

Complementary analyses of transcriptome and proteome revealed the formation mechanism of ethyl acetate, ethanol and organic acids in *Kluyveromyces marxianus* L1-1 in Chinese fermented acid rice soup

Na Liu

Guizhou University

Song Miao

Teagasc Food Research Centre Moorepark

Likang Qin (✉ lkqin@gzu.edu.cn)

Guizhou University <https://orcid.org/0000-0002-4156-2937>

Research

Keywords: *Kluyveromyces marxianus*, Ethyl acetate, Organic acids, RNA-seq, Proteome

DOI: <https://doi.org/10.21203/rs.3.rs-74924/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Recently, more chemical and biotechnological applications have been found in *Kluyveromyces marxianus* than *Saccharomyces cerevisiae* in the food field because they show advantageous metabolism features in the production of flavor components of interest. However, most of study demonstrated *Kluyveromyces marxianus* involved in ethanol synthesis in the dairy products in food fields. Our study aims to clarify the formation mechanism of ethyl acetate and organic acids in acid rice soup inoculated with *Kluyveromyces marxianus*.

Results: The higher concentration of ethyl acetate than ethanol and organic acids in fermented acid rice soup inoculated with *Kluyveromyces marxianus*. Up-regulated genes/proteins, including ADH1, ADH2, ADH6, ATF1, ACCT, and TES1, and down-regulated ALD family involved in glycolysis/gluconeogenesis and pyruvate metabolism played the crucial roles in the formation of ethyl acetate and other esters. In addition, up-regulated genes/proteins involved in starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, glycolysis/gluconeogenesis, TCA cycle, and pyruvate metabolism played the important roles in the formation of organic acids, ethanol and esters.

Conclusion: Our results reveals the formation mechanism of ethyl acetate and organic acids in acid rice soup inoculated with *K. marxianus* L1-1. This study provides the basis for improving aroma and taste of fermented foods and reveals the formation mechanisms of flavors in no-dairy products.

Background

Recently, more chemical and biotechnological applications have been found in non-saccharomyces yeasts than *Saccharomyces cerevisiae* in the food field because they show advantageous metabolism features in the production of flavor components of interest. *Kluyveromyces marxianus* (*K. marxianus*), as a non-saccharomyces yeast, has various advantages over mesophilic yeasts, such as fast growth rate, wide spectrum of substrates and reduced cooling cost [1]. It is a haploid, homothallic, thermotolerant, and hemiascomycetous yeast and is closely related to *Kluyveromyces lactis*. Unlike *S. cerevisiae*, *K. marxianus* has the assimilating capability of lactose, glucose and xylose. Notably, compared to *S. cerevisiae* and *K. lactis*, *K. marxianus* has the more significant ability concentrate on the intrinsic fermentation capability of various sugars at high temperatures [2]. *K. marxianus* grows in a wide temperature range from 4 to 52 °C, indicating it is thermotolerant and can be applied in the processes of low temperatures and high temperatures, which can also prevent the growth of microorganisms sensitive to heat [3]. Furthermore, the probiotic properties of *K. marxianus* have been extensively explored [4].

Interestingly, *K. marxianus* shows a great potential in the production of esters, which are key aromatic compounds in the food industry [5]. Ethyl acetate and other short-chain volatile esters are used as industrial solvents and perfume ingredients. It was reported that the global market demand of ethyl acetate was more than 1.7 million tons per year [6]. Therefore, it is significant to produce ethyl acetate in the food production industry. Three synthesis ways of ethyl acetate have been reported [7]: hemiacetal

oxidation (spontaneous formation of hemiacetal from acetaldehyde and ethanol under the action of enzymatic oxidization), condensation of ethanol and acetyl-CoA, and esterification of acetate and ethanol (the reverse synthesis of ethyl acetate from ethanol and acetate). However, it is difficult to produce ethyl acetate through esterification of ethanol and acetate because the ester-hydrolyzing activity of esterase is much higher than its ester-synthesizing activity [5]. The previous report demonstrated that ethyl acetate synthesis was characterized by the direct utilization of ethanol as a substrate or the hemiacetal reaction between sugar and acetaldehyde [5]. Ethyl acetate as an aroma component plays an increasingly important role in foods and other fields. However, the synthesis mechanism of ethyl acetate in *K. marxianus* in rice-acid is unknown. Most studies on *K. marxianus* focused on its role in alcoholic fermentation in the fuel industry and food field [8], however, the production of ethyl acetate or organic acids with *K. marxianus* was seldom reported. The metabolism of *K. marxianus* is less well understood than that of *S. cerevisiae*. The formation mechanisms of ethyl acetate, other esters or flavor compounds in fermented foods inoculated with *K. marxianus* are not yet completely understood.

More multi-omics analysis technologies have been used to explore the synthesis, metabolism and accumulation of nutrients and flavor components in foods. RNA Sequencing is a novel high-throughput sequencing technology with many advantages, such as much information, low data redundancy and accurate analysis. In addition, it does not require the background of genomics, but it can analyze the transcriptional expression of multiple materials [9]. Proteomics can enhance the understanding of the biochemical processes of flavor development in fermented foods [10]. However, due to non-coding RNA regulation, protein degradation, protein secretion, the quantity of differently expressed proteins (DEPs) are often less than differently expressed genes (DEGs). In addition, the detectable protein content, protease hydrolysis and operational errors in proteomics detection also limit the application of proteomic analysis [11]. Many factors lead to the difference in the analysis results of transcriptomics and proteomics. Notably, a previous report proved that sequence features contributed to 15.2–26.2% of total variations of mRNA and proteins [12]. Therefore, the combination of transcriptomics and proteomics may reveal the flavor formation mechanism in *K. marxianus* L1-1 in acid rice soup (rice-acid).

In order to clarify the formation mechanism of ethyl acetate and organic acids in rice-acid inoculated with *Kluyveromyces marxianus* L1-1, volatile compounds and organic acids in rice-acid were measured in this study. In addition, we analyzed the differently expressed genes and proteins of *K. marxianus* L1-1 in rice-acid in the key fermentation days (the first and third days) through the complementary analysis of mRNA sequencing and proteomics. Through GO enrichment analysis and KEGG pathway enrichment analysis, the formation mechanisms of ethyl acetate and organic acids in *K. marxianus* L1-1 in rice-acid were explored in this study.

Results

Variations of the quantities of *K. marxianus* L1-1 in the fermentation process of rice-acid

K. marxianus, which is able to utilize various sugars, may be a suitable microbe for lignocellulose hydrolysis and grain matrix at 30 °C [13]. In this study, the fermentation temperature was determined as 30 °C based on our previous study. The number of *K. marxianus* L1-1 changed significantly in the fermentation process. The number of *K. marxianus* L1-1 gained the most significant increase rate from 0 d to 1 d (Fig. 1a) and decreased from 1d to 2 d. Interestingly, it increased from 2 d to 3 d. The above variations may be related to the oxygen content in the fermentation tank. The limited supply of oxygen (the terminal electron acceptor) also initiated the synthesis of some esters, but it primarily forced ethanol production during the growth of *K. marxianus* DSM 5422 [14]. We explored the formation of ethyl acetate, other esters and organic acids based on the growth of *K. marxianus* L1-1 in this study. In the study, we investigated the key fermentation process of rice-acid with *K. marxianus* L1-1 in 1 d and 3 d and analyzed the formation mechanism of flavors in *K. marxianus* L1-1.

Variations of key volatile compounds in the fermentation process of rice-acid

The variations of flavor compounds in 1 d and 3 d were explored. The basic conditions for the formation of ethyl acetate are acetic acid, ethanol and some key enzymes which were discussed in this study. In the obtained volatile compounds, 5 key acids, 13 key alcohols and 12 key esters were found (Table 1). From 1 d to 3 d, ethyl acetate content increased from 162.98 ± 5.02 to 241.37 ± 6.20 g/kg; ethanol content increased from 36.11 ± 4.54 to 52.68 ± 14.45 g/kg; acetic acid content increased from 0.21 ± 0.06 to 32.67 ± 1.57 g/kg. Acetic acid, 2-phenylethyl ester, 2-methyl-propanoic acid, ethyl ester and 9 other esters were also found. Interestingly, ethyl acetate, ethanol and acetic acid are important volatile compounds because of their high contents and low odor thresholds [15]. Ethyl acetate made a significant contribution to the formation of the fruity flavor and promoted the overall flavor balance in rice-acid. Moreover, ethyl acetate exhibits probiotic properties such as being closely linked to the antioxidant function in the fruit [16]. The formation of esters in the alcoholization stage was closely related to the enzyme activity of yeasts. Therefore, it is necessary to explore the formation mechanism of ethyl acetate and other esters.

Table 1

The key volatile compounds (mg/L) and organic acids (mg/L) in fermented rice-acid inoculated with *K. marxianus* L1-1 at 1 d and 3 d.

volatile Compound	RT	RI	RIL	1 d	3 d
Acids (5)					
Acetic acid	12.69	1508	1449	0.21 ± 0.06	32.67 ± 1.57
Propanoic acid, 2-methyl-	14.69	1671	1570	0.55 ± 0.13	1.12 ± 0.21
Butanoic acid, 2-methyl-	16.05	1791	1662	0	0.33 ± 0.19
Butanoic acid, 3-methyl-	16.22	1806	1666	0.09 ± 0.01	0.93 ± 0.32
Hexanoic acid	18.17	1991	1846	0.36 ± 0.05	0.73 ± 0.29
Alcohol (13)					
Ethanol	3.76	937	932	36.11 ± 4.54	52.68 ± 14.45
1-Propanol	5.52	1043	1036	1.30 ± 0.76	1.13 ± 0.25
1-Propanol, 2-methyl-	6.47	1096	1092	20.89 ± 6.43	38.77 ± 10.50
1-Butanol, 3-methyl-	8.39	1209	1209	17.04 ± 4.79	31.84 ± 5.27
1-Pentanol	9.07	1252	1250	0.36 ± 0.11	0.89 ± 0.15
1-Hexanol	10.60	1355	1355	5.57 ± 0.35	3.38 ± 0.92
1-Octen-3-ol	11.92	1450	1450	0.89 ± 0.31	0.43 ± 0.08
1-Heptanol	12.01	1457	1453	1.49 ± 0.51	0.81 ± 0.30
1-Hexanol, 2-ethyl-	12.46	1490	1491	4.23 ± 1.12	4.87 ± 1.79
2-Nonanol	12.83	1519	1521	0.03 ± 0.00	0.64 ± 0.19
1-Octanol	13.33	1559	1557	1.79 ± 0.67	1.69 ± 0.11
1-Nonanol	14.57	1661	1660	2.26 ± 1.10	1.75 ± 0.48
Phenylethyl Alcohol	17.43	1919	1906	12.22 ± 4.65	13.68 ± 3.62

RI: the linear retention indices calculated from a series of n-alkanes (C6-C26).

RIL: retention indices referred to the literature value with same or equivalent chromatographic column shown on <http://www.flavornet.org/flavornet.html> and NIST Chemistry WebBook (<http://webbook.nist.gov/chemistry/>).

Method of identification: (A) by comparison of the mass spectrum with the NIST/Wiley mass spectral library; (B) by comparison of RI (Kovats indices) with RI of an authentic compound; and (C), by comparison of retention time and spectrum of an authentic compound.

volatile Compound	RT	RI	RIL	1 d	3 d
Esters (12)					
Ethyl Acetate	3.05	886	888	162.98 ± 5.02	241.37 ± 6.20
Propanoic acid, ethyl ester	4.06	956	953	5.03 ± 1.44	3.06 ± 0.69
Propanoic acid, 2-methyl-, ethyl ester	4.19	964	961	6.43 ± 2.84	8.83 ± 1.17
Butanoic acid, methyl ester	4.51	984	982	0.75 ± 0.25	0.65 ± 0.17
Isobutyl acetate	4.98	1013	1012	3.56 ± 0.72	9.26 ± 1.53
1-Butanol, 3-methyl-, acetate	6.88	1120	1122	2.36 ± 0.76	3.30 ± 0.84
1-Butanol, 2-methyl-, acetate	6.90	1121	1125	0	4.74 ± 6.70
Acetic acid, pentyl ester	7.73	1169	1176	0.06 ± 0.01	0.83 ± 0.16
2-Propenoic acid, ethenyl ester	8.42	1211		0	6.85 ± 1.68
Acetic acid, hexyl ester	9.37	1271	1272	0.29 ± 0.11	0.39 ± 0.11
Acetic acid, 2-phenylethyl ester	16.41	1823	1813	28.29 ± 7.93	38.11 ± 4.26
Propanoic acid, 2-methyl-, 2-phenylethyl ester	17.11	1889	1896	1.61 ± 0.21	2.26 ± 0.62
Organic acids				1d	3d
L-lactic acid				3.01 ± 0.61	6.02 ± 1.67
Acetic acid				0.007 ± 0.003	0.51 ± 0.06
Malic acid				0.77 ± 0.06	0.14 ± 0.08
Succinic acid				0.022 ± 0.004	0.072 ± 0.006
Citric acid				0	0.005 ± 0.001
Oxalic acid				0.036 ± 0.005	0.43 ± 0.03

RI: the linear retention indices calculated from a series of n-alkanes (C6-C26).

RIL: retention indices referred to the literature value with same or equivalent chromatographic column shown on <http://www.flavornet.org/flavornet.html> and NIST Chemistry WebBook (<http://webbook.nist.gov/chemistry/>).

Method of identification: (A) by comparison of the mass spectrum with the NIST/Wiley mass spectral library; (B) by comparison of RI (Kovats indices) with RI of an authentic compound; and (C), by comparison of retention time and spectrum of an authentic compound.

volatile Compound	RT	RI	RIL	1 d	3 d
Tartaric acid				0.029 ± 0.007	0.043 ± 0.005
RI: the linear retention indices calculated from a series of n-alkanes (C6-C26).					
RIL: retention indices referred to the literature value with same or equivalent chromatographic column shown on http://www.flavornet.org/flavornet.html and NIST Chemistry WebBook (http://webbook.nist.gov/chemistry/).					
Method of identification: (A) by comparison of the mass spectrum with the NIST/Wiley mass spectral library; (B) by comparison of RI (Kovats indices) with RI of an authentic compound; and (C), by comparison of retention time and spectrum of an authentic compound.					

Variations of organic acids in the fermentation process of rice-acid

Seven organic acids were found in rice-acid, including L-lactic acid, acetic acid, malic acid, succinic acid, citric acid, oxalic acid and tartaric acid (Table 1). Among 7 organic acids, L-lactic acid had the highest content. The content of L-lactic acid increased from 3.01 ± 0.61 g/kg in Day 1 to 6.02 ± 1.67 g/kg in 3 d. The contents of the other 6 organic acids did not show the significant increase during the fermentation process. However, although the contents of the 6 other organic acids were low, they interacted with each other to promote the formation of the sourness and taste of rice-acid. In our study, both volatile components and organic acids affected the formation of the flavor of rice-acid. Lactic acid exists in two isomeric forms which include L-(+) and D-(-)-Lactic acid. It is produced by microbial fermentation and chemical synthesis and used in food, cosmetic, pharmaceutical, and chemical industries [17]. In the study, we mainly focused on the increase in L-(+) -Lactic acid caused by microbial fermentation. L-(+) -Lactic acid with a high enantiomeric purity is required in many industries, especially in medical, pharmaceutical and food industries, since D-(-)- Lactic acid is harmful to humans and can cause decalcification or acidosis [18]. L-(+) -Lactic acid not only promoted the formation of the flavor in rice-acid, but also had an important effect on the health. We will further explore the genes, proteins and enzymes associated with the formation of organic acids.

Transcriptomic analysis of *K. marxianus* L1-1 during the key fermentation period of rice-acid

In the fermentation process of rice-acid, Y 1 d and Y 3 d initiated the changes in the expressions of a series of genes. The two independent cDNA libraries (Y 1 d and Y 3 d) constructed for high-throughput sequencing respectively yielded 19,617,191 ~ 27,627,279 pair-end reads and 5,885,157,300 ~ 8,288,183,700 clean reads after stringent quality check and data filtering (Q20 bases > 97.28%, Q30 bases > 92.01%, G + C 42.63%~43.03%) (Table 2). All of high-quality clean reads were used for gene comparison and more than 88% of total reads of clean reads were mapped to the database by software Bowtie 2. By using $|\log_2FC| > 1.5$ and $FDR < 0.05$, we identified 1390 DEGs (788 up-regulated and 602 down-regulated) between Y 1 d and Y 3 d (Fig. 2a).

Table 2
Quality assessment results of RNA sequencing in *K. marxianus* L1-1 when inoculated rice-acid.

Sample ID	Read Sum	Base Sum	GC (%)	Q20 (%)	Q30 (%)
Y1d-1	19,617,191	5,885,157,300	42.90%	97.33%	92.07%
Y1d-2	24,816,214	7,444,864,200	42.91%	97.36%	92.19%
Y1d-3	22,852,031	6,855,609,300	42.63%	97.28%	92.01%
Y3d-1	27,627,279	8,288,183,700	42.99%	97.44%	92.34%
Y3d-2	22,987,221	6,896,166,300	42.95%	97.35%	92.18%
Y3d-3	25,331,202	7,599,360,600	43.03%	97.36%	92.18%

GO analysis of the DEGs showed the enrichment of three major cellular components, biological processes and molecular functions (Fig. 3). In terms of cellular components, most of the DEGs were enriched in nucleolus, small-subunit processome and preribosome, and large-subunit precursor. In terms of biological processes, most of the DEGs were enriched in endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, and LSU-rRNA). In terms of molecular functions, these DEGs were enriched in structural constituent of ribosome, snoRNA binding and rRNA binding. All DEGs were subjected to KEGG pathway enrichment analysis. KEGG analysis assigned the DEGs of Y1-d and Y3-d to metabolic pathways. At least 4231 genes were identified, including 1390 DEGs annotated to 279 KEGG pathways (Table S1). The significantly enriched pathways (p -value < 0.05 and q -value < 0.05) were ribosome, cytosolic DNA-sensing pathway, RNA polymerase, DNA replication, ribosome biogenesis in eukaryotes, pyrimidine metabolism and purine metabolism. Importantly, the crucial KEGG pathway related to the ethyl acetate and organic acids included amino sugar and nucleotide sugar metabolism, starch and sucrose metabolism, glycolysis/gluconeogenesis, pyruvate metabolism and TCA cycle. These pathways showed the different roles of *K. marxianus* L1-1 in the formation of flavor and taste of rice-acid.

Ethyl acetate and organic acids promoted the maturity of flavor and taste of rice-acid inoculated with *K. marxianus* L1-1. Key genes were found to be involved in the ethyl acetate metabolism process in glycolysis, including GLK1 (K00844), GPD2 (K00134), 3 genes of ALD family (K00129 and 2 K00128), 6 genes of ADH family (K13953, K13953 and 4 genes of K13953) and ATF1 protein (BA042650, BA042650 and BA042650) (Table 3). The genes of GLK1, GPD2, ALD family, ADH family and ATF1 encode glucokinase-1, glyceraldehyde 3-phosphate dehydrogenase, aldehyde dehydrogenase, alcohol dehydrogenase and alcohol O-acetyltransferase, respectively. In addition, the gene ERG10 encoding acetyl-CoA C-acetyltransferase was found to be involved in the ethyl acetate metabolism process in pyruvate metabolism. The key genes related to organic acids found in pyruvate metabolism included CYB2 (K00101), DLD1 (K00102), 2 genes of MDH family (2 K00026), and FUM1 (K01679) respectively encoding L-lactate dehydrogenase, D-lactate dehydrogenase, malate dehydrogenase and fumarate

hydratase. The key genes related to organic acids found in the citrate cycle included 2 genes of CIT family (2 K01647), 3 genes of SDH family (K00234, K00236 and K00237) and 2 genes of MDH2 (K00026) respectively encoding citrate synthase succinate dehydrogenase (ubiquinone) flavoprotein subunit and malate dehydrogenase. The key KEGG pathways related to the formation mechanism of the flavor based on the DEGs are discussed below.

Table 3

The different expression of genes and proteins (DEGs and DEPs) in key KEGG pathways of *K. marxianus* L1-1 when inoculated rice-acid.

Seq_ID	Description	KO/Gene_ID	KEGG/ Genename	Log2FC/ Transcription	Log2FC/ Proteomics
Starch and sucrose metabolism					
KLMA_10051	glucokinase-1,hexokinase [EC:2.7.1.1]	K00844	GLK1, HK	1.83	13.384
KLMA_30608	glucan 1,3-beta-glucosidase [EC:3.2.1.58]	K01210	SCW4	1.53	6.986
Amino sugar and nucleotide sugar metabolism					
KLMA_10051	glucokinase-1,hexokinase [EC:2.7.1.1]	K00844	GLK1	1.83	13.384
Glycolysis / Gluconeogenesis					
KLMA_10051	glucokinase-1 [EC:2.7.1.1]	K00844	GLK1	1.83	13.384
KLMA_10836	alcohol dehydrogenase, propanol-preferring [EC:1.1.1.1]	K13953	ADH2	1.47	None
KLMA_20673	aldehyde dehydrogenase (NAD(P)+) [EC:1.2.1.5]	K00129	ALD2	-2.15	None
KLMA_30203	alcohol O-acetyltransferase [EC:2.3.1.84]	K00664	ATF1	2.90	None
KLMA_40102	alcohol dehydrogenase, propanol-preferring [EC:1.1.1.1]	K13953	ADH2	-1.77	1.65
KLMA_40404	aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]	K00128	ALD5	-1.65	None
KLMA_40624	alcohol dehydrogenase	None	ADH	None	2.047
KLMA_50012	aldehyde dehydrogenase (NAD+) [EC:1.2.1.3]	K00128	ALD4	-1.07	None

Note: FC value represented the differential expression multiple of mRNA and protein.

Seq_ID	Description	KO/Gene_ID	KEGG/ Genename	Log2FC/ Transcription	Log2FC/ Proteomics
KLMA_70007	alcohol dehydrogenase, propanol-preferring [EC:1.1.1.1]	K13953	ADH2	2.10	None
KLMA_70462	alcohol dehydrogenase, propanol-preferring [EC:1.1.1.1]	K13953	ADH2	3.50	None
KLMA_80306	alcohol dehydrogenase, propanol-preferring [EC:1.1.1.1]	K13953	ADH3	-1.42	None
KLMA_80339	NADP-dependent alcohol dehydrogenase 6	none	ADH6	none	1.539
KLMA_80427	alcohol dehydrogenase, propanol-preferring [EC:1.1.1.1]	K13953	ADH1	2.83	None
Pyruvate metabolism					
KLMA_10621	L-lactate dehydrogenase (cytochrome) [EC:1.1.2.3]	K00101	CYB2	2.93	None
KLMA_10649	D-lactate dehydrogenase (cytochrome) [EC:1.1.2.4]	K00102	DLD1	-1.20	None
KLMA_10771	homocitrate synthase [EC:2.3.3.14]	K01655	LYS21	1.58	2.073
KLMA_30013	L-lactate dehydrogenase (cytochrome) [EC:1.1.2.3]	K00101	CYB2	-3.88	None
KLMA_40055	malate dehydrogenase [EC:1.1.1.37]		MDH3	None	1.523
KLMA_40341	L-lactate dehydrogenase (cytochrome) [EC:1.1.2.3]	K00101	CYB2	1.13	None

Note: FC value represented the differential expression multiple of mRNA and protein.

Seq_ID	Description	KO/Gene_ID	KEGG/ Genename	Log2FC/ Transcription	Log2FC/ Proteomics
KLMA_40583	D-lactate dehydrogenase (cytochrome) [EC:1.1.2.4]	K00102	DLD1	-2.06	None
KLMA_60086	acyl-coenzyme A, acyl-coenzyme A thioester hydrolase	K01068	TES1	-1.13	1.517
KLMA_60167	malate dehydrogenase	K00026	MDH1	1.12	2.984
KLMA_60383	acetyl-CoA C-acetyltransferase [EC:2.3.1.9]	K00626	ACCT	1.07	None
KLMA_80028	fumarate hydratase	K01679	FUM1	1.03	2.034
TCA cycle					
KLMA_10253	pyruvate carboxylase [EC:6.4.1.1]	K01958	PYC2	1.50	None
KLMA_20105	citrate synthase [EC:2.3.3.1]	K01647	CIT1	2.28	2.47
KLMA_20466	succinate dehydrogenase (ubiquinone) flavoprotein subunit [EC:1.3.5.1]	K00234	SDH1	1.25	1.75
KLMA_30124	succinate dehydrogenase (ubiquinone) cytochrome b560 subunit	K00236	SDH4	1.28	None
KLMA_40055	malate dehydrogenase	None	MDH3	None	1.523
KLMA_40391	malate dehydrogenase	K00026	MDH2	2.07	3.539
KLMA_50550	succinate dehydrogenase (ubiquinone) membrane anchor subunit	K00237	SDH4	1.36	None
KLMA_60167	malate dehydrogenase [EC:1.1.1.37]	K00026	MDH1	1.12	None
KLMA_60167	malate dehydrogenase [EC:1.1.1.37]	K00026	MDH2	1.12	2.984

Note: FC value represented the differential expression multiple of mRNA and protein.

Seq_ID	Description	KO/Gene_ID	KEGG/ Genename	Log2FC/ Transcription	Log2FC/ Proteomics
KLMA_80028	fumarate hydratase, class II [EC:4.2.1.2]	K01679	FUM1	1.03	2.034
KLMA_80408	succinate dehydrogenase [ubiquinone] iron-sulfur subunit	None	SDH2	None	1.651
Note: FC value represented the differential expression multiple of mRNA and protein.					

Proteomics Characterization

The total proteins were extracted from the Y1-d and the Y3-d at filling stage and subjected to 4D label-free proteomics analysis to complement the transcriptome analysis. According to the abundance levels of proteins, 610 proteins were identified as DEPs at p-value < 0.05, including 135 proteins with increased abundance levels and 475 proteins with decreased abundance levels (Fig. 2b), and the difference ratio reached > 1.5. The number of up-regulated proteins was smaller than that of down-regulated proteins at the filling stage since the growth of *K. marxianus* L1-1 was inhibited due to the acid environment in the later fermentation stage of rice-acid.

To obtain a global diagram of proteomic changes, at least 2937 proteins were identified and 187 DEPs were annotated with GO analysis and KEGG analysis (Table S2). In GO functional analysis, 187 proteins were annotated to 59 GO terms. The results of GO analysis showed that the distributions of DEPs in functional classification were consistent with the distributions of transcription levels of DEGs (Fig. 4). In terms of cellular components, most of the up-regulated DEPs were enriched in mitochondrion, tricarboxylic acid cycle enzyme complex and mitochondrial matrix. In terms of the biological process, most of the up-regulated DEPs were enriched in citrate metabolic process, tricarboxylic acid metabolic process and galactose catabolic process. In terms of molecular functions, these up-regulated DEPs were enriched in oxidoreductase activity, L-malate dehydrogenase activity, malate dehydrogenase activity and alcohol dehydrogenase activity. The 6 GO terms involved in alcohol dehydrogenase activity had the smallest p-value (p-value < 0.01) and were related to the formation of ethyl acetate. The 187 DEPs were annotated to 30 KEGG pathways (Table S2). Most up-regulated KEGG pathways analyzed by proteomics were related the formation of ethyl acetate and organic acids, which explained reasonably the formation of flavor and taste in rice-acid inoculated with *K. marxianus* L1-1. Meanwhile, the down-regulated KEGG pathways obtained by proteomics analysis were most related to the growth of *K. marxianus* L1-1 (Fig. 5b), verifying that it was reasonable to select the third day as the ending of rice-acid fermentation. According to the pathway analysis (Fig. 5a), we could conclude that many proteins took part in various metabolic pathways including amino sugar and nucleotide sugar metabolism, starch and sucrose metabolism, glycolysis/gluconeogenesis, pyruvate metabolism and citrate cycle, which might affect many aspects of the metabolism of *K. marxianus* L1-1 during the key fermentation period of rice-acid.

However, the different KEGG pathways obtained by proteomics analysis played different roles in the formation of flavor and taste in rice-acid.

Ethyl acetate is the most important volatile compound in rice-acid inoculated with *K. marxianus* L1-1. Four key proteins were found to be involved in the ethyl acetate metabolism process in glycolysis, including GLK1(BAO37673), GAP1 (BAO40242), 2 ADH1 (BAO40648 and BAO40126) and ADH6 (BAO42650) (Table 2). GLK1, GAP1, ADH1 and ADH6 encode glucokinase-1, glyceraldehyde-3-phosphate dehydrogenase 1, alcohol dehydrogenase and NADP-dependent alcohol dehydrogenase, respectively. In addition, four key proteins were found to be involved in organic acids metabolic process in pyruvate metabolism, including 2 MDH family (BAO41458 and BAO40079), FUM1 (BAO42339) and LYS21 (BAO38393). MDH, FUM1 and LYS21 encode malate dehydrogenase, fumarate hydratase and homocitrate synthase, respectively. Four key proteins were found to be involved in the metabolism process of organic acids in citrate cycle, including CIT1 (BAO38563), SDH1 (BAO38924), MDH2 (BAO40415) and FUM1 (BAO42339), which respectively encode citrate synthase, succinate dehydrogenase (ubiquinone) flavoprotein subunit, malate dehydrogenase and fumarate hydratase. The proteins related to the metabolism of organic acids were consistent with the genes, indicating that the combination of transcriptomics and proteomics were useful tools to analyze the formation of ethyl acetate and organic acids during the key fermentation period of rice-acid inoculated with *K. marxianus* L1-1.

Correlation Analysis Of Transcriptome And Proteome Data

Transcriptomic and proteomic analysis results are shown in Fig. 1b-1d. At least 4231 genes were identified and 1390 DEGs were annotated to 279 KEGG pathways by using transcriptomic analysis. In addition, 2937 proteins were identified and 610 DEPs were annotated to 30 KEGG pathways by using proteomics analysis. Pearson correlation coefficient was 0.3761 and the results of the two analysis methods were significantly different. Therefore, the combination of transcriptome and proteome data could be an effective way to reveal the formation mechanism of the flavor in rice-acid inoculated with *K. marxianus* L1-1. The 5 KEGG pathways related to the synthesis of ethyl acetate and organic acids are shown in Table 3.

Discussion

Rice-acid, as a cereal-based fermented food used for seasoning, is famous in China. However, the traditional rice-acid process requires two times of fermentation and the long-term fermentation time may lead to the unstable and non-persistent flavor. In this study, we adopted a novel inoculation strain (*K. marxianus* L1-1), the inoculation of *K. marxianus* L1-1 promoted the fermented rice-acid has the unique flavor and shorten the fermentation period of rice-acid from 40 d to 4 d. Our previous study proved that this fermentation method could realize the high-quality flavor. However, the formation mechanism of the flavor in rice-acid inoculated with *K. marxianus* L1-1 is not clear. In this study, the RNA-seq and 4D label-free technologies were used to explore the genes and proteins in the formation mechanism of ethyl

acetate and organic acids in rice-acid inoculated with *K. marxianus* L1-1. DEGs and DEPs were identified and annotated to key KEGG metabolism pathways, including starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, glycolysis/gluconeogenesis, pyruvate metabolism and TCA cycle. Furthermore, the results of transcriptome and proteome were combined to reveal the formation mechanism of ethyl acetate and organic acids. We provided a comprehensive interpretation and exact measurements of genes and protein expressions involved in the changes of the flavor of rice-acid for the first time.

Up-regulated proteins and genes involved in starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism provide the energy for the formation of acids, ethanol and esters

Starch and sucrose metabolism provides an important transient pool in the sugar accumulation pathways. The genes and proteins could provide the energy for the formation of the flavor and taste in rice-acid, including GLK1 encoding glucokinase-1 and KLMA_10051 encoding hexokinase in the KEGG pathways of starch and sucrose metabolism and amino sugar and nucleotide sugar metabolism (Table 3). Protein GLK1 was a glycolysis-initiating enzyme [19] and showed the 13.384-Log2FC upregulation. The gene of KLMA_10051 encodes a hexokinase, which showed the 1.83-Log2FC upregulation. Up-regulated GLK1 indicated an increase in NADPH amount and more energy generated in *K. marxianus* L1-1. Lane et al. [20] also reported that the catabolite repression could be reduced by modulating the expression of glucose-phosphorylating enzymes, such as GLK1 and hexokinase (HK). In addition, SCW4 (KLMA_30608) encoding glucan 1,3-beta-glucosidase showed the upregulation of 6.986-Log2FC, as indicated by proteomic analysis, and the upregulation of 1.53-Log2FC, as analyzed by transcriptomic analysis. SCW4 (KLMA_30608) encoding glucan 1,3-beta-glucosidase might have the hydrolytic activity and provide the energy to promote the formation of the flavor and taste in rice-acid. In a previous report, it was also indicated that a 1,3- β -glucosidase BGL1 purified from the pilei showed the hydrolytic activity toward laminarin, laminarioligosaccharides including laminaribiose, and p-nitrophenyl- β -d-glucopyranoside (pNPG) [21]. In addition, another study also showed that 1,3- β -glucosidase had a certain hydrolytic activity towards gentiobiose, cellobiose, and related polysaccharides [22]. Therefore, the hydrolysis of sugar compounds may provide substrates and energy for orderly metabolism. Eventually, with hydrolytic enzymes, carbohydrates were easily converted into glucose in *K. marxianus* L1-1. Our fermentation belonged to a static fermentation method. *K. marxianus* seemed to enhance glucose metabolism and shift to fermentation, implying the connection between oxygen and glucose-sensing pathways [13]. In future study we will explore the content on oxygen in the fermenter and its correlation with flavor in rice-acid. However, genes and proteins were differentially expressed in the different fermentation days (the first and third days) of rice-acid with the inoculation of *K. marxianus* L1-1. The correlation between transcriptomics and proteomics data was not perfect and there are some differences in the results of the two methods [23]. Protein expressions may be affected by various factors in the translational stage [12]. The efficiency of protein biosynthesis and accumulation depends on various factors in the biological regulation process. Therefore, we adopted the complementary analysis method of transcriptomics and proteomics to reveal formation mechanism of ethyl acetate, ethanol and organic acids in Chinese rice-acid inoculated with *K. marxianus*.

Up-regulated proteins and genes involved in glycolysis/gluconeogenesis and pyruvate metabolism played an important role in the formation of ethyl acetate and other esters

Volatile esters are secondary metabolites produced by yeasts and fungi during fermentation of fermented foods [24]. Interestingly, our study showed compared to alcohol, more esters existed in rice-acid inoculated with *K. marxianus* L1-1. This difference contributed to the flavor formation in rice-acid. Ethyl acetate was one of most important volatile esters in rice-acid. Both DEGs and DEPs in glycolysis/gluconeogenesis and pyruvate metabolism played an important role in the formation of ethyl acetate and other esters (Fig. 7). Glycolysis is the cytosolic pathway that converts glucose to pyruvate. ADH1, ADH2, ADH3 (alcohol dehydrogenase) and ADH6 (NADP-dependent alcohol dehydrogenase 6) involved in glycolysis/gluconeogenesis respectively encode the proteins of BAO40648, BAO40126 and BAO42650 (Table 3 and Fig. 6a). This type of esterification is carried out from primary and secondary alcohols, aldehydes or ketones [25]. ADH enzymes catalyze the synthesis of ethyl acetate through the oxidation of hemiacetal. ADH3 was only found in the transcriptomic analysis in this study. A previous study proved that ADH2 was constitutively expressed in aerobic growth with glucose as a carbon source, whereas ADH3 expression increased as cells reached the stationary phase. These results were in agreement with the previous analyses of *K. marxianus* transcriptome [26]. Our study proved that the ADH family was important in the complementary analysis of transcriptomics and proteomics. ADH family had a significant effect on the formation of ethyl acetate and ethanol, and the up-regulated genes and proteins suggested that ADH1, ADH2 and ADH6 were the dominant enzymes in ethanol production when glucose was used as a carbon source. In the identified up-regulated genes and proteins, ADH2 and ADH6 were critical in the reduction of acetaldehyde to ethanol (a precursor to ethyl acetate). This result was consistent with the previous study [26]. Another report proved that the alcohol acetyltransferase was associated with intracellular lipid particles in cytosol [27]. We analyzed the up-regulated genes and proteins of ADH family and found that they promoted the formation of ethanol and ethyl acetate. Herein, one possible formation pathway was the oxidation of hemiacetal (the spontaneous product of ethanol and acetaldehyde) under the catalysis action of ADH activity. Although the two methods of transcriptomics and proteomics showed some differences, all the up-regulated ADH genes and proteins indicated that the alternative biosynthetic routes of ethyl acetate existed in *K. marxianus* L1-1.

In addition, the formation of ethyl acetate in rice-acid fermentation inoculated with *K. marxianus* L1-1 was mainly catalyzed by two enzymes named ATF1 (alcohol O-acetyltransferase) and TES1 (acyl-coenzyme A) in this study (Table 3 and Fig. 6b), which possessed an acyl-coenzyme A: ethanol O-acyltransferase (AEATase) activity as well as the esterase activity. We found that ATF1 could use ACCT (acetyl-CoA C-acetyltransferase) to synthesize acetate esters, including ethyl acetate, acetic acid, 2-phenylethyl ester and isobutyl acetate (Fig. 7). The previous analysis of the ATF1p also found that acetyl-CoA was used to synthesize acetate esters [28, 29]. It was demonstrated that acetyl-CoA taking a crucial role to produce the generation of more NADH and more ATP in glucose [13]. The previous study postulated that ester biosynthesis in *K. marxianus* may also occur through homologs to the medium-chain acyltransferases from *S. cerevisiae*, the isoamyl acetate-hydrolyzing esterase, the N-acetyltransferase Sli1 and/or the alcohol-O-acetyltransferase [30]. Interestingly, we also found that the

another important ester family including propanoic acid, 2-methyl-propanoic acid, ethyl ester, 2-propenoic acid, ethenyl ester and propanoic acid, 2-methyl-2-phenylethyl ester, and the increase in propanoic acid, ethyl ester was closely related to propanoate metabolism. The propanoate is expected to be converted to acetyl-CoA or pyruvate, as suggested by examination of likely propanoate metabolism. It was also demonstrated that propanoate was converted to acetyl-CoA in three classes of mycolate [31]. However, the differentially expressed gene ATF1 was only found under the transcriptomics analysis. The combination method of transcriptomics and proteomics could provide the more reasonable explanation for the formation of ethyl acetate and propanoic acid, ethyl ester.

Furthermore, some genes and proteins related to the formation of ethyl acetate and other esters included GLK1 (glucokinase-1), KLMA_10051 (hexokinase), GPD2 (glyceraldehyde 3-phosphate dehydrogenase) and ALD family (aldehyde dehydrogenase) (Table 3 and Fig. 7). Protein GLK1 was a glycolysis-initiating enzyme, which promoted the formation of ethyl acetate and other esters and also played an important role in glycolysis/gluconeogenesis and pyruvate metabolism. Glyceraldehydes-3-P and pyruvate were the intermediate products in the glycolysis process and provided the carbon skeleton for volatile compound biosynthesis in rice-acid. A recent transcriptomic study suggested that a β -glucosidase homolog in *K. marxianus* may be responsible for cellobiose degradation [13]. Interestingly, the genes of ALD family (aldehyde dehydrogenase showed the downregulation, indicating that more energy was used in ADH family encoding alcohol dehydrogenase. In the further study, we will explore the change of aldehyde dehydrogenase. Some different genes were involved in the formation of ethyl acetate in *K. marxianus* and *S. cerevisiae*, although the metabolism of ethyl acetate in *K. marxianus* was seldom reported. The previous study suggested that the biosynthesis of acetate ester could be interpreted with the antagonistic activity of esterase IAH1 [32], but we did not find a reverse esterase playing a role in the formation of ethyl acetate and other esters in our study (Fig. 6c). A previous study also reported that there was no esterase involved in the biosynthesis ethyl acetate or other esters in *K. marxianus* CBS 6556 [26].

Up-regulated genes and proteins involved in TCA cycle and pyruvate metabolism played important roles in the formation of organic acids

The mitochondrial TCA cycle, also known as Krebs cycle, is one of the major pathways of carbon metabolism in higher organisms that provides electrons during oxidative phosphorylation within the inner mitochondrial membrane. TCA cycle is crucial in mitochondrial membranes for respiration. Both TCA cycle and pyruvate metabolism played important roles in the formation of organic acids (Table 3 and Fig. 7). It was also reported that organic acids were closely related to TCA cycle in rice [33]. The up-regulated gene PYC2 (KLMA_10253) encoding pyruvate carboxylase played the crucial role in TCA cycle and pyruvate metabolism. It was demonstrated that pyruvate carboxylase as an anaplerotic enzyme had a special effect and played an essential role in various cellular metabolic pathways including gluconeogenesis, glucose-induced insulin secretion, de novo fatty acid synthesis and amino acid synthesis [34]. The DEGs and DEPs related to TCA cycle and pyruvate metabolism reasonably interpreted the formation of organic acids during the key fermentation period of rice-acid inoculated with *K. marxianus* L1-1.

The up-regulated gene LYS21 (KLMA_10771) encoding homocitrate synthase, has been identified by the combined analysis of transcriptomics and proteomics in this study. The previous study proved that homocitrate synthase was responsible for the first important step of the pathway and played the crucial role in pyruvate metabolism [35]. Notably, homocitrate synthase LYS21 was linked to the key process of DNA damage repair in a nucleus and TCA cycle in the cytoplasm [13, 36]. The up-regulated gene CIT1 (KLMA_20105) encoding citrate synthase reasonably interpreted the increase in citric acid in the fermentation process of rice-acid. CIT1 acts as a quantitative marker for healthy mitochondrion and is encoded by the nuclear DNA [37]. Interestingly, lactic acid has optical isomers: L-lactic acid and D-lactic acid, which can be produced by chemical synthesis (DL-lactic acid) or microbial fermentation (L-lactic acid, D-lactic acid, or DL-lactic acid). Compared to chemical synthesis processes, microbial fermentation processes present more advantages since they make use of renewable substrates from lactic acid bacteria [38]. Consistently, our study demonstrated that fermented rice-acid inoculated with *K. marxianus* L1-1 could produce more concentration of L-lactic acid in 3 day than 1 day, and the L-lactic acid has some advantages for healthy. The genes CYB2 (KLMA_10621 and KLMA_40341) encoding L-lactate dehydrogenase were up-regulated, whereas another CYB2 (KLMA_30013) was down-regulated. DLD1 (KLMA_40583 and KLMA_10649) genes encoding D-lactate dehydrogenase were down-regulated in pyruvate metabolism and the down-regulation of DLD1 partially interpreted the increase in L-lactic acid during the fermentation process. The up-regulated genes and proteins including MDH1 (KLMA_60167), MDH2 (KLMA_60167), MDH2 (KLMA_40391), and MDH3 (KLMA_40055) encoding malate dehydrogenase in pyruvate metabolism and TCA cycle were related to the increase in malic acid. FUM1 (KLMA_80028) encoding fumarate hydratase in TCA cycle and pyruvate metabolism was up-regulated and this upregulation reasonably explained the conversion from fumarate to malate. SDH1 (KLMA_20466) encoding succinate dehydrogenase in TCA cycle was related to the increase in succinate. The up-regulated proteins SDH1 (KLMA_80408) was only detected in proteomics and the up-regulated genes SDH1 (KLMA_30124 and KLMA_50550) were only detected in transcriptomics. Furthermore, it was reported proved that fumarate hydratase (FUM1) and succinate dehydrogenase (SDH) were tumour suppressors [39]. Indicating *K. marxianus* L1-1 could have potential probiotic characteristics. Therefore, the enhanced activities of proteins and enzymes in TCA cycle and pyruvate metabolism indicated the increased organic acids in rice-acid inoculated with *K. marxianus* L1-1.

Down-regulated proteins and genes indicated the stable formation of the flavor

Interestingly, most of the genes and proteins involved in the 5 KEGG pathways includes starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, glycolysis/gluconeogenesis, pyruvate metabolism and TCA cycle were up-regulated except the genes CTS1, ALD family and DLD1. The reason had been discussed above. The up-regulated genes and proteins played the active role in the formation of flavor during rice-acid fermentation, whereas the down-regulated genes and proteins played the important role in maintaining the stable key flavor. Many reports focused on the up-regulated genes or proteins in the role of promoting the flavor maturity of fermented foods, the down-regulated genes or proteins were seldom reported. In our study, as seen from Fig. 5B, the down-regulated genes and proteins were involved in the KEGG pathways, including DNA replication, meiosis-yeast, homologous

recombination, mismatch repair, autophagy-yeast, MAPK signaling pathway-yeast, phenylalanine metabolism, base excision repair, nucleotide excision repair, other types of O-glycan biosynthesis, tryptophan metabolism, cell cycle-yeast, penicillin and cephalosporin biosynthesis and D-Arginine and D-ornithine metabolism. Most of the down-regulated pathways were related to the growth of *K. marxianus* L1-1 and this result provided a reasonable explanation of the decrease in the quantity of *K. marxianus* L1-1 in rice-acid during the key fermentation period (Day 3). The down-regulated proteins in DNA replication had the small Log2FC value (data is not displayed), indicating that related DEPs had small effects on the growth of *K. marxianus* L1-1. Therefore, the third day was the suitable ending of the fermentation process. In addition, the down-regulated KEGG pathways were related to the decomposition and utilization of substrates by *K. marxianus* L1-1. Consistently, a previous report demonstrated that down-regulated proteins related to advanced glycation end products were implicated in the aging process [40]. In the future, we will focus on the influences of the content of substrate of *K. marxianus* L1-1 on the flavor formation of rice-acid.

Conclusion

The transcriptome and proteome of *K. marxianus* L1-1 in rice-acid were determined in the study. The differentially expressed genes and proteins related to the formation of ethyl acetate and organic acids were determined. DEGs and DEPs were identified and found to be enriched in the key KEGG metabolism pathways, including starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, glycolysis/gluconeogenesis, pyruvate metabolism and TCA cycle. With the complementary analyses of the transcriptome and proteome, we revealed the formation mechanism of ethyl acetate and organic acids in Chinese rice-acid inoculated with *K. marxianus* L1-1. This study provides the basis for improving aroma and taste of fermented foods and reveals the formation mechanisms of flavors.

Methods

Strain culture and growth determination

The strain of *K. marxianus* L1-1 was previously screened and isolated in the traditional fermented rice-acid and could produce high concentration of aroma compounds, was used in the fermentation experiments. The strain *K. marxianus* L1-1 (CICC 33373) was grown at 30 °C for 48–72 h in YPD medium. The culture was stored at 4 °C and subcultured bi-weekly in order to maintain the viability. For a long-term storage, stock cultures were maintained in 30% glycerol at – 80 °C.

Preparation Methods Of Fermented Rice-acid

Rice-acid was fermented according to the following method. Firstly, the selenium rice (Danzhai, Guizhou, China) was broken with a high-speed pulverizer and sieved twice with an 80-mesh sieve to prepare rice flour. Then water was added into rice flour according to the proportions of 8.0% rice flour and 92.0% water and boiled under stirring conditions to obtain rice soup. Then, boiled rice soup was gelatinized for 30 min in a water bath at 60 °C under stirring conditions to prevent uneven gelatinization and local deterioration

of the gelatinization solution. Then 1.0% α -amylase was added into gelatinized rice soup for 30-min liquefaction at 90 °C. Then, 0.02% saccharifying enzyme (glucoamylase) was added for 2-h saccharification at 60 °C. Sterilize the mixture at 90 °C for 20 min. After the rice soup was cooled to about 30 °C, the concentration of inoculation of *K. marxianus* L1-1 was 1.0%. Then the prepared raw materials were poured into the sterilized rice-acid fermenter and sealed immediately. Finally, the mixture was fermented in a constant-temperature incubator at 30 °C for 4 d. We compared the different fermentation process (1 d and 3 d) in two rice-acid samples inoculated with *K. marxianus* L1-1, namely Y 1 d and Y 3 d.

Determination of the counts of *K. marxianus* L1-1

Samples (5 g) were mixed separately with 45 mL of 0.85% (w/v) NaCl solution and shaken at 150 rpm for 30 min at room temperature. Then, the supernatant was serially diluted by 10^5 times with 0.85% (w/v) NaCl solution and spread on YPD agar medium. *K. marxianus* L1-1 were cultured at 30 °C for 72 h, and the number of *K. marxianus* L1-1 was count at the 0 d, 1d, 2 d, 3 d and 4 d, respectively.

Determination Of Volatile Compounds

Through SPME-GC-MS analysis, volatile compounds were determined according with the method of Molyneux and Schieberle [41]. Retention times and mass spectral data were used to identify each compound. The retention times of the volatile compounds were determined using a C6–C26 alkane standard. The concentrations of volatile components were calculated with the peak areas of the internal standard (10 μ L of 2-methyl-3-heptanone, 10 mg/L). The mass spectra and retention indices were determined on at least two different GC columns that have stationary phases of different polarities and results were compared to spectra and retention indices.

Determination Of Organic Acids

After the settlement, the rice-acid samples inoculated with *K. marxianus* L1-1 were filtered with double-layer filter paper. The obtained filtrate was filtered through a 0.22- μ m microporous membrane and then passed through a ZORBAX SB-AQ solid phase cartridge for HPLC analysis (equipped with G1329B autosampler, G1311C quaternary low pressure ladder, G1316A column oven, and G1315D diode array UV-visible light detector) with ZORBAX SB-Aq column (4.6 \times 250 mm, 5 μ m, American Agilent Corporation) according to the following parameters: 0.02 mol/L NaH₂PO₄ solution as the mobile phase (pH 2.7), the injection volume (10 μ L), flow rate (0.9 mL/min), column temperature (35 °C), and detector wavenumber (UV 210 nm). Oxalic acid, tartaric acid, malic acid, acetic acid, citric acid, and succinic acid in rice-acid samples were determined. L-lactic acid concentration was measured with Amplite™ Colorimetric L-Lactate Assay Kit (AAT Bioquest Inc.USA).

Transcriptome Sequencing And Data Analysis

Three independent biological replicates of transcriptome sequencing of *K. marxianus* L1-1 were used in the samples Y 1 d and Y 3 d. The RNA extraction of *K. marxianus* L1-1 cells, libraries construction, RNA-

seq, and primary data analysis by PTM BIO Co. Ltd (Hangzhou, China). The *Kluyveromyces marxianus* dmku3 genome was downloaded from the EnsemblFungi website (<http://fungi.ensembl.org/index.html>) and used as reference genome. The filtered reads were aligned to the reference genome using TopHat (v2.0.12), and the novel transcripts were identified using both Cufflinks2.2.1 and TopHat2.0.12. The genes with fold change ≥ 2 and False Discovery Rate (FDR) < 0.05 in a comparison were considered as significant differentially expressed genes (DEGs). We used the enrichment analysis Gene Ontology (GO) functions and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways to compare the DEGs in the pairwise comparison (Y 3 d vs Y 1 d).

Proteomic Sequencing And Data Analysis

Similarly, three independent biological replicates of proteomic sequencing of *K. marxianus* L1-1 were used in the samples Y 1 d and Y 3 d. The preparation for proteomic analysis of *K. marxianus* L1-1 cell, liquid chromatography and mass spectrometry, peptide and protein identification and quantification according to the method by Xu et al [42]. All data with a 95% confidence and false discovery rate (FDR) less than 1% were considered to result in false positive results. According to the protein abundance level, the difference of more than 1.5-fold change (FC) and the statistical test of the p-value less than 0.05 were deemed to be differentially expressed proteins (DEPs) between Y 3 d and Y 1 d. All of DEPs were analyzed by GO and KEGG. The FASTA protein sequences of DEPs were blasted against KEGG database to retrieve their KEGG Orthologies (KOs) and were subsequently mapped to the pathways in KEGG. The corresponding KEGG pathways were extracted.

Correlation Analysis Between Proteomic And Transcriptomic Results

The DEGs and the DEPs were separately counted, and the Venn diagrams were plotted according to the counted results. Correlation analysis (Pearson correlation coefficient) was performed by Origin Pro 2018, and the four-quadrant maps were drawn based on changes in the transcriptome and proteome analysis.

Statistical analysis

All experiments were conducted in triplicate. Data were represented as the means \pm standard deviation. Duncan's multiple range test and t-test were carried out to analyze significant differences in SPSS version 20.0 (SPSS Inc., Chicago, IL, USA). whereby $P \leq 0.05$ or $P \leq 0.01$ were considered to be statistically significant.

Abbreviations

K. marxianus L1-1: *Kluyveromyces marxianus* L1-1; rice-acid: acid rice soup.

Declarations

Acknowledgments

We would like to thank MG Yang for their technical assistance.

Authors' contributions

NL conceived the study, screened and isolated the *Kluyveromyces marxianus* L1-1, carried out the fermentation experiments, performed the data analysis and wrote the paper. SM revised and edited the paper. LQ organized this research project and provided fund support. All authors read and approved the final manuscript.

Funding

We acknowledge the financial support of Technology platform and talent team plan of Guizhou, China ([2018]5251), Graduate Research Fund Project of Guizhou (YJSCXJH[2019]028). Industry-University-Research Cooperation Project of Guizhou (701/700465172217) and China Scholarship Council (201906670006).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

Author details

¹Key laboratory of Plant Resource Conservation and Germplasm Innovation in Mountainous Region (Ministry of Education), Collaborative Innovation Center for Mountain Ecology & Agro-Bioengineering (CICMEAB), College of Life Sciences/Institute of Agro-bioengineering, Guizhou University, Guiyang 550025, China. ²School of Liquor and Food Engineering, Guizhou University, Guiyang, 550025, China. ³Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland.

References

1. Nurcholis M, Murata M, Lertwattanasakul N, Kosaka T, Rodrussamee N, Limtong S, Yamada M. Characteristics of kanMX4-inserted mutants that exhibit 2-deoxyglucose resistance in thermotolerant yeast *Kluyveromyces marxianus*. *Open Biotechnol J.* 2016; 10(1).

2. Lertwattanasakul N, Rodrussamee N, Limtong S, Thanonkeo P, Kosaka T, Yamada M. Utilization capability of sucrose, raffinose and inulin and its less-sensitiveness to glucose repression in thermotolerant yeast *Kluyveromyces marxianus* DMKU 3-1042. *AMB Express*. 2011; 1(1): 20.
3. Inokuma K, Ishii J, Hara KY, Mochizuki M, Hasunuma T, Kondo A. Complete genome sequence of *Kluyveromyces marxianus* NBRC1777, a nonconventional thermotolerant yeast. *Genome Announc*. 2015; 3(2).
4. Zhao W, Shu Q, He G, Qihe C. Reducing antigenicity of bovine whey proteins by *Kluyveromyces marxianus* fermentation combined with ultrasound treatment. *Food Chem*. 2020; 311:125893.
5. Reyes-Sánchez FJ, Páez-Lerma JB, Rojas-Contreras JA, López-Miranda J, Soto-Cruz NÓ, Reinhart-Kirchmayr M. Study of the Enzymatic Capacity of *Kluyveromyces marxianus* for the Synthesis of Esters. *J Mol Microb Biotech*. 2020; 29 (1-6):1-9.
6. Posada JA, Patel AD, Roes A, Blok K, Faaij AP, Patel MK. Potential of bioethanol as a chemical building block for biorefineries: preliminary sustainability assessment of 12 bioethanol-based products. *Bioresou Technol*. 2013;135:490-499.
7. Park YC, Shaffer CEH, Bennett GN. Microbial formation of esters. *Appl Microbiol Biotechnol*. 2009; 85(1):13.
8. Villarreal-Morales SL, Muñiz-Márquez DB, Michel-Michel M, González-Montemayor ÁM, Escobedo-García S, Salas-Tovar JA, Rodríguez-Herrera R. Aguardiente a Fresh Beverage from Agave spp. Sap with Functional Properties. In *Natural Beverages*. 2019;(pp. 179-208). Academic Press.
9. Huang Q, Xu M, Zhang H, He D, Kong Y, Chen L, Song H. Transcriptome and proteome analyses of the molecular mechanisms associated with coix seed nutritional quality in the process of breeding. *Food Chem*. 2019; 272:549-558.
10. Zhou CY, Wang C, Tang CB, Dai C, Bai Y, Yu XB, Cao JX. Label-free proteomics reveals the mechanism of bitterness and adhesiveness in Jinhua ham. *Food Chem*. 2019; 297:125012.
11. Ma Q, Shi C, Su C, Liu Y. Complementary analyses of the transcriptome and iTRAQ proteome revealed mechanism of ethylene dependent salt response in bread wheat (*Triticum aestivum* L.). *Food Chem*. 2020;126866.
12. Nie L, Wu G, Culley DE, Scholten JC, Zhang W. Integrative analysis of transcriptomic and proteomic data: challenges, solutions and applications. *Crit Rev Biotechnol*.2007; 27(2):63-75.
13. Lertwattanasakul N, Kosaka T, Hosoyama A, Suzuki Y, Rodrussamee N, Matsutani M, Fujita N. Genetic basis of the highly efficient yeast *Kluyveromyces marxianus*: complete genome sequence and transcriptome analyses. *Biotechnol Biofuels*. 2015; 8(1):47.
14. Löser C, Urit T, Keil P, Bley T. Studies on the mechanism of synthesis of ethyl acetate in *Kluyveromyces marxianus* DSM 5422. *Appl Microbiol Biotechnol*.2015; 99(3):1131-1144.
15. Fan G, Teng C, Xu D, Fu Z, Liu P, Wu Q, Li X. Improving ethyl acetate production in Baijiu manufacture by *Wickerhamomyces anomalus* and *Saccharomyces cerevisiae* mixed culture fermentations. *BioMed Res Int*.2019; 2019.

16. Oloyede O, Franco J, Roos D, Athayde JRM, Boligon A. Antioxidative properties of ethyl acetate fraction of unripe pulp of *Carica papaya* in mice. *J Microbiol Biotechnol Food Sci.* 2020; 9(4): 409-425.
17. Rodrigues C, Vandenberghe LPS, Woiciechowski AL, de Oliveira J, Letti LAJ, Soccol CR. Production and application of lactic Acid. In *Current Developments in Biotechnology and Bioengineering.* 2017; (pp. 543-556). Elsevier.
18. de Oliveira RA, Komesu A, Rossell CEV, Maciel Filho R. Challenges and opportunities in lactic acid bioprocess design—From economic to production aspects. *Biochem Eng J.* 2018;133:219-239.
19. Kim SB, Kwon DH, Park JB, Ha SJ. Alleviation of catabolite repression in *Kluyveromyces marxianus*: the thermotolerant SBK1 mutant simultaneously coferments glucose and xylose. *Biotechnol Biofuels.* 2019; 12(1):90.
20. Lane S, Xu H, Oh EJ, Kim H, Lesmana A, Jeong D, Kim SR. Glucose repression can be alleviated by reducing glucose phosphorylation rate in *Saccharomyces cerevisiae*. *Sci Rep.* 2018; 8(1):1-12.
21. Zhou Y, Zhang W, Liu Z, Wang J, Yuan S. Purification, characterization and synergism in autolysis of a group of 1, 3- β -glucan hydrolases from the pilei of *Coprinopsis cinerea* fruiting bodies. *Microbiology.* 2015; 161(10):1978-1989.
22. Kang L, Zhang X, Wang R, Liu C, Yi L, Liu Z, Yuan S. β -Glucosidase BGL1 from *Coprinopsis cinerea* Exhibits a Distinctive Hydrolysis and Transglycosylation Activity for Application in the Production of 3-O- β -d-Gentiobiosyl-d-laminarioligosaccharides. *J Agric Food Chem.* 2019; 67(38):10744-10755.
23. Huang H, Shao X, Xie Y, Wang T, Zhang Y, Wang X, Deng X. An integrated genomic regulatory network of virulence-related transcriptional factors in *Pseudomonas aeruginosa*. *Nat Commun.* 2019;10(1):1-13.
24. Xu X, Zhou S, McClements DJ, Huang L, Meng L, Xia X, Dong M. Multistarter fermentation of glutinous rice with Fu brick tea: Effects on microbial, chemical, and volatile compositions. *Food Chem.* 2020;309:125790.
25. Selvaraju K, Gowsalya R, Vijayakumar R, Nachiappan V. MGL2/YMR210w encodes a monoacylglycerol lipase in *Saccharomyces cerevisiae*. *FEBS Lett.* 2016;590(8):1174-1186.
26. Löbs AK, Engel R, Schwartz C, Flores A, Wheeldon I. CRISPR–Cas9-enabled genetic disruptions for understanding ethanol and ethyl acetate biosynthesis in *Kluyveromyces marxianus*. *Biotechnol Biofuels.* 2017;10(1):164.
27. Verstrepen KJ, Van Laere SD, Vercammen J, Derdelinckx G, Dufour JP, Pretorius IS, Delvaux FR. The *Saccharomyces cerevisiae* alcohol acetyl transferase Atf1p is localized in lipid particles. *Yeast.* 2004;21(4):367-377.
28. Nancolas B, Bull ID, Stenner R, Dufour V, Curnow P. *Saccharomyces cerevisiae* Atf1p is an alcohol acetyltransferase and a thioesterase *in vitro*. *Yeast.* 2017;34(6): 239-251.
29. Marsafari M, Xu P. Debottlenecking mevalonate pathway for antimalarial drug precursor amorphadiene biosynthesis in *Yarrowia lipolytica*. *Metab Eng Commun.* 2020;10:e00121.

30. Kruis AJ, Levisson M, Mars AE, van der Ploeg M, Daza FG, Ellena V, Weusthuis RA. Ethyl acetate production by the elusive alcohol acetyltransferase from yeast. *Metab Eng.*2017;41:92-101.
31. Wheeler PR, Brosch R, Coldham NG, Inwald JK, Hewinson RG, Gordon SV. Functional analysis of a clonal deletion in an epidemic strain of *Mycobacterium bovis* reveals a role in lipid metabolism. *Microbiology.* 2008;154(12):3731-3742.
32. Parapouli M, Sfakianaki A, Monokrousos N, Perisynakis A, Hatziloukas E. Comparative transcriptional analysis of flavour-biosynthetic genes of a native *Saccharomyces cerevisiae* strain fermenting in its natural must environment, vs. a commercial strain and correlation of the genes' activities with the produced flavour compounds. *J Biol Res.*2019; 26(1): 5.
33. Das P, Manna I, Sil P, Bandyopadhyay M, Biswas AK. Exogenous silicon alters organic acid production and enzymatic activity of TCA cycle in two NaCl stressed indica rice cultivars. *Plant Physiol Biochem.*2019;136:76-91.
34. Lao-On U, Attwood PV, Jitrapakdee S. Roles of pyruvate carboxylase in human diseases: from diabetes to cancers and infection. *J Mol Med.* 2018;96(3-4): 237-247.
35. Suzuki T, Akiyama N, Yoshida A, Tomita T, Lassak K, Haurat MF, Nishiyama M. Biochemical characterization of archaeal homocitrate synthase from *Sulfolobus acidocaldarius*. *Febs Lett.*2020;594(1):126-134.
36. Scott EM, Pillus L. Homocitrate synthase connects amino acid metabolism to chromatin functions through Esa1 and DNA damage. *Genes Dev.*2010;24(17), 1903-1913.
37. Sienkiewicz-Porzucek A, Nunes-Nesi A, Sulpice R, Lisek J, Centeno DC, Carillo P, Fernie AR. Mild reductions in mitochondrial citrate synthase activity result in a compromised nitrate assimilation and reduced leaf pigmentation but have no effect on photosynthetic performance or growth. *Plant Physiol.*2008;147(1): 115-127.
38. Wischral D, Arias JM, Modesto LF, de França Passos D, Pereira Jr N. Lactic acid production from sugarcane bagasse hydrolysates by *Lactobacillus pentosus*: Integrating xylose and glucose fermentation. *Biotechnol Prog.*2019;35(1): e2718.
39. Frezza C, Zheng L, Folger O, Rajagopalan KN, MacKenzie ED, Jerby L, Kalna G. Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. *Nature*; 2011; 477(7363):225-228.
40. Kazi RS, Banarjee RM, Deshmukh AB, Patil GV, Jagadeeshaprasad MG, Kulkarni MJ. Glycation inhibitors extend yeast chronological lifespan by reducing advanced glycation end products and by back regulation of proteins involved in mitochondrial respiration. *J Proteomics.* 2017;156:104-112.
41. Molyneux RJ, Schieberle P. Compound identification: a journal of agricultural and food chemistry perspective. *J Agric Food Chem.* 2007;55:4625-4629.
42. Xu D, Behr J, Geißler AJ, Bechtner J, Ludwig C, Vogel RF. Label-free quantitative proteomic analysis reveals the lifestyle of *Lactobacillus hordei* in the presence of *Saccharomyces cerevisiae*. *Int J Food Microbiol.* 2019; 294: 18-26.

Figures

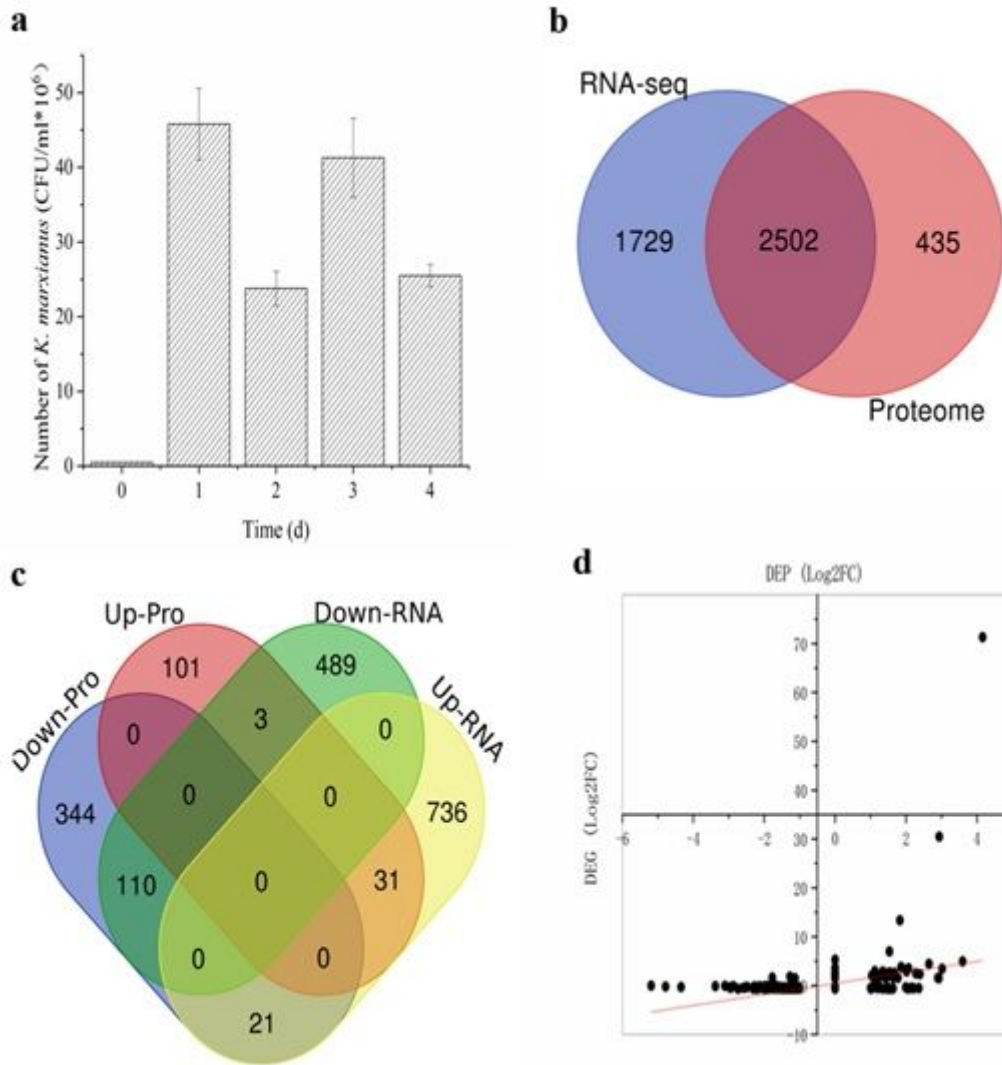


Figure 1

(a) The number of *K. marxianus* L1-1 during the rice-acid fermentation process (0 d, 1 d, 2 d, 3 d and 4 d). (b) Venn diagram of gene and protein in RNA-seq and proteome. (c) Venn diagram of up-RNA, down-RNA, up-protein and down-protein. (d) Scatter diagram of the correlation between the proteins and transcripts of *K. marxianus* L1-1 in 3-d and 1-d when rice-acid fermentation (Pearson correlation=0.32016, p-value \leq 0.01).

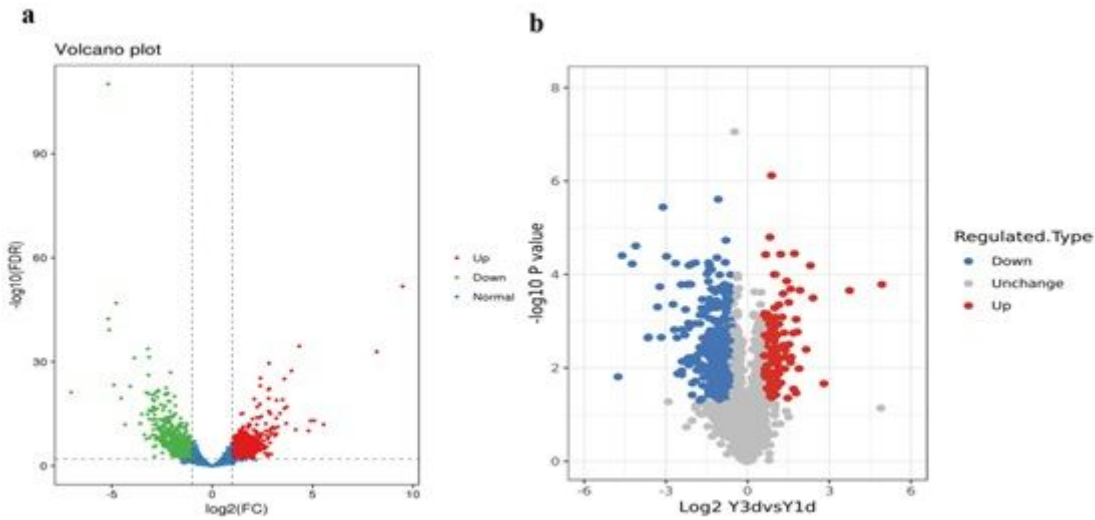


Figure 2

The volcano map of differentially expressed genes (a) and differentially expressed proteins (b) in *K. marxianus* L1-1 when inoculated rice-acid (Y 3 d vs Y 1 d).

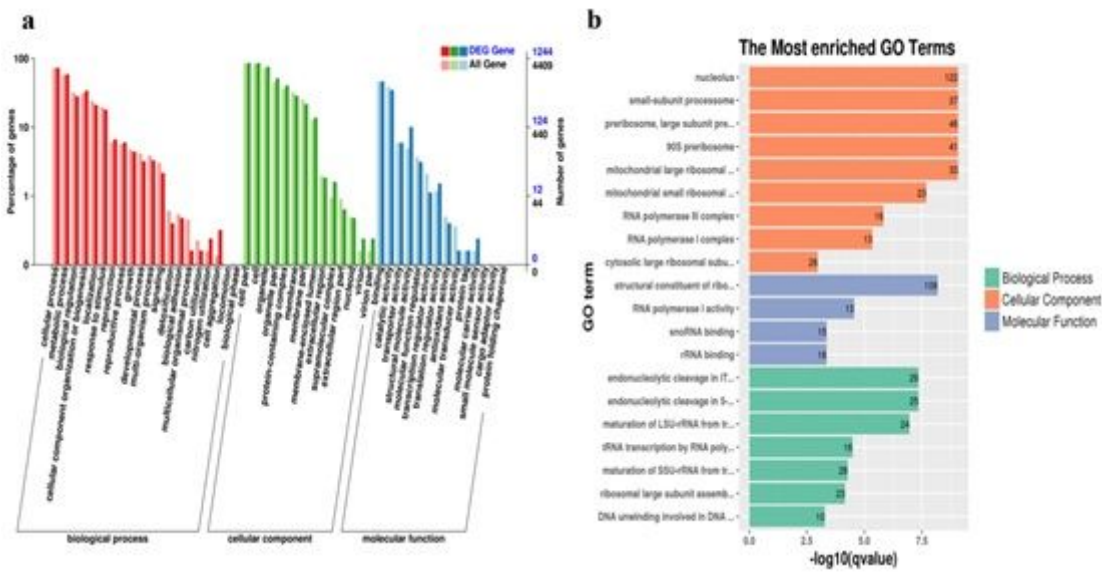


Figure 3

Statistical diagram of the second node annotation (a) and the most enriched GO Terms (b) of the differentially expressed genes in *K. marxianus* L1-1 when inoculated rice-acid (Y 3 d vs Y 1 d).

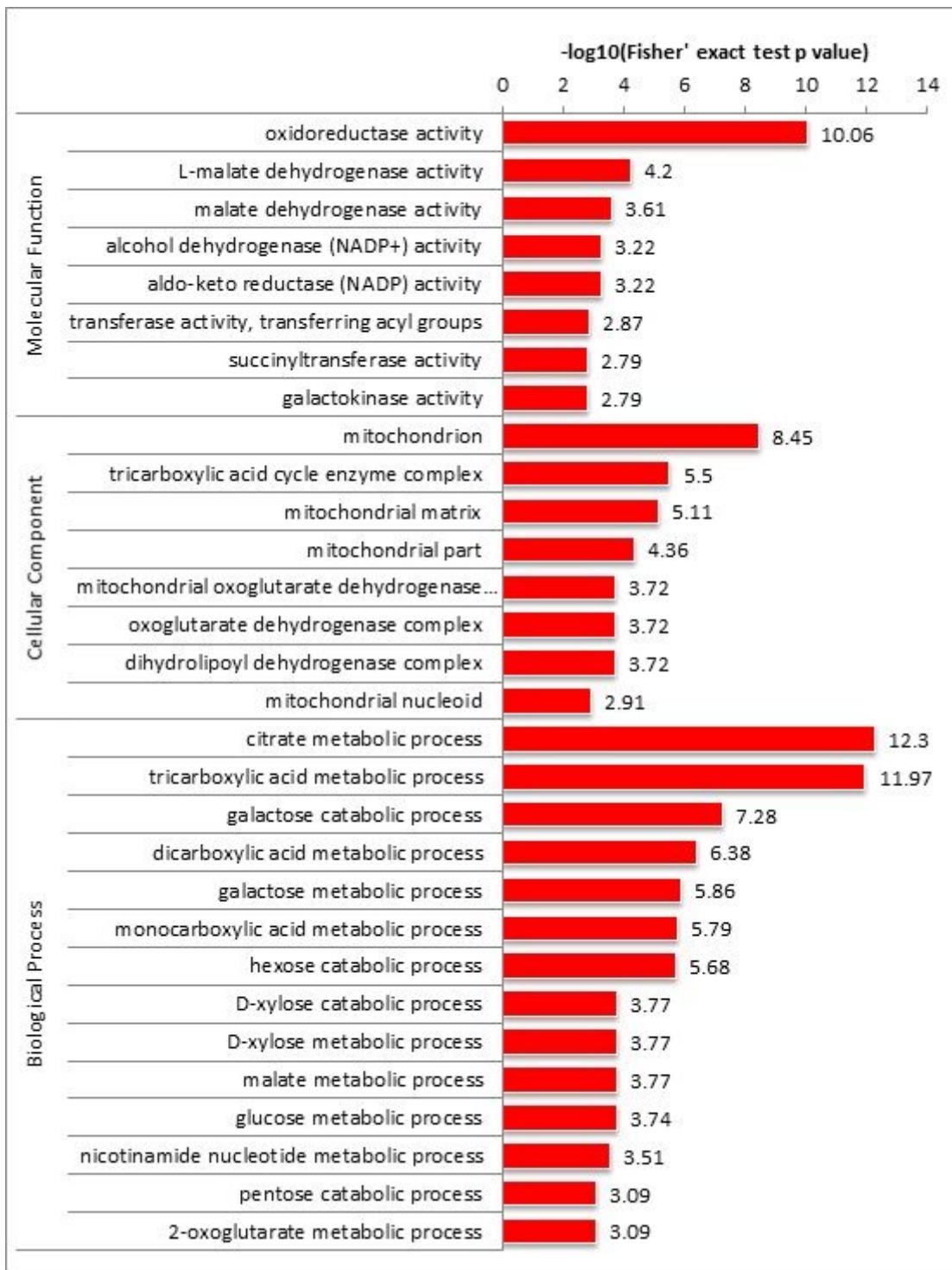


Figure 4

The GO analysis (molecular function, cellular component and biological process) of differentially expressed proteins in *K. marxianus* L1-1 when inoculated rice-acid (Y 3 d vs Y 1 d).

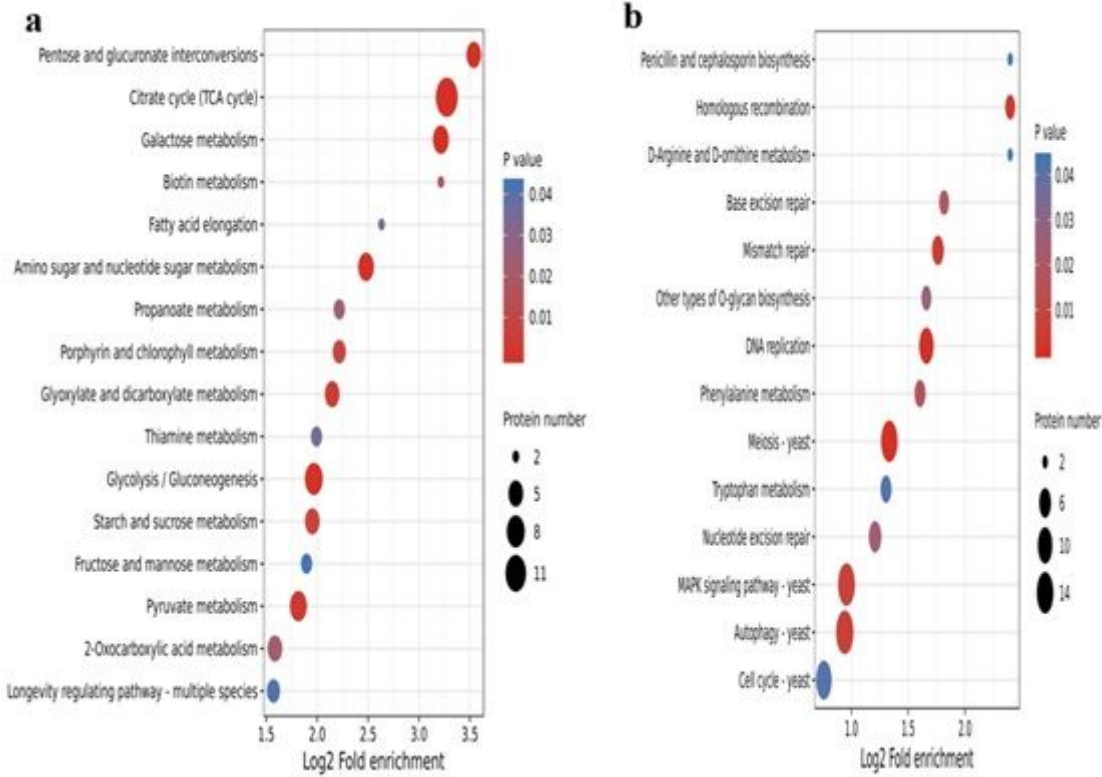


Figure 5

The top 16 up-regulated KEGG pathways (a) and 14 down-regulated KEGG pathways (b) enriched scatter plot of DEPs in *K. marxianus* L1-1 which inoculated in rice-acid (Y 3 d vs Y 1 d). Rich factor means the ratio of the protein number enriched in the pathway to the number of annotated proteins. The bigger the Rich factor, the more significant the enrichment was. The Q-value was the corrected p-value after multiple hypotheses testing, which was ranged from 0 to 1. The closer to zero, the more significant the enrichment was (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this study).

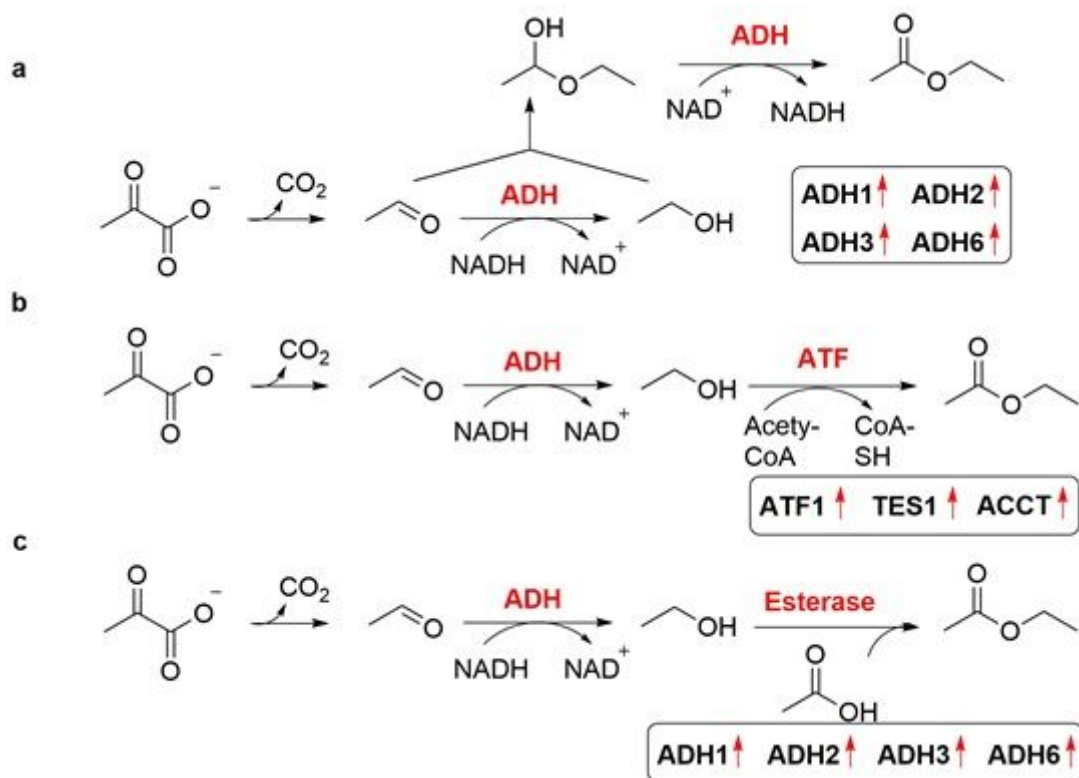


Figure 6

Ethyl acetate biosynthetic pathways and synthesis activities in *K. marxianus* L1-1 when inoculated rice-acid. (a) Alcohol dehydrogenase (ADH) oxidation of hemiacetal; (b) Ethyl acetate biosynthesis via alcohol-O-acetyltransferase (ATF); (c) Reverse esterase activity.

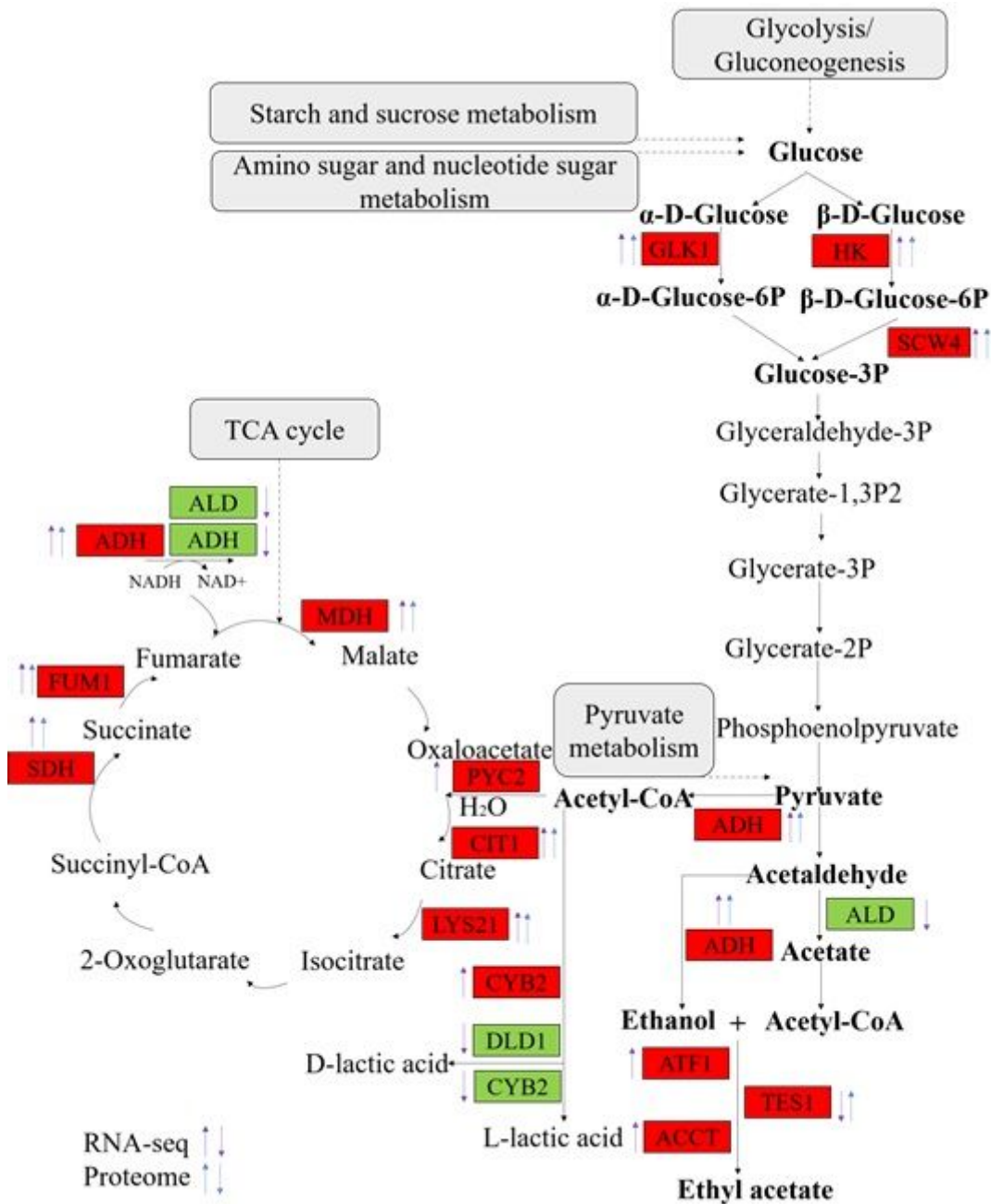


Figure 7

The metabolic pathway of synthesis of ethyl acetate and organic acids in *K. marxianus* L1-1 when inoculated in rice-acid (Y 3 d vs Y 1 d). Red square means the up-regulate gene and protein, green square means the down-regulate gene and protein, purple up arrow means the up-regulate gene, purple down arrow means the down-regulate gene, blue up arrow means the up-regulate protein, blue down arrow means the down-regulate protein.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.TranscriptomeKEGG.enrich.xls](#)
- [TableS2.ProteomeKEGGpathwayenrichment.xlsx](#)