Transcriptional regulatory networks of methanol-independent protein expression in *Pichia pastoris* under the *AOX1* promoter with trans-acting elements engineering

Lei Shi¹, Jinjia Wang¹, Xiaolong Wang¹, Yuanxing Zhang¹,², Zhiwei Song³,⁴, Menghao Cai¹ and Xiangshan Zhou¹*

**Abstract**

To explore the differences in the intracellular transcriptional mechanism in carbon-derepressed and wild-type *Pichia pastoris* strains fed with three different carbon sources. RNA in carbon-derepressed (Δmig1Δmig2Δnrg1-Mit1; Mut) and wild-type (WT) *P. pastoris* fed with three different carbon sources (dextrose, glycerol, and methanol) were sequenced. Differentially expressed genes (DEGs) associated with these carbon sources were obtained and clustered into modules using weighted gene co-expression network analysis (WGCNA). Signaling pathway enrichment analysis was performed using KEGG, and protein to protein interaction (PPI) network was also constructed. A total of 2536 DEGs were obtained from three intersections, and some of them were enriched in carbon sources and involved in carbon metabolism, secondary metabolisms, and amino acid biosynthesis. Two modules, MEgreenyellow (involved in protease, oxidative phosphorylation, endoplasmic reticulum protein processing, folate carbon pool, and glycerol phospholipid metabolism pathways) and MEmidnightblue (involved in protease, endocytosis, steroid biosynthesis, and hippo signaling pathways) were significantly correlated with the strain type. Eight hub genes and two sub-networks were obtained from PPI network. Sub-network A enriched in proteasomes pathway while sub-network B enriched in ribosome pathway. The genes involved in carbon metabolism, secondary metabolic, and amino acid biosynthesis pathways changed significantly under different carbon sources. The changes in proteasome and ribosome activities play roles in carbohydrate metabolism in the methanol-free *P.AOX1* start-up Mut strain.

**Keywords:** *Pichia pastoris*, RNA-seq, Carbon sources, AOX1 promoter

**Introduction**

*Pichia pastoris* (*P. pastoris*) is a widely used recombinant expression system for exogenous genes having prokaryotic advantages of having a simple operation and rapid growth, and also has eukaryotic cell functions of post-translational modifications of proteins. Yeast expression systems include *Saccharomyces cerevisiae* system, methylotrophic yeast expression system, and merozoite yeast expression system (van den Hazel et al. 1996). The methylotrophic yeast *P. pastoris* expression system is one of the most commonly used expression systems.

The most commonly used promoter on *P. pastoris* expression system is the alcohol oxidase 1 promoter (*AOX1*), which is a carbon source-dependent promoter whose transcription expression is strongly induced by methanol, but inhibited by other carbon sources, such as glucose, glycerol, or ethanol (Liang et al. 2014). *P. pastoris* synthesizes a large number of enzymes when using methanol as the sole carbon source. However, the
use of methanol has many disadvantages as it is toxic, flammable, and volatile, which makes methanol harmful to the human body (Polupanov et al. 2012; Portnoy et al. 2011). Methanol fermentation requires large amounts of oxygen, around three to four times higher than the amount needed when the carbon source is glucose (Lin et al. 2000). It also produces hydrogen peroxide ($H_2O_2$) during methanol metabolism, causing oxidative stress and proteolytic degradation of some proteins (Wang et al. 2010).

Researchers have worked on modifying the expression system of $P.\ pastoris$ by various methods which avoid the use of methanol as a carbon source (Lu et al. 2007). The most convenient method to avoid methanol dependence was to engineer other highly expressed promoters that can replace $AOX1$ and work without methanol induction (Chen et al. 2007). $AOX1$ promoter has several regulatory factors, such as Mig1, Mig2, Nrg1, Mit1, and Prm1. Among these factors, Mig1, Mig2, and Nrg1 inhibit $AOX1$ promoter in cells grown in glucose and glycerol, while Mit1 and Prm1 activate $AOX1$ promoters in cells grown in methanol (Prielhofer et al. 2013). Nevertheless, the regulatory mechanism of $AOX1$ promoter in $P.\ pastoris$ remains unclear.

An earlier study (Lin-Cereghino et al. 2006) carried out a large-scale mutation screening analysis to identify regulatory factors and to elucidate their mechanisms and the studies also not recorded how to construct an engineered strain or to verify the success of the strain construction. To understand the effects of transcriptional repressors on $AOX1$ and the transcriptional differences between carbon-derepressed strain (Mut) and wild-type (WT) strain, we modified expression levels of carbon source repressors in $P.\ pastoris$ (Wang et al. 2017). Mig1, Mig2, and Nrg1 which act as carbon source repressors were inhibited and Mit1 levels were enhanced in the carbon-derepressed strain (Wang et al. 2017). In the present study, we analyzed the transcriptomes of WT and the carbon-derepressed strain $\Delta$mig1$\Delta$mig2$\Delta$nrg1$\Delta$mit1 (Mut) of $P.\ pastoris$ under different carbon sources using RNA-seq method. Differential expression genes (DEGs) analysis was applied on different carbon sources with Mut and WT strains. Weighted gene co-expression network analysis (WGCNA) and Kyoto encyclopedia of genes and genomes (KEGG) enrichment were used to explore the effect of different carbon sources on the efficiency of gene expression. Protein–protein interaction (PPI) network analysis was used to obtain key sub-networks and hub genes with Mut and WT strains. The goal of the study was to investigate the mechanisms underlying the carbon derepression and $AOX1$ promoter regulation of the methanol-free $P_{AOX1}$ start-up $\Delta$mig1$\Delta$mig2$\Delta$nrg1$\Delta$mit1 strain.

Materials and methods

$P.\ pastoris$ strains and culture conditions

$P.\ pastoris$ wild-type (WT) strain GS115 (Life Technologies, Carlsbad, CA, USA) and the genetically engineered $P.\ pastoris$ strain $\Delta$mig1$\Delta$mig2$\Delta$nrg1$\Delta$mit1 (Mut) (constructed and preserved in our laboratory) were used for this study. A chemostatic cultivation medium for recombinant $P.\ pastoris$ was used to provide a stable environment to ensure the reliability and repeatability of RNA-seq analysis. In brief, Mut and WT strains (stored at $-80^\circ C$) were inoculated directly into 250-mL shaking flasks containing 50 mL yeast nitrogen-base/dextrose (YNB) medium (0.67% yeast nitrogen base (YNB) and 1% glucose) or into 50 mL minimal glycerol media (MGY) as the primary seed culture till $OD_{600}=2–6$, respectively, and cultivated at 30 °C for 24 h with agitation (200 rpm). Then, Mut and WT seed cultures were used to inoculate into 500 mL YND or MGY medium (0.67% YNB+1% glycerin) till $OD_{600}=2–6$, respectively, as the second seed cultures and the cultures were then cultivated for 16 h at 30 °C (200 rpm). Batch cultivation was started by inoculating the secondary seeds at a concentration of 1:10 (v/v) into a fermenter containing basal salts media (BSM). Batch fermentations were carried out at 30 °C, pH 5.0 (controlled by 25% ammonium hydroxide), and a minimum of 30% dissolved oxygen concentration. When carbon source in the medium was depleted and the concentration of dissolved oxygen increased sharply, the chemostat medium using glucose (20.0 g L$^{-1}$), glycerol (20.0 g L$^{-1}$), or methanol (20.0 g L$^{-1}$) as the sole carbon source was supplemented. The fermentation broth is drained by a peristaltic pump to maintain a constant volume of the fermentation broth and the flowrate of glucose, glycerol, or methanol. Finally, they were inoculated into a 50 mL liquid MGY medium with an initial concentration of $OD_{600}=1$, and two parallel samples were inoculated for each strain. Samples were taken when the specific growth rate reached $\mu=0.03$ h$^{-1}$.

RNA extraction, construction of cRNA Library and High-throughput sequencing

The detection of RNA samples was provided by Beijing Novosource Bioinformation Technology Co. RNA concentration was measured using Qubit® RNA Assay Kit with Qubit® 2.0 Fluorometer (Life Technologies, CA, USA), and RNA integrity was assessed with the RNA Nano 6000 Assay Kit by the Agilent 2100 Bioanalyzer (Agilent Technologies, USA), and only the qualified samples were used for library construction and sequencing experiments. As shown in Table 1, a total of 12 samples were used for RNA-seq analysis, including WT strains in glucose, glycerol, and methanol conditions (named as $W_D$, $W_G$, and $W_M$) and Mut strain in the same three
conditions (named as M_D, M_G, and M_M). Each sample was done in duplicates. Illumina Hiseq2000 (Illumina, NY, USA) high-throughput sequencing system was used to sequence RNA of the samples.

Data pre-processing
Quality control of raw reads was processed by removing reads containing adapter sequences and reads with N (information cannot be determined) bases greater than 10%. Consequently, clean reads were obtained for downstream analysis. Alignment of the paired-end clean reads to the coding sequence (CDS) of GS115 in The European Bioinformatics Institute (EMBL) database was performed using TopHat v2 (Kim et al. 2013). Based on the predicted gene model of Cufflinks (Trapnell et al. 2012), the alternative splicing event on each sample was classified and counted by Asprofile software (Florea et al. 2013). All the data obtained from reads were assembled together with Cufflinks and then compared with known gene models by Cuffcompare software to predict new transcripts. Gene expression levels were quantified as the number of reads per kilobase of exon region per million reads (RPKM).

DEG analysis
DEGs were analyzed using the R package DEseq (Anders and Huber 2010; Anders and Huber 2012). For the samples with biological repetition, the original count was corrected by DEseq and the log of the count ratio between the two groups was calculated. Finally, log2FoldChange (FC) value was obtained. Benjamini & Hochberg’s method (Benjamini and Hochberg 1995) was used for multiple testing correction and genes with an adjusted \( P \)-value < 0.05 was selected as the cut-off criterion for the identification of DEGs.

DEGs were analyzed for growth in different sources, including W_G vs. W_D, M_G vs. M_D, W_M vs. W_D, M_M vs. M_D, W_M vs. W_G, and M_M vs. M_G and a total of 6 groups of DEGs were analyzed. Then, we analyzed the DEGs intersection between Mut and WT strains in the same carbon comparison, such as the intersection DEGs between W_G vs. W_D and M_G vs. M_D. Then, a total of three groups of intersection DEGs were determined. Finally, we combined these three groups of intersection DEGs together, which were thought to be genes related to carbon sources.

WGCNA network construction
Using weighted gene co-expression analysis (WGCNA) R package (Langfelder and Horvath 2008), a co-expressed network of DEGs related to their carbon sources was constructed, and further screening of the gene sets which were significantly related to their strain type, and genes which were considered to be both related to the carbon source and the strain type were selected. Two methods were used to find modules related to the strain type. First, we calculated the correlation between each module feature vector gene and strain types. Second, the mean value of the correlation between the trait and gene expression in each module was taken as the significance. The greater the significance, the more the correlation between the module and the trait. By selecting the weight value of power, the whole network was distributed as a scale-free network, and the power with structure in \( R^2 \) reached 0.8 would be generally selected. Because of batch-related effects, sample heterogeneity, and experimental conditions used, we selected power \( = 9 \) and then, based on clustering and dynamic pruning of the data (Barabási 2009), we clustered the highly correlated genes into modules. Finally, gene sets were identified significantly related to the type of strain.

Functional enrichment analysis
KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/anno_iden.php) (Xie 2011; Wu et al. 2006) software was used to analyze the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa and Goto 2000) with DEGs related to carbon sources. Then, we also analyzed the KEGG database for genes obtained from the modules of WGCNA analysis with a cut-off of \( P \) value < 0.05 and gene count \( \geq 2 \), and these pathways were significantly related to the strain type and carbon source.

Construction of protein to protein (PPI) networks
Search tool for the retrieval of interacting genes (STRING v 10.0) (Szklarczyk 2014) was used to explore interactions between proteins. Komagataella pastoris was selected as the reference species to predict whether there were interactions between matched protein sets. Parameter PPI score was set to 0.7 (high confidence), which requires the interacting protein nodes to be included in the above module genes. Then, the PPI network was

| Table 1 The P. pastoris samples used for sequencing |
|-----------------------------------------------|
| Carbon sources | Dextrose (D) | Glycerol (G) | Methanol (M) |
| Mut (M) | |
| Duplicate 1 | M_D_1 | M_G_1 | M_M_1 |
| Duplicate 2 | M_D_2 | M_G_2 | M_M_2 |
| Wild-type (W) | |
| Duplicate 1 | W_D_1 | W_G_1 | W_M_1 |
| Duplicate 2 | W_D_2 | W_G_2 | W_M_2 |

Mut indicates the genetically engineered P. pastoris strain Δmig1Δmig2Δnrg1-Met1
built by Cytoscape software (version 3.2.0) (Shannon 2003) and cytoNCA plug-in (Tang et al. 2015) (v.2.1.6, http://apps.cytoscape.org/apps/cytonca) was used to analyze the topological properties of the node network. The results included degree centrality (DC), Bweenness Centrality (BC), and closeness centrality (CC), and the hub proteins were obtained by ranking the network topological properties of each node. We used MCODE (Degree Cutoff = 2, Node Score Cutoff = 0.2, K-core = 2, MaxDepth = 100) (Bandettini et al. 2012) in Cytoscape to analyze the subnetworks in the PPI network. The modules with score > 5 were chosen to perform the KEGG pathway analysis using KOBAS 3.0.

**Results**

**DEGs analysis**

After data pre-processing, a total of 5381 genes in the gene matrix were obtained from 12 samples. According to the cut-off criterion used, a pairwise comparison was used to compare the three different carbon sources (Fig. 1a). There were 399 overlapping DEGs between W_G vs. W_D and M_G vs. M_D. There were 1863 overlapping DEGs between W_M vs. W_D and M_M vs. M_D. In addition, there were 2098 overlapping DEGs between W_M vs. W_G, and M_M vs. M_G. By combining the union DEGs of the Mut strain and WT strain grown under different carbon sources, 2536 DEGs were detected as genes related to carbon sources. Bidirectional hierarchical clustering was carried out to analyze these 2536 DEGs (Fig. 1b). The results of bidirectional hierarchical clustering showed that DEG analysis only separated methanol samples from the three carbon sources, while glucose and glycerol samples were aggregated in the other branches. However, DEGs in glucose and glycerol samples were separated according to the Mut or WT types. Thus, we inferred that these genes were not only associated with carbon source treatment, but also potentially related to the strain type, laying the foundation for our next co-expression analysis.
The KEGG enrichment analysis results of the 2536 DEGs showed that with the change of carbon source, the genes involved in carbon metabolism, secondary metabolic pathway, and amino acid biosynthesis changed significantly (Table 2).

Screening of related modules and genes by WGCNA
To further screen genes related to strain type, we performed WGCNA analysis for the 2536 DEGs (Fig. 2). Based on the clustering and dynamic pruning method, 2536 DEGs were divided into 20 modules (the grey module was a collection of genes that cannot be aggregated into other modules) (Fig. 2a). Modules with heterogeneity coefficient < 0.2 were merged together, and finally 10 modules were obtained (Fig. 2b, c). By calculating the correlation coefficient between the module eigengene (ME) and sample traits, 4 modules (green-yellow: 122 genes, midnight-blue: 266 genes, grey60: 44 genes, light green: 41 genes) were significantly correlated with the strain type ($p < 0.05$) (Fig. 3a). Both ME_{green-yellow} and ME_{midnight-blue} modules were significantly correlated with the type of strain (correlation coefficient > 0.9).

Functional enrichment pathway analysis
The genes in the ME_{green-yellow} and ME_{midnight-blue} modules were analyzed by KEGG pathway enrichment, and the results are shown in Fig. 3b. The genes in the ME_{green-yellow} module were mainly enriched in proteases, oxidative phosphorylation, endoplasmic reticulum protein processing, folate carbon pool, and glycerol phospholipid metabolism pathway, while the genes in the ME_{midnight-blue} module were significantly correlated with the type of strain (correlation coefficient > 0.9). Genes in the ME_{midnight-blue} module were down-regulated in Mut vs. WT, and genes in the ME_{green-yellow} module were up-regulated in Mut vs. WT (Fig. 3c–d).

Table 2 Five KEGG pathways related to carbon source in \textit{P. pastoris}

| Term                      | Gene number | P-value        | Corrected P-value |
|---------------------------|-------------|----------------|------------------|
| Metabolic pathways        | 395         | 3.20E−06       | 0.000329689      |
| Biosynthesis of secondary metabolites | 179         | 7.32E−05       | 0.003657969      |
| Biosynthesis of antibiotics | 140        | 0.000106543    | 0.003657969      |
| Biosynthesis of amino acids | 83          | 0.000570221    | 0.014683197      |
| Carbon metabolism         | 73          | 0.001622285    | 0.033419065      |

Discussion
The commonly used and tightly regulated promoter used to drive heterologous protein expression in \textit{P. pastoris} is the AOX1 promoter; the gene expression of AOX1 is repressed in carbon sources (Brierley et al. 1990). A \textit{P. pastoris} mutant strain expressing AOX1-lac2 fusion has been demonstrated to be able to use sorbitol, alanine, trehalose, or mannitol as the sole carbon source (Inan and Meagher 2001). The deletion of the hexose transporter PpHXT1, leading to glucose or fructose-induced expression of the AOX1 gene and pexophagy, is associated with catabolite repression of Aox in \textit{P. pastoris} (Zhang et al. 2010). We previously described a novel Mut strain of the AOX1 promoter-based \textit{P. pastoris} expression system, Δmig1Δmig2Δnrg1-Mit1 (Wang et al. 2017). However, the mechanism underlying the catabolite and AOX1 regulation of this expression system should be explored.
In this study, we compared the gene expression patterns of the Mut strain cultured in three carbon sources (glucose, glycerol, and methanol). A total of 2536 carbon source-associated DEGs were detected from the intersection of three groups, and 388 DEGs were found to be associated with both the carbon source type and strain type (Mut or WT). Further investigation of these DEGs indicated the significant involvement of the proteasome and ribosome pathways. Moreover, several hub genes related to AOX1 regulation and carbon sources were identified.

When the two P. pastoris strains (WT and Mut) were induced by different carbon sources, the DEGs were mainly associated with carbon metabolism, secondary metabolic biosynthesis, and amino acid biosynthesis. Carbon metabolism in microbial cells mainly involves glycolysis, the pentose phosphate pathway, and the three
carboxylic acid (TCA) cycle (Zhang et al. 2012). Both the TCA and glyoxylic acid cycles provide energy and precursors for microorganisms (Trapnell et al. 2014). Wang and Yan (2017) showed that when P. pastoris was cultured with glycerol and methanol. Genes associated with the first stage of glycolysis were expressed at the basal level. Dihydroxy acetone phosphate and dihydroxy acetone phosphate were the methanol metabolites in the peroxisome. These metabolites were then transported through the peroxisomal membrane to the cytoplasm and entered the second stage of glycolysis. In a previous study, the growth status of yeast with glucose and methanol as carbon sources was investigated. The metabolic flow of TCA was found to have decreased by three times under methanol growth conditions, while no significant change in transcription or protein levels in the TCA cycle was observed (Russmayer et al. 2015). These results indicate that the TCA cycle may supply different precursor
substances for other metabolic pathways when cells are grown under different carbon sources.

The glycerol transporter GT1 may participate in the process of P\textsubscript{AOX1} inhibition of \textit{P. pastoris} in a medium containing glycerol, and knockdown of this gene can eliminate the glycerol repression of P\textsubscript{AOX1} (Zhan et al. 2016). In this study, we found several hub proteins in the PPI network of the 388 genes in two significant modules, including XP\_002489693.1 (one of six ATPases of the 19S regulatory particle of the 26S proteasome involved in degradation), XP\_002490020.1 (ubiquitin-protein ligase (E3) that interacts with Rpt4p and Rpt6p), XP\_002491845.1 (metalloprotease subunit of the 19S regulatory particle of the 26S proteasome lid), XP\_002492676.1 (Alpha 4 subunit of the 20S proteasome), XP\_002493471.1 (Alpha 1 subunit of the 20S core complex of the 26S proteasome), XP\_002493717.1 (one of six ATPases of the 19S regulatory particle of the 26S proteasome), XP\_002493779.1 (Adenylate cyclase, required for cAMP production and cAMP-dependent protein kinase signaling), and XP\_002494220.1 (Beta 6 subunit of the 20S proteasome). These proteins were mainly associated with the proteasome and ribosome; one of these proteins was ubiquitin ligases E3, which was a proteolytic enzyme guiding relative protein degradation in the proteasome. The proteasome, which has many catalytic functions and can selectively degrade proteins in cells, is a polyclonal macromolecule complex widely distributed in organisms (Tomko and Hochstrasser 2011). To remove the proteins that the cell does not need for functioning as soon as possible, eukaryotic cells can increase the expression of the proteasome by upregulated transcription of genes involved in the proteasome. Ubiquitin–proteasome-linked protein degradation of the key gluconeogenic enzyme, fructose-1,6-bisphosphatase, is associated with glucose uptake and carbohydrate metabolism of \textit{P. pastoris} (Santt et al. 2008). Ribosomes are the synthesis machine of proteins, and they are also an important site for cells to perceive the nutrition level within cells. Studies (Beilharz and Preiss 2004; Prielhofer et al. 2015) reported that the proportion of polyribosomes in total ribosomes (40s, 60s, 80s/monosomes and polysomes) increased by approximately twofold, as detected by RNA-seq sequencing, when the different carbon sources induced the expression of proteins in \textit{P. pastoris}. When methanol was used as the carbon source, high levels of protein expression were observed. At this time, a high number of amino acids need to be

![Fig. 4 Carbon metabolism of M_D vs. W_D, M_G vs. W_D (a) and M_M vs. M_D, M_M vs. M_G (b) in P. pastoris. Red words represented the up-regulated genes, and blue words the down-regulated genes](image-url)
synthesized and more ribosomes need to participate in protein synthesis (Prielhofer et al. 2015). Therefore, we suggested that the changes in ribosomal and proteasomal activities were associated with carbohydrate metabolism and glucose-mediated as well as glycerol-mediated induction of PAOX1 in P. pastoris, indicating that the carbon derepressed expression system actually causes protein degradation and amino acid synthesis to meet the demand for heterologous protein secretion. However, the signal transduction mechanism that induces the derepression of AOX1 in ∆mig1∆mig2∆nrg1-Mit1 remains to be investigated further.

In the carbon metabolism network, several genes were up-regulated in M_G and M_D compared with W_G and W_D, such as DAS1 and FLD1. DAS1 encodes dihydroxyacetone synthase Das and FLD1 encodes formaldehyde dehydrogenase Fld. They are all critical enzymes in the methanol utilization pathway. Both P_{DAS1} and P_{FLD1}
are regulated by methanol (Vogl and Glieder 2013). In the present study, the higher expression level of DASL and FLD1 may play a critical role in carbon metabolism of the dextrose and glycerol medium.

TKT plays an important role in the pentose phosphate acid cycle and the photosynthetic reduced pentose phosphate cycle. In the group M_M, the expression level of TKT was down-regulated compared with the group M_G and M_D. This result may indicate that in the case of methanol as the sole carbon source, biotransformation efficiency of the carbon source may be lower in M_M than M_G and M_D. Further results require experimentation to verify.

**Conclusion**

In conclusion, with the change of the carbon source, the genes involved in the carbon metabolism pathway, secondary metabolic pathway, and amino acid biosynthesis pathway changed significantly in *P. pastoris*. Changes in the proteasome and ribosome activities played a significant role in the carbohydrate metabolism of the methanol-free P_AOX1 start-up *Δmig1Δmig2Δnrg1-Mit1* strain. Such an expression system could be an efficient alternative to conventional production methods using methanol as the carbon source.

**Supplementary information**

Supplementary information accompanies this paper at https://doi.org/10.1186/s40643-020-00306-w.

**Additional file 1:** Table S1. Proteins in the PPI network.

**Abbreviations**

- WT: Wildtype
- DEGs: Differentially expressed genes
- PPI: Protein–protein interaction
- WGCNA: Weighted gene co-expression network analysis
- KEGG: Kyoto encyclopedia of genes and genomes
- CDS: Coding sequence
- RPKM: Reads per kilo bases per million reads
- FC: Fold-change
- DC: Degree centrality
- BC: Betweenness centrality
- CC: Closeness centrality
- TCA: Three carboxylic acid

### Table 3

**Top 30 genes with higher topological properties in the protein–protein interaction network**

| Gene          | Degree | Gene          | Betweenness | Gene          | Closeness |
|---------------|--------|---------------|-------------|---------------|-----------|
| XP_002493717.1| 28     | XP_002490020.1| 6223.7      | XP_002493717.1| 0.02147   |
| XP_002489693.1| 27     | XP_002493779.1| 6097.7      | XP_002490020.1| 0.02142   |
| XP_002491845.1| 26     | XP_002490773.1| 5526.0      | XP_002489693.1| 0.02136   |
| XP_002493471.1| 26     | XP_002493873.1| 5250.0      | XP_002493873.1| 0.02136   |
| XP_002492676.1| 26     | XP_002493717.1| 4834.5      | XP_002491845.1| 0.02136   |
| XP_002494220.1| 26     | XP_002493265.1| 4645.0      | XP_002493471.1| 0.02132   |
| XP_002494177.1| 26     | XP_002491363.1| 1818.0      | XP_002492676.1| 0.02132   |
| XP_002493117.1| 26     | XP_002489675.1| 1454.0      | XP_002493779.1| 0.02131   |
| XP_002493814.1| 26     | XP_002490815.1| 1346.0      | XP_002490773.1| 0.02124   |
| XP_002491581.1| 25     | XP_002492039.1| 1050.0      | XP_002494220.1| 0.02122   |
| XP_002493329.1| 25     | XP_002489799.1| 864.0       | XP_002494177.1| 0.02122   |
| XP_002492434.1| 25     | XP_002489693.1| 802.5       | XP_002493117.1| 0.02122   |
| XP_002489540.1| 25     | XP_002489961.1| 763.7       | XP_002493814.1| 0.02122   |
| XP_002492684.1| 25     | XP_002491845.1| 753.8       | XP_002493930.1| 0.02122   |
| XP_002494172.1| 25     | XP_002493614.1| 646.0       | XP_002492940.1| 0.02122   |
| XP_002493375.1| 25     | XP_002492074.1| 618.0       | XP_002491581.1| 0.02122   |
| XP_002494106.1| 25     | XP_002493040.1| 618.0       | XP_002493291.1| 0.02122   |
| XP_002491076.1| 25     | XP_002493471.1| 577.6       | XP_002492434.1| 0.02122   |
| XP_002492927.1| 25     | XP_002492676.1| 577.6       | XP_002489540.1| 0.02122   |
| XP_002492977.1| 25     | XP_002493501.1| 500.0       | XP_002492684.1| 0.02122   |
| XP_002493941.1| 25     | XP_002490025.1| 437.7       | XP_002494172.1| 0.02122   |
| XP_002493111.1| 24     | XP_002492667.1| 434.0       | XP_002493375.1| 0.02122   |
| XP_002489719.1| 24     | XP_002491242.1| 432.0       | XP_002494106.1| 0.02122   |
| XP_002493134.1| 24     | XP_002491269.1| 432.0       | XP_002491076.1| 0.02122   |
| XP_002493007.1| 22     | XP_002493930.1| 277.3       | XP_002492927.1| 0.02122   |
| XP_002493779.1| 19     | XP_002492193.1| 249.0       | XP_002492977.1| 0.02122   |
| XP_002491334.1| 15     | XP_002491431.1| 227.0       | XP_002493941.1| 0.02122   |
| XP_002491293.1| 15     | XP_002494220.1| 219.6       | XP_002493311.1| 0.02122   |
| XP_002493265.1| 15     | XP_002489862.1| 218.0       | XP_002489719.1| 0.02122   |
| XP_002490020.1| 10     | XP_002492146.1| 218.0       | XP_002493134.1| 0.02122   |
Acknowledgements
Not applicable.

Authors' contributions
LS and MC contributed to the conception and design of the research, JW, XW, and YZ acquired data, KZ and ZS performed all the statistical analysis, LS drafted the manuscript, and MC revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Funding
None.

Availability of data and materials
All datasets used and analyzed in the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors agree to publish this manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
1 State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China. 2 Shanghai Collaborative Innovation Center for Biomanufacturing, Shanghai 200237, China. 3 Bioprocessing Technology Institute, Singapore, Singapore. 4 Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore.

Received: 15 November 2019 Accepted: 16 March 2020
Published online: 25 March 2020

References
Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol
Anders S, Huber W (2012) Differential expression of RNA-Seq data at the gene level-the DESeq package. Genome Biol
Bandettini WP et al (2012) MultiContrast Delayed Enhanced Movement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. J Cardiovasc Magn Reson 14:83
Barabási AL (2009) Scale-free networks: a decade and beyond. Science 325(5939):412–413
Beilharz TH, Preiss T (2004) Translational profiling: the genome-wide measure of the nascent proteome. Brief Funct Genomic Proteomic 3(2):103–111
Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57(1):289–300
Brierley RA, Russineau C, Kosson R, Metton A, Siegel RS (1990) Fermentation development of recombinant Pichia pastoris expressing the heterologous gene: bovine lysozyme. Ann N Y Acad Sci 589:350–362
Chen HX, Chu J, Zhang SL et al (2007) Intracellular expression of Vitreoscilla hemoglobin improves S-adenosylmethionine production in a recombinant Pichia pastoris. Appl Microbiol Biotechnol 74(6):1205–1212
Florea L, Song L, Salzberg SL (2013) Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. F1000 Res 2:188
Inan M, Meagher MM (2001) Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of Pichia pastoris. J Biosci Bioeng 92(6):585–589
Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27–30
Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14:R36
Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9(1):559
Liang S, Wang B, Pan L, Ye Y, He M, Han S et al (2014) Comprehensive structural annotation of Pichia pastoris transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. BMC Genomics 9(4):511–525
Lin J, Panigraphy D, Trinh L, Folkman J, Shiioa P (2006) Production process for recombinant human angiotatin in Pichia pastoris. J Ind Microbiol Biotechnol 32(1):31–35
Lin-Cereghino GP, Godfrey L, de la Cruz BJ, Johnson S, Khuospherei S, Tolstorukov I et al (2006) Mxtr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in Pichia pastoris. Mol Cell Biol 26(3):883–897
Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol 25(1):117–124
Polupanov AS, Nazarko VY, Sibirra AA (2012) Gs1 protein of the methylotrophic yeast Pichia pastoris is involved in glucose sensing, pexophagy and catabolite repression. Int J Biochem Cell Biol 44(11):1906–1918
Portnoy T, Margot A, Linke R, Atanasova L, Fekete E, Sander E et al (2011) The CRE1 carbon catabolite repressor of the fungus Trichoderma reesei: a master regulator of carbon assimilation. BMC Genomics 12:269
Priehofer R, Maurer M, Klein J, Wenger J, Kiziac K, Gasser B, Mattanovich D (2013) Induction without methanol: novel regulated promoters enable high-level expression in Pichia pastoris. Microb Cell Fact 12(1):5
Priehofer R, Cartwright SP, Graf AB, Valti M, Bill RM, Mattanovich D, Gasser B (2015) Pichia pastoris regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics 16:167
Russeymeyer H, Buchetcs M, Gruber C, Valti M, Grilli K, Modarres G, Guerrios R, Klaivos K, Neubauer S, Dreksler H, Steiger M, Troyer C, Al Chalabi A, Krebieg C, Sonntag D, Zellieg C, Daua M, Graf AB, Altrmann F, Koffensberger G, Hann S, Sauer M, Mattanovich D, Gasser B (2015) Systems-level organization of yeast methylotrophic lifestyle. BMC Syst Biol 13:80
Santt O, Pfirrmann T, Braun B, Jureschke J, Kimmig P, Scheel H, Hofmann K, Thumm M, Wolf DH (2008) The yeast GID complex, a novel ubiquitin ligase (E3) involved in the regulation of carbohydrate metabolism. Mol Biol Cell 19(8):3323–3333
Shannon P et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504
Sklarzczyk D, et al (2014) STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res.1003
Tang Y et al (2015) CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks. BioSystems 127:67–72
Tomko RJ Jr, Hochstrasser M (2011) Order of the proteasomal ATPases and eukaryotic proteasome assembly. Cell Biochem Biophys 60(1–2):13–20
Trapnell C, Roberts A, Goff L, Pertege A, Kim D, Kelley DR et al (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562–578
Trapnell CR, Roberts A, Goff L, Pertege A, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2014) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 9(10):2513
van den Hazel HB, Kieland-Brandt MC, Winther JR (1996) Biosynthesis and function of yeast vacuolar proteases: review. Yeast 12:1–66
Vogt T, Gleder A (2013) Regulation of Pichia pastoris promoters and its consequences for protein production. New Biotechnol 30(4):385–404
Wang XY, Yan X (2017) The metabolic difference and mechanism of Kama gataella phaffi in glycerol and methanol culture based on transcriptional and translational analysis. Acta Microbiol Sin:1–17
Wang Z, Wang Y, Zhang D, Li J, Hua Z, Du G, Chen J (2010) Enhancement of methylotrophic lifestyle. BMC Syst Biol 13:80
Wang Z, Wang Y, Zhang D, Li J, Hua Z, Du G, Chen J (2010) Enhancement of cell viability and alkaline polyglycustarinate lyase production by sorbo tol co-feeding with methanol in Pichia pastoris fermentation. Bioreactors 101(4):1318–1323
Wang J, Wang X, Shi L, Qi F, Zhang P, Zhang Y, Zhou X, Song Z, Cai M (2017) Methanol-independent protein expression by AOX1 promoter with trans-acting elements engineering and glucose–glycerol-shift induction in Pichia pastoris. Sci Rep 7:41850. https://doi.org/10.1038/srep41850
Wu J, Mao X, Cai T, Luo J, Wei L (2006) KOBAS server: a web-based platform for automated annotation and pathway identification. Nucleic Acids Res 34:720–724
Xie CEA (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res 39:316–322
Zhan C, Wang S, Sun Y, Dai X, Liu X, Harvey L, McNeil B, Yang Y, Bai Z (2016) The Pichia pastoris transmembrane protein GT1 is a glycerol transporter and relieves the repression of glycerol on AOX1 expression. FEMS Yeast Res. https://doi.org/10.1093/femsyr/fow033
Zhang P, Zhang W, Zhou X, Bai P, Cregg JM, Zhang Y (2010) Catabolite repression of Aox in Pichia pastoris is dependent on hexose transporter PpHxt1 and pexophagy. Appl Environ Microbiol 76(18):6108
Zhang CL, Qin ZJ, Wang GZ, Ji ZB, Wang JM (2012) Transcriptome and RNA-Seq technology. Biotechnol Bull 12:51–56

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