org/wiki/Hypergeometric\\_distribution](https://en.wikipedia.org/wiki/Hypergeometric_distribution)) based on Hypergeometric test. The significant levels of terms and pathways were corrected by Q value with a rigorous threshold (Q value  $\leq 0.05$ ) by Bonferroni. Rank Genes by TF were analysed on N.C. Yeasttract *K.phaffii* (<http://yeastract-plus.org/ncyeastract/kphaffii/formrankbytf.php>) [59, 60].

### Metabolome data analysis

In this experiment, the metabolic profile changes of the samples were analysed by a metabonomics method based on UHPLC-Q-TOFMS technology. Metabonomics usually takes strict OPLS-DA VIP > 1 and P value < 0.05 as the screening criteria for significantly different metabolites. Before the annotation and analysis of KEGG pathways, the differential metabolites screened by positive and negative ion patterns were combined. The significant levels of metabolite enrichment in each pathway were analysed and calculated by Fisher's accurate test to determine the metabolic and signal transduction pathways that were significantly affected [61].

### qRT-PCR

After extraction using the hot acidic phenol method, total RNA was treated with RNase-free DNase I (Thermo Scientific) for 42 °C, 5 min according to the

manufacturer's protocol. Subsequently, first-strand cDNA synthesis and qPCR were performed using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) and the PowerUp™ SYBR™ Green Master Mix (Thermo Scientific). For the negative control, the reverse transcriptase was omitted in the cDNA synthesis reaction with the DNase-treated total RNA as the template. No PCR product was obtained using the negative control as the template, suggesting that the genomic DNA was removed completely [53]. Data are presented as the mean of six replicates  $\pm$  standard deviation.

The relative amount of fold change in gene expression was calculated using the comparative Ct method as follows.

$$\text{Log}_2^{-1}[(\text{Ct value of the target gene of the sample to be tested} - \text{Ct value of the internal reference gene of the sample to be tested}) - (\text{Ct value of the target gene of the control sample} - \text{Ct value of the internal reference gene of the control sample})]$$

### Abbreviations

AOX: Alcohol Oxidase; FLD: Formaldehyde Dehydrogenase; FGH: S-Hydroxymethyl Glutathione Hydrolase; FDH: Formate Dehydrogenase; DPCs: DNA-Protein Crosslinks; DAS: Dihydroxyacetone Synthase; ACAS: Acetyl-Coenzyme A Synthetase; ADH: Alcohol Dehydrogenase; ALDH: Aldehyde Dehydrogenase; FBA: Fructose-Bisphosphate Aldolase; CAT: Catalase; GPX: Glutathione Peroxidase; GSR: Glutathione Reductase; COX7C: Cytochrome C Oxidase Subunit 7c; COX10: Haem O Synthase; ROS: Reactive Oxygen Species; GSH: Reduced Glutathione; GSSG: Oxidized Glutathione; metE: 5-Methyltetrahydropteroyltrimethylamino-Homocysteine Methyltransferase; CTH: Cystathionine Gamma-lyase; AOC: Amine Oxidase; amiE: Amidase; PRODH: Proline Dehydrogenase; GCLC: Glutamate-Cysteine Ligase Catalytic Subunit; GST: Glutathione S-transferase; DUG: Cys-Gly Metallodipeptidase; UPS: Ubiquitin-Proteasome System.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-08592-8>.

**Additional file 1: Figure 1.** KEGG clustering analysis of differential metabolites between formaldehyde dehydrogenase knockout strains and wild strains. Each bubble in the bubble plot represents a metabolic pathway (the top 20 most significant were selected based on *P*-value).

**Additional file 2: Figure 2.** Transcription-related DEGs in *P. pastoris*. Red represents significant up-regulation and blue represents significant down-regulation, with shades of colour indicating the degree of up- and down-regulation.

**Additional file 3: Supplemental table 1.** List of DEGs encoding transcription factors in *p. pastoris*.

**Additional file 4: Supplemental table 2.** TF rank of the DEGs in *p.pastoris*.

### Acknowledgements

Not applicable.

### Authors' contributions

YY carried out the whole experiment design, operation and data analysis, and was a major contributor in writing the manuscript. JY conducted experimental operations and data analysis of the transcriptome. FZ provided *Pichia pastoris* CRISPR/Cas9 technology and analyzed some transcriptome data. YL guided the conception and design of the experiment. SH guided the research

throughout the whole process. She analyzed and summarized all experimental contents. All authors read and approved the final manuscript.

### Funding

This report research was supported by the National Key R&D Program of China [2018YFA0901500]. It bears the cost of testing the transcriptome and metabolome and the labor cost of the researchers.

### Availability of data and materials

The RNA sequence datasets generated and/or analysed during the current study are available in the [NCBI Sequence Read Archive (SRA)] repository, [<https://www.ncbi.nlm.nih.gov/biosample/>]. Our recent submission to the BioSample database has been successfully processed and will be released on the date specified. BioSample accessions: SAMN20981851, SAMN20981852, SAMN20981853, SAMN20981854. Temporary Submission ID: SUB10230863. Release date: 2022–09-08, or with the release of linked data, whichever is first. All metabolomics data generated or analysed during this study are included in this published article [and its supplementary information files].

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

Received: 3 September 2021 Accepted: 27 April 2022

Published online: 12 May 2022

### References

- Bernauer L, Radkohl A, Lehmayr LGK, Emmerstorfer-Augustin A. Komagataella phaffii as Emerging Model Organism in Fundamental Research. *Front Microbiol.* 2021;11:607028. <https://doi.org/10.3389/fmicb.2020.607028>.
- Karbalaeei M, Rezaee SA, Farsiani H. *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *J Cell Physiol.* 2020;235(9):5867–81. <https://doi.org/10.1002/jcp.29583>.
- Zepeda AB, Pessoa A, Fariás JG. Carbon metabolism influenced for promoters and temperature used in the heterologous protein production using *Pichia pastoris* yeast. *Braz J Microbiol.* 2018;49:119–27. <https://doi.org/10.1016/j.bjbm.2018.03.010>.
- Yurimoto H, Kato N, Sakai Y. Assimilation, dissimilation, and detoxification of formaldehyde, a central metabolic intermediate of methylotrophic metabolism. *Chem Rec.* 2005;5(6):367–75. <https://doi.org/10.1002/tcr.20056>.
- Johnson MA, Waterham HR, Ksheminska GP, Fayura LR, Cereghino JL, Stasyk OV, Veenhuis M, Kulachkovsky AR, Sibirny AA, Cregg JM. Positive Selection of Novel Peroxisome Biogenesis-Defective Mutants of the Yeast *Pichia pastoris*. *Genetics.* 1999;151(4):1379–91. <https://doi.org/10.1093/genetics/151.4.1379>.
- Vorholt J. Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Arch Microbiol.* 2002;178(4):239–49. <https://doi.org/10.1007/s00203-002-0450-2>.
- Hartner FS, Glieder A. Regulation of methanol utilisation pathway genes in yeasts. *Microb Cell Fact.* 2006;5(1):39. <https://doi.org/10.1186/1475-2859-5-39>.
- Geier M, Brandner C, Strohmeier GA, Hall M, Hartner FS, Glieder A. Engineering *Pichia pastoris* for improved NADH regeneration: A novel chassis strain for whole-cell catalysis. *Beilstein J Org Chem.* 2015;11:1741–8. <https://doi.org/10.3762/bjoc.11.190>.
- Schroer K, Peter Luef K, Stefan Hartner F, Glieder A, Pscheidt B. Engineering the *Pichia pastoris* methanol oxidation pathway for improved NADH regeneration during whole-cell biotransformation. *Metab Eng.* 2010;12(1):8–17. <https://doi.org/10.1016/j.ymben.2009.08.006>.
- Liu T, Liu B, Zhou H, Zhang J. Knockout of the DAS gene increases S-adenosylmethionine production in *Komagataella phaffii*. *Biotechnol Biotechnol Equip.* 2020;35(1):29–36. <https://doi.org/10.1080/13102818.2020.1837012>.
- Çelik E, Çalık P, Oliver SG. Metabolic flux analysis for recombinant protein production by *Pichia pastoris* using dual carbon sources: Effects of methanol feeding rate. *Biotechnol Bioeng.* 2010;105(2):317–29. <https://doi.org/10.1002/bit.22543>.
- Vanz AL, Lünsdorf H, Adnan A, Nimtz M, Gurrnkonda C, Khanna N, Rinas U. Physiological response of *Pichia pastoris* GS115 to methanol-induced high level production of the Hepatitis B surface antigen: catabolic adaptation, stress responses, and autophagic processes. *Microb Cell Fact.* 2012;11(1):103. <https://doi.org/10.1186/1475-2859-11-103>.
- Jordà J, Suarez C, Carnicer M, ten Pierick A, Heijnen J, van Gulik W, Ferrer P, Albiol J, & Wahl A. Glucose-methanol co-utilization in *Pichia pastoris* studied by metabolomics and instantaneous 13C flux analysis. *BMC Syst Biol.* 2013;7(1). <https://doi.org/10.1186/1752-0509-7-17>.
- Vartiainen E, Blomberg P, Ilmén M, Andberg M, Toivari M, & Penttilä M. Evaluation of synthetic formaldehyde and methanol assimilation pathways in *Yarrowia lipolytica*. *Fungal Biol Biotechnol.* 2019;6(1). <https://doi.org/10.1186/s40694-019-0090-9>.
- Guo F, Dai Z, Peng W, Zhang S, Zhou J, Ma J, Dong W, Xin F, Zhang W, Jiang M. Metabolic engineering of *Pichia pastoris* for malic acid production from methanol. *Biotechnol Bioeng.* 2020;118(1):357–71. <https://doi.org/10.1002/bit.27575>.
- Shen S, Sulter G, Jeffries TW, Cregg JM. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene.* 1998;216(1):93–102. [https://doi.org/10.1016/S0378-1119\(98\)00315-1](https://doi.org/10.1016/S0378-1119(98)00315-1).
- Wakayama K, Yamaguchi S, Takeuchi A, Mizumura T, Ozawa S, Tomizuka N, Hayakawa T, Nakagawa T. Regulation of intracellular formaldehyde toxicity during methanol metabolism of the methylotrophic yeast *Pichia methanolica*. *J Biosci Bioeng.* 2016;122(5):545–9. <https://doi.org/10.1016/j.jbiosc.2016.03.022>.
- Yuan W, Zhong X, Han Q, Jiang Y, Shen J, & Wang B. A novel formaldehyde fluorescent probe based on 1, 8-naphthalimide derivative and its application in living cell. *J Photochem Photobiol, A.* 2020;400: 112701. <https://doi.org/10.1016/j.jphotochem.2020.112701>.
- Nakamura J, Holley D, Kawamoto T, & Bultman S. J. (2020). The failure of two major formaldehyde catabolism enzymes (ADH5 and ALDH2) leads to partial synthetic lethality in C57BL/6 mice. *Genes Environ.* 2020;42(1). <https://doi.org/10.1186/s41021-020-00160-4>.
- North M, Gaytán BD, Romero C, De La Rosa VV, Loguinov A, Smith MT, Zhang L, Vulpe CD. Functional Toxicogenomic Profiling Expands Insight into Modulators of Formaldehyde Toxicity in Yeast. *Front Genet.* 2016;7:200. <https://doi.org/10.3389/fgene.2016.00200>.
- Anandarajan V, Noguchi C, Oleksak J, Grothusen G, Terlecky D, Noguchi E. Genetic investigation of formaldehyde-induced DNA damage response in *Schizosaccharomyces pombe*. *Curr Genet.* 2020;66(3):593–605. <https://doi.org/10.1007/s00294-020-01057-z>.
- Gonzalez-Rivera JC, Sherman MW, Wang DS, Chuvalo-Abraham JCL, Hildebrandt Ruiz L, Contreras LM. RNA oxidation in chromatin modification and DNA-damage response following exposure to formaldehyde. *Sci Rep.* 2020;10(1):16545. <https://doi.org/10.1038/s41598-020-73376-7>.
- Chen FY-H, Jung H-W, Tsuei C-Y, Liao J-C. Converting *Escherichia coli* to a Synthetic Methylotroph Growing Solely on Methanol. *Cell.* 2020;182(4):933–946.e14. <https://doi.org/10.1016/j.cell.2020.07.010>.
- Yasokawa D, Murata S, Iwahashi Y, Kitagawa E, Nakagawa R, Hashido T, Iwahashi H. Toxicity of Methanol and Formaldehyde Towards *Saccharomyces cerevisiae* as Assessed by DNA Microarray Analysis. *Appl Biochem Biotechnol.* 2009;160(6):1685–98. <https://doi.org/10.1007/s12010-009-8684-y>.
- Jordà J, Jouhten P, Cámara E, Maaheimo H, Albiol J, & Ferrer P. (2012). Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose:methanol mixtures. *Microb Cell Fact.* 2012;11(1). <https://doi.org/10.1186/1475-2859-11-57>.
- Zhu T, Zhao T, Bankefa OE, Li Y. Engineering unnatural methylotrophic cell factories for methanol-based biomanufacturing: Challenges and opportunities. *Biotechnol Adv.* 2020;39:107467. <https://doi.org/10.1016/j.biotechadv.2019.107467>.

27. De Schutter K, Lin Y-C, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y, Callewaert N. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol*. 2009;27(6):561–6. <https://doi.org/10.1038/nbt.1544>.
28. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res*. 2020;49(D1):D545–51. <https://doi.org/10.1093/nar/gkaa970>.
29. van Zutphen T, Baerends R J, Susanna K A, de Jong A, Kuipers O P, Veenhuis M, van der Klei I. J. . Adaptation of *Hansenula polymorpha* to methanol: a transcriptome analysis. *BMC Genomics*. 2010;11(1). <https://doi.org/10.1186/1471-2164-11-1>.
30. Nakagawa T, Ito T, Fujimura S, Chikui M, Mizumura T, Miyaji T, Yurimoto H, Kato N, Sakai Y, Tomizuka N. Molecular characterization of the glutathione-dependent formaldehyde dehydrogenase gene FLD1 from the methylotrophic yeast *Pichia methanolica*. *Yeast*. 2004;21(5):445–53. <https://doi.org/10.1002/yea.1110>.
31. Sakai Y, Nakagawa T, Shimase M, Kato N. Regulation and Physiological Role of the DAS1 Gene, Encoding Dihydroxyacetone Synthase, in the Methylotrophic Yeast *Candida boidinii*. *J Bacteriol*. 1998;180(22):5885–90. <https://doi.org/10.1128/jb.180.22.5885-5890.1998>.
32. Lu K, Ye W, Gold A, Ball LM, Swenberg JA. Formation of S-[1-(N2-Deoxyguanosinyl)methyl]glutathione between Glutathione and DNA Induced by Formaldehyde. *J Am Chem Soc*. 2009;131(10):3414–5. <https://doi.org/10.1021/ja808048c>.
33. Pietzke M, Burgos-Barragan G, Wit N, Tait-Mulder J, Sumpton D, Mackay G M, Patel K J, Vazquez A. Amino acid dependent formaldehyde metabolism in mammals. *Communications Chem*. 2020;3(1). <https://doi.org/10.1038/s42004-020-0324-z>.
34. Torrellas M, Rozès N, Aranda A, Matallana E. Basal catalase activity and high glutathione levels influence the performance of non-Saccharomyces active dry wine yeasts. *Food Microbiol*. 2020;92:103589. <https://doi.org/10.1016/j.fm.2020.103589>.
35. Yano T, Takigami E, Yurimoto H, Sakai Y. Yap1-Regulated Glutathione Redox System Curtails Accumulation of Formaldehyde and Reactive Oxygen Species in Methanol Metabolism of *Pichia pastoris*. *Eukaryot Cell*. 2009;8(4):540–9. <https://doi.org/10.1128/ec.00007-09>.
36. Yurimoto H, Lee B, Yano T, Sakai Y, Kato N. Physiological role of S-formylglutathione hydrolase in C1 metabolism of the methylotrophic yeast *Candida boidinii*. *Microbiology*. 2003;149(8):1971–9. <https://doi.org/10.1099/mic.0.26320-0>.
37. Skrzydlewska E. Toxicological and Metabolic Consequences of Methanol Poisoning. *Toxicol Mech Methods*. 2003;13(4):277–93. <https://doi.org/10.1080/713857189>.
38. Rußmayer, H., Buchetics, M., Gruber, C., Valli, M., Grillitsch, K., Modarres, G., Guerrasio, R., Klavins, K., Neubauer, S., Drexler, H., Steiger, M., Troyer, C., Al Chalabi, A., Krebiehl, G., Sonntag, D., Zellnig, G., Daum, G., Graf, A. B., Altmann, F., & Koellensperger, G. (2015). Systems-level organization of yeast methylotrophic lifestyle. *BMC Biol*. 2015;13(1). <https://doi.org/10.1186/s12915-015-0186-5>.
39. Stinglele J, Bellelli R, Boulton SJ. Mechanisms of DNA–protein crosslink repair. *Nat Rev Mol Cell Biol*. 2017;18(9):563–73. <https://doi.org/10.1038/nrm.2017.56>.
40. Zhao S, Kieser A, Li H-Y, Reinking HK, Weickert P, Euteneuer S, Yaneva D, Acampora A, Götz M, Feederle R, Stinglele J. A ubiquitin switch controls autocatalytic inactivation of the DNA–protein crosslink repair protease SPRTN. *Nucleic Acids Res*. 2020;49(2):902–15. <https://doi.org/10.1093/nar/gkaa1224>.
41. Marshall RS, Vierstra RD. Dynamic Regulation of the 26S Proteasome: From Synthesis to Degradation. *Front Mol Biosci*. 2019;6:40. <https://doi.org/10.3389/fmolb.2019.00040>.
42. Tanaka K. The proteasome: Overview of structure and functions. *Proceedings of the Japan Academy, Series B*. 2009;85(1):12–36. <https://doi.org/10.2183/pjab.85.12>.
43. Klages-Mundt NL, Li L. Formation and repair of DNA–protein crosslink damage. *Science China Life Sciences*. 2017;60(10):1065–76. <https://doi.org/10.1007/s11427-017-9183-4>.
44. Jayachandran C, Palanisamy Athiyaman B, Sankaranarayanan M. Cofactor engineering improved CALB production in *Pichia pastoris* through heterologous expression of NADH oxidase and adenylate kinase. *PLoS ONE*. 2017;12(7):e0181370. <https://doi.org/10.1371/journal.pone.0181370>.
45. Kern A, Hartner FS, Freigassner M, Spielhofer J, Rumpf C, Leitner L, Fröhlich K-U, Glieder A. *Pichia pastoris* “just in time” alternative respiration. *Microbiol*. 2007;153(4):1250–60. <https://doi.org/10.1099/mic.0.2006/001404-0>.
46. van Zutphen T, Veenhuis M, van der Klei J. Damaged peroxisomes are subject to rapid autophagic degradation in the yeast *Hansenula polymorpha*. *Autophagy*. 2011;7(8):863–72. <https://doi.org/10.4161/aut.7.8.15697>.
47. Zhang C, Ma Y, Miao H, Tang X, Xu B, Wu Q, Mu Y, Huang Z. Transcriptomic Analysis of *Pichia pastoris* (*Komagataella phaffii*) GS115 During Heterologous Protein Production Using a High-Cell-Density Fed-Batch Cultivation Strategy. *Front Microbiol*. 2020;11:463. <https://doi.org/10.3389/fmicb.2020.00463>.
48. Stinglele J, Jentsch S. DNA–protein crosslink repair. *Nat Rev Mol Cell Biol*. 2015;16(8):455–60. <https://doi.org/10.1038/nrm4015>.
49. Mota MN, Martins LC, Sá-Correia I. The Identification of Genetic Determinants of Methanol Tolerance in Yeast Signifies Differences in Methanol and Ethanol Toxicity Mechanisms and Candidates for Improved Methanol Tolerance Engineering. *J Fungi*. 2021;7(2):90. <https://doi.org/10.3390/jof7020090>.
50. Gu Y, Gao J, Cao M, Dong C, Lian J, Huang L, Cai J, Xu Z. Construction of a series of episomal plasmids and their application in the development of an efficient CRISPR/Cas9 system in *Pichia pastoris*. *World J Microbiol Biotechnol*. 2019;35(6):79. <https://doi.org/10.1007/s11274-019-2654-5>.
51. Naito Y, Hino K, Bono H, Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*. 2014;31(7):1120–3. <https://doi.org/10.1093/bioinformatics/btu743>.
52. Liang S, Wang B, Pan L, Ye Y, He M, Han S, Zheng S, Wang X, Lin Y. Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. *BMC Genomics*. 2012;13(1):738. <https://doi.org/10.1186/1471-2164-13-738>.
53. Tyurin OV, Kozlov DG. Deletion of the FLD gene in methylotrophic yeasts *Komagataella phaffii* and *Komagataella kurtzmanii* results in enhanced induction of the AOX1 promoter in response to either methanol or formate. *Microbiol*. 2015;84(3):408–11. <https://doi.org/10.1134/s0026261715030212>.
54. Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. *Bioinformatics*. 2008;24(5):713–4. <https://doi.org/10.1093/bioinformatics/btn025>.
55. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60. <https://doi.org/10.1038/nmeth.3317>.
56. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923>.
57. Li B, Dewey C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12(1). <https://doi.org/10.1186/1471-2105-12-323>.
58. Love M I, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12). <https://doi.org/10.1186/s13059-014-0550-8>.
59. Monteiro PT, Oliveira J, Pais P, Antunes M, Palma M, Cavalheiro M, Galocha M, Godinho CP, Martins LC, Bourbon N, Mota MN, Ribeiro RA, Viana R, Sá-Correia I, Teixeira MC. YEASTRACT+: a portal for cross-species comparative genomics of transcription regulation in yeasts. *Nucleic Acids Res*. 2019;48(D1):D642–9. <https://doi.org/10.1093/nar/gkz859>.
60. Godinho, C. P., Palma, M., Oliveira, J., Mota, M. N., Antunes, M., Teixeira, M. C., Monteiro, P. T., & Sá-Correia, I. (2021). The N.C.Yeasttract and CommunityYeasttract databases to study gene and genomic transcription regulation in non-conventional yeasts. *FEMS Yeast Research*, 21(6). <https://doi.org/10.1093/femsyr/foab045>.
61. Shen, X., Gong, X., Cai, Y., Guo, Y., Tu, J., Li, H., Zhang, T., Wang, J., Xue, F., & Zhu, Z.-J. Normalization and integration of large-scale metabolomics data using support vector regression. *Metabolomics*. 2016;12(5). <https://doi.org/10.1007/s11306-016-1026-5>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.