Characterization of JEN family carboxylate transporters from the acid-tolerant yeast *Pichia kudriavzevii* and their applications in succinic acid production

Yongyan Xi,† Tao Zhan,‡ Hongtao Xu,† Jing Chen,† Changhao Bi,† Feiyu Fan,‡ and Xueli Zhang,†

1Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 West 7th Ave, Tianjin Airport Economic Park, Tianjin, 300308, China.
2Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin, China.
3University of Chinese Academy of Sciences, Beijing, China.
4National Technology Innovation Center of Synthetic Biology, Tianjin, China.

Summary

The unconventional yeast *Pichia kudriavzevii* is renowned for its ability to survive at low pH and has been exploited for the industrial production of various organic acids, especially succinic acid (SA). However, *P. kudriavzevii* can also utilize the di- and tricarboxylate intermediates of the Krebs cycle as the sole carbon sources for cell growth, which may adversely affect the extracellular accumulation of SA. Because the carboxylic acid transport machinery of *P. kudriavzevii* remains poorly understood, here, we focused on studying its SA transportation process from the perspective of mining and characterization of dicarboxylate transporters in a newly isolated acid-tolerant *P. kudriavzevii* strain CY902. Through genome sequencing and transcriptome analysis, two JEN family carboxylate transporters (PkJEN2-1 and PkJEN2-2) were found to be involved in SA transport. Substrate specificity analysis revealed that both PkJEN proteins are active dicarboxylate transporters, that can effectively import succinate, fumarate and L-malate into the cell. In addition, PkJEN2-1 can transport α-ketoglutarate, while PkJEN2-2 cannot. Since PkJEN2-1 shows higher transcript abundance than PkJEN2-2, its role in dicarboxylate transport is more important than PkJEN2-2. In addition, PkJEN2-2 is also responsible for the uptake of citrate. To our best knowledge, this is the first study to show that a JEN subfamily transporter is involved in tricarboxylate transport in yeast. A combination of model-based structure analysis and rational mutagenesis further proved that amino acid residues 392-403 of the tenth transmembrane span (TMS-X) of PkJEN2-2 play an important role in determining the specificity of the tricarboxylate substrate. Moreover, these two PkJEN transporters only exhibited inward transport activity for SA, and simultaneous inactivation of both PkJEN transporters reduced the SA influx, resulting in enhanced extracellular accumulation of SA in the late stage of fermentation. This work provides useful information on the mechanism of di-/tricarboxylic acid utilization in *P. kudriavzevii*, which will help improve the organic acid production performance of this microbial chassis.

Introduction

Succinic acid (SA) is an important platform C4 chemical with immense commercial value and a wide range of applications in the chemical, food and pharmaceutical industries (Jansen and van Gulik, 2014). SA is not only an important building block for the biodegradable plastic polybutylene succinate (PBS), but can also be catalytically converted into a variety of high-value products, such as 1,4-butanediol (1,4-BDO), tetrahydrofuran, γ-butyrolactone and 2-pyrrolidone (Ahn et al., 2016). The global SA market is projected to grow at a Compound Annual Growth Rate (CAGR) of around 27.4% to reach approximately USD 1.8 billion by 2025 (Nghiem, 2017). Compared with traditional petroleum-based approaches, microbial fermentation of renewable carbon sources to produce SA is more economical and has greater potential for sustainable development in the future. Currently, the industrial microorganisms most commonly used to produce bio-based SA are bacteria such as *Escherichia coli*.
coli, Corynebacterium glutamicum, Actinobacillus succinogenes and Mannheimia succiniciproducens (Beau-prez et al., 2010). However, the acid resistance of bacterial cells is generally poor, leading to the need for neutralizing reagents to maintain growth. In fact, the production of 1 metric ton of SA requires about half a metric ton of alkali (KOH), after which 1 metric ton of sulfuric acid will be consumed to obtain free acid. Therefore, the downstream purification cost accounts for 60–70% of the total production cost (Yuzbashev et al., 2011; Chen and Patel, 2012). By contrast, yeast cells have excellent acid tolerance and can maintain normal growth at pH values below 3. Under these conditions, SA is predominantly in its undissociated form, which greatly simplifies the downstream processing. The unconventional yeast Pichia kudriavzevii (syn. Issatchenka orientalis or Candida krusei) can withstand multiple environmental stresses, including low pH, and has been developed as a suitable host for the production of organic acids (Toivari et al., 2013; Park et al., 2018). The BioAmber company has begun to use P. kudriavzevii to produce bio-based SA and has been operating a plant with a capacity of 30,000 metric tons since 2014 (Ferone et al., 2019). A patent held by Myriant shows that this company has also studied acid-resistant yeast since 2013. The P. kudriavzevii strain SD108 screened by the company can grow in medium containing 6% w/v SA (pH 3.0), reaching a biomass that is 7.5 times higher than that of Saccharomyces cerevisiae BY4742 under the same culture conditions (Yocum et al., 2013).

The relatively low production titre of SA is one of the main problems faced by low-pH yeast fermentation technology (Ferone et al., 2019), which may be partly attributed to the transport characteristics of weak carboxylic acids. As a typical carboxylic acid, SA can partially dissociate in aqueous solution, and a chemical equilibrium will be established between the undissociated molecules and dissociated carboxylic anions following the Henderson–Hasselbalch equation \( [\text{pH} = \text{pK}_a + \log (A^-/HA]) \) (Casal et al., 2008, 2016). In solutions with a pH value lower than \( \text{pK}_{a1} \) (SA: \( \text{pK}_{a1} \approx 4.21 \) and \( \text{pK}_{a2} \approx 5.64 \)), the undissociated SA molecules dominate in the solution. However, the undissociated acid form is electrically neutral and thus can cross the cell membrane through simple passive diffusion or permease/channel-mediated facilitative diffusion. At the same time, the partially dissociated carboxyl anions can also be transported by proton symporters, which depend not only on the transmembrane gradient of the substrate but also on the \( \Delta \text{pH} \) component of the proton motive force (Casal et al., 2008). Once inside the cell, the high cytoplasmic pH of P. kudriavzevii of about 8.0 (Halm et al., 2004) will cause the rapid dissociation of SA, thereby releasing protons and carboxylic anions in the cytoplasm. Since the charged protons and carboxylic anions cannot exit the cell by passive diffusion, they will cause cytotoxicity if they cannot be rapidly metabolized or pumped out of the cell. It has been reported that the accumulation of SA leads to oxidative stress and elevated mitochondrial reactive oxygen species (ROS) levels (Zhang et al., 2020). In addition, cytoplasmic acidification caused by proton accumulation can significantly reduce the activities of several glycolytic enzymes, thereby hindering the normal metabolism and utilization of carbohydrates (Pampulha and Loureirodias, 1990). Therefore, cells must pump out excess protons and carboxyl anions using specific proton pumps (P-type H⁺-ATPase) and ATP-Binding Cassette (ABC) family transporters located in the cell membrane to maintain a normal physiological state. These processes are energy-dependent and consume ATP. The exported protons and anions again form undissociated molecules, in the low-pH environment outside the cell, and can re-enter the cell to form an energy draining futile cycle. This process is believed to eventually waste a considerable portion of carbon source molecules for energy generation, thereby reducing SA synthesis. One possible strategy to block the futile cycle is to reduce the uptake of extracellular SA by yeast cells as much as possible. However, there are few reports on the mechanism of carboxylate transport in P. kudriavzevii, and the potential transporter involved in SA metabolism remains to be identified and functionally characterized.

In this study, we systematically studied the SA transport mechanism in the newly isolated acid-tolerant P. kudriavzevii strain CY902. We first obtained a high-quality genome sequence using Oxford Nanopore MinION and Illumina HiSeq sequencing, after which we fully annotated all putative transporter genes in the CY902 genome. We further compared the transcriptome profiles between cells growing on SA and glucose and identified two JEN family transporters with potential SA transport activity. Through gene knockout, overexpression and rational protein engineering, we found that these two proteins are responsible for the transport of di/tricarboxylic acids, but have different expression patterns and substrate specificities. Finally, we evaluated the impact of the two JEN family transporters on SA production.

**Results**

*Isolation and genome sequencing of acid-tolerant P. kudriavzevii CY902*

By screening acid-tolerant yeasts isolated from the peel of wild fruits under low-pH culture conditions containing different inorganic and organic acids, we found that one of them, named CY902, showed the best growth performance under all these low-pH culture conditions (Fig. 1A), and was therefore selected for further study. Based on the D1/D2 region of the 26S rDNA (data not available).
shown), the strain CY902 was identified as *P. kudriavzevii*. To better understand the genetic features of this acid-tolerant yeast, we sequenced the complete genome of CY902. A total of 1 158 003 Oxford Nanopore ultra-long reads were generated using the MinION™ platform, which is equivalent to 33.29 Gb of clean data (≈3000 × genome coverage). The N50 value of the reads was 34 934 bp, and the longest read was 523 434 bp (Fig. S1). A preliminary de novo assembly based on these long reads yielded a total of 63 contigs. Then, we used 15 772 689 clean Illumina short reads (3.943 Gb, ~358 × genome coverage) for genome polishing and error correction, and finally merged the 63 contigs into a 10.8 Mb nuclear genome containing five chromosomes (Fig. 1B) and a circle mitochondrial genome of 51 kb. When the Illumina data were realigned using BWA, over 99.61% of reads were successfully mapped to the CY902 genome, indicating that the CY902 genome assembly has good structural integrity and accuracy. According to the mapping results, a total of 30 127 heterozygous single-nucleotide variations (SNVs) were detected. Most of our calling SNVs had an allele frequency of 0.5, indicating that the CY902 strain was diploid. Benefitting from the oxford nanopore long reads, we obtained more complete rDNA and telomere structures at the ends of chromosomes compared with the previous assembly based on the PacBio strategy (Douglass et al., 2018). We found that the rDNA unit is a sequence of about 11 800 bp that repeats at least seven times on the left end of chromosome 1 (1L), and at least twice on the right ends of chromosomes 1(1R), 2(2R), and 3(3R), but was not found at the ends of other chromosomes. Each rDNA unit contained 18S, 5.8S and 26S rRNA genes transcribed towards the end of the chromosome, as well as a 5S gene in the opposite orientation (Fig. S2). Consistent with a previous report (Douglass et al., 2018), telomere repeats (TTACAATATGAACTAGGAGCGAGGTGTG) were only found on chromosomes 2L, 3L, 4L, 4R, 5L and 5R. Based on extrinsic evidence obtained from strand-specific RNA-Seq and protein homology information, we predicted a total of 5464 protein-coding genes in the CY902 nuclear genome (Table S1). Among them, 4185 genes (76.59%) have homologues in *S. cerevisiae*, 431 genes (7.88%) were assigned to putative functions using the Pfam and Swiss-Prot databases, while 848 genes (15.51%) without any information were annotated as hypothetical genes (Table S1).

**RNA-Seq analysis revealed potential SA transporters in CY902**

To explore the mechanism of carboxylate transport in *P. kudriavzevii*, we first systematically annotated and classified all transporters in CY902. A total of 341 putative transporters were annotated by searching against the

---

**Fig. 1.** Isolation and genome sequencing of acid-tolerant *P. kudriavzevii* CY902.  
A. The acid-resistant phenotype of CY902 cells grown in media containing different inorganic or organic acids. The diploid *S. cerevisiae* BY4743 strain was used as a control.  
B. The chromosome structure of CY902. The organization of the five chromosomes was plotted using Integrative Genomics Viewer (IGV, http://www.igv.org/). Genomic features such as genes, tRNAs and rRNAs are shown in blue, and the centromeric region of each chromosome is shown in a red box.
SGD, Swiss-Prot and Pfam databases using the BLASTP algorithm (Table S2). Among the 341 putative transporters, 327 were found to have homologues in S. cerevisiae, 11 were homologous to the proteins found in the Swiss-Prot database, and only three were predicted at the protein domain (Pfam) level. According to the Transport Classification (TC) system of the TransportDB database, all putative transporters were classified into four families (ATP-dependent transporters, ion channels, outer membrane porins and secondary transporters), and 13 transporters were found to be involved in carboxylic acid transport, including eight MCH-like transporters, two JEN family transporters and three ESBP6-like transporters (Table S2). The RNA-sequencing data were used to profile the expression of 341 putative transporter genes in SA-based culture. We then performed a comparative analysis of the transcriptional profiling data of a ΔURA3 mutant grown either in mixed carbon source comprising 2% w/v glucose and 2% w/v SA (referred to as GS), 2% w/v SA as the sole carbon source (SoleS) or 2% w/v glucose as the sole carbon source (SoleG). Among the 341 putative transporter genes, 43 (12.6%) showed significantly altered their expression levels between the GS and SoleG samples (7 up- and 37 downregulated). However, the 13 putative carboxylic acid transporter genes were not presented in this DEG group and exhibited low mRNA abundance with normalized expression value (FPKM) of less than 100 (Fig. 2; Table S2). A different situation was observed in the SoleS sample, where a total of 149 genes (43.7%) showed differential expression (74 up- and 75 downregulated) compared with the SoleG sample. Among the 74 upregulated genes, the mRNA levels of four putative carboxylic acid transporter genes increased 23-fold, fivefold, fourfold and threefold after SA induction respectively (Fig. 2). These included two JEN-like genes (gene identifiers PK0403 and PK0627) and two ESBP6-like genes (PK0737 and PK0738).

![Fig. 2. RNA-Seq analysis revealing potential SA transporters in CY902. The expression patterns of 341 putative transporter genes were plotted by calculating the log2 normalized fold changes in ΔURA3 cells grown on glucose/SA (GS) or SA (SoleS) versus the glucose (SoleG) control. Each point represents a putative transporter transcript. Abs was defined as the absolute difference of the normalized transcript abundance (FPKM) between the glucose control and SA cultures. The dimensions of the points were adjusted in relation to the scale of the Abs value. The transcripts for JEN family transporters are indicated by blue points, while other putative carboxylic acid transporters are marked in red. The black dotted line represents a manually defined threshold (log2 ratio = 1).](image-url)
Considering the low expression levels of the two ESBP6-like genes (FPKM value of less than 15), the two JEN-like genes remained as the only bona fide SA-induced genes. The average FPKM value of PK0403 reached 1975 upon SA induction, which was nearly 3.6 times that of PK0627 (Table S2). PK0403 encodes a protein of 539aa with 67.0% identity and 76.3% similarity with PK0627 protein (523aa). Protein BLAST search in the Swiss-Prot database showed that PK0403 and PK0627 proteins, respectively, shared 100% sequence identity to carboxylic acid transporter proteins A0A099NZE8 and A0A099P158 of *P. kudriavzevii*. Both PK0403 and PK0627 contain the conserved signature sequence of the JEN family proteins (NXXS/THXS/TQDXXXT) (Soares-Silva *et al.*, 2007), except that the first amino acid residue of PK0627 protein is not a conserved asparagine residue but a serine residue (Fig. 3A). Phylogenetic analysis (Fig. 3B) revealed that JEN proteins from yeast and fungal species can be divided into four major clades, corresponding to the JEN1 subfamily (Clade A), JEN2 subfamily (Clade B), *Yarrowia lipolytica* specific JEN3 subfamily (Clade C) and other fungal JEN proteins (Clade D). Since both PK0403 and PK0627 proteins clustered closely with JEN2 homologues in Clade B, therefore, we named them as *PkJEN2-1* and *PkJEN2-2* respectively.

### Substrate specificities of PkJEN2-1 and PkJEN2-2

To assess the possible roles of the *PkJEN2-1* and *PkJEN2-2* genes in carboxylate transport, single-gene (Δ*PkJEN2-1* and Δ*PkJEN2-2*) and double-gene (ΔΔ*PkJEN2*) null mutants were generated using CRISPR/Cas9-aided homologous recombination. To investigate the substrate specificity of *PkJEN2-1* and *PkJEN2-2*, their null mutants and the parental ΔURA3 strain were grown in media containing monocarboxylate (L-lactate or pyruvate), dicarboxylate (L-malate, fumarate, succinate or α-ketoglutarate), tricarboxylate (citrate or DL-isocitrate), or glucose as sole carbon and energy source (Fig. 4A) respectively. The pH of the medium was adjusted to 5.0 (Queiros *et al.*, 2007) or 7.0 (Vieira *et al.*, 2010; Guo *et al.*, 2015) as described before. In addition, we also adjusted the medium pH to 2.5 to test the effect of low pH on the carboxylate transport of PkJEN2 proteins. In addition to DL-isocitrate (data not shown), CY902 can utilize other carboxylates mentioned above as the sole carbon source for growth (Fig. 4A). We found that inactivating either or both PkJEN2-1 and PkJEN2-2 in CY902 did not affect cell growth in L-lactate and pyruvate medium at pH 2.5 and 5.0 (Fig. 4A), suggesting that JEN family transporters do not affect monocarboxylic acid transport in *P. kudriavzevii*.
Interestingly, we observed that at pH 7.0, CY902 cells could grow normally in L-lactate medium, but could hardly grow in pyruvate medium (Fig. 4A and B), suggesting that different transport systems were used for L-lactate and pyruvate transport in CY902. Distinct phenotypes were observed when the null mutants were grown in dicarboxylate media at different pH values: At pH 2.5, the growth of ΔPkJEN2-1 in the medium containing succinate or α-ketoglutarate as the sole carbon source was strongly reduced (Fig. 4A). The consumption of succinate and α-ketoglutarate in ΔPkJEN2-1 was strongly reduced by ~72 and 96% compared to the parental ΔURA3 strain after 24 h of cultivation, and the corresponding biomass (OD$_{600}$) decreased by ~73% and 94% respectively (Fig. 4B). However, deletion of PkJEN2-1 did not significantly affect the utilization of L-malate and fumarate at pH 2.5 (Fig. 4A and B). Under four kinds of dicarboxylate growth conditions, ΔPkJEN2-2 did not show a statistically significant difference in growth compared with the control (Fig. 4A and B). Interestingly, we found that deletion of both PkJEN2 genes simultaneously leads to almost lost growth on L-malate, fumarate or α-ketoglutarate, and showed a significant decrease in growth on succinate at pH 2.5 (Fig. 4A). The consumption of l-malate, fumarate, α-ketoglutarate and succinate in ΔΔPkJEN2 was strongly reduced by ~91%, 96%, 97% and 78% compared to the parental ΔURA3 strain after 24 h of cultivation, and the corresponding biomass (OD$_{600}$) decreased by ~92%, 95%, 96% and 87% respectively (Fig. 4B). Different from the phenotype observed at pH 2.5, the growth of ΔPkJEN2-1 in the medium containing l-malate and fumarate as the sole carbon source was also strongly reduced at pH 5.0 (Fig. 4A). The consumption of l-malate, fumarate, succinate and α-ketoglutarate in ΔPkJEN2-1 was strongly reduced by ~77, 94, 83 and 94% compared to the parental ΔURA3 strain after 24 h of cultivation, and the corresponding biomass (OD$_{600}$) decreased by ~72%, 88%, 80% and 95% respectively (Fig. 4B). However, ΔPkJEN2-2 did not show a statistically significant difference in growth compared with the control under the same growth conditions (Fig. 4A and B). At pH 5.0, ΔΔPkJEN2 strain showed a similar growth phenotype when grown on dicarboxylate as that under the condition of pH 2.5 (Fig. 4A). The consumption of l-malate, fumarate, succinate and α-ketoglutarate by ΔΔPkJEN2, respectively, decreased to ~6%, 5%, 7% and 6% of the ΔURA3 strain and only a very small amount of biomass (OD$_{600}$ < 0.4) was visible after 24 h of cultivation (Fig. 4A and B). These results suggest that both PkJEN proteins play a role in dicarboxylate transport in P. kudriavzevii, whereby PkJEN2-1 is more important. Interestingly, all null mutants and the controls exhibited a longer lag phase when grown on L-malate, fumarate, succinate or α-ketoglutarate at pH 7.0 (Fig. 4A and B), suggesting that the transport of dicarboxylate through PkJEN2-1 and PkJEN2-2 is pH-dependent. Moreover, we found that ΔPkJEN2-2 and ΔΔPkJEN2 could hardly grow in citrate medium at pH 5.0, while ΔΔPkJEN2-1 showed normal growth under the same conditions (Fig. 4A and B), suggesting that PkJEN2-2 has broader substrate specificity than PkJEN2-1, including tricarboxylate transport activity. Accordingly, all null mutants and the control showed growth retardation when utilizing citrate at pH 7.0. Considering that there is a significant difference in the mRNA abundance of PkJEN2-1 and PkJEN2-2 upon SA induction, both genes were overexpressed in ΔΔPkJEN2 under the control of the well-established TDH3 constitutive promoter and GAL2 terminator (Cao et al., 2020) to compare their carbohydrate transport efficiency. Different from the growth phenotype of single-gene null mutants, the PkJEN gene overexpression strains OV$_{PkJEN2-1}$-RI and OV$_{PkJEN2-2}$-RI did not exhibit a noticeable difference in carbohydrate consumption and biomass accumulation when grown on L-malate, fumarate or succinate (Fig. 4A and B), suggesting that PkJEN2-1 and PkJEN2-2 may have comparable transport efficiency for these dicarboxylates. In addition, overexpression of the PkJEN2-1 or PkJEN2-2 gene did not improve the growth of ΔΔPkJEN2 on citrate or α-ketoglutarate respectively. These results indicate that the PkJEN2-1 transporter does not have tricarboxylate transport activity, while PkJEN2-2 transporter can transport citrate but not α-ketoglutarate.

The structural basis of substrate specificity

To decipher the mechanisms determining the substrate specificity of PkJEN transporters, homology models of S. cerevisiae JEN1 (ScJEN1), as well as P. kudriavzevii JEN2-1 and JEN2-2, were constructed using the I-TASSER server. The crystal structure of the high-affinity phosphate transporter PiPT from the fungus Piriformospora indica (PDB ID:4J05) was used as template for modelling because it exhibits the highest structural similarity score (TM-score) (Roy et al., 2010) with all three JEN family transporters of P. kudriavzevii. The TM-scores of the modelled ScJEN1, PkJEN2-1 and PkJEN2-2 were 0.68, 0.71 and 0.73, respectively, and a TM-score > 0.5 indicates a model of correct topology (Roy et al., 2010). Parts of the N-terminal (residues 1-143 of ScJEN1, 1-86 of PkJEN2-1 and 1-71 of PkJEN2-2) and C-terminal sequences (residues 574-616 of ScJEN1, 520-539 of PkJEN2-1 and 478-523 of PkJEN2-2) were truncated from the three initial models because they could not be successfully folded into reliable structures (Fig. 5A). Then, the remaining parts were subjected to multiple rounds of refinement using the
GalaxyRefine program until structure had a Ramachandran score of > 90% of the amino acids in the favoured region (Ko et al., 2012). The Rama-favoured score of the refined models of ScJEN1, PkJEN2-1 and PkJEN2-2 was 90.1, 91.0 and 91.6 respectively. The final adopted models of the three proteins were all composed of 12 putative α-helical transmembrane spans (TMS) (Fig. 5A) and adopted the canonical Major Facilitator Superfamily (MFS) fold, in which the N-terminal (TMS I-VI) and C-terminal (TMS VII-XII) six-helix bundles are gathered together in the membrane to form a 'V' shaped transporter (Marger and Saier, 1993) (Fig. 5E). Previous
Effect of PkJEN transporters on SA production

Inactivation of the mitochondrial succinate dehydrogenase complex (SDH) has been used as a simple strategy to achieve intracellular accumulation of SA in Y. lipolytica (Gao et al., 2016). To achieve the initial production of SA in CY902, we followed the same strategy and disrupted the SDH complex subunit gene SDH5 in the parental ΔURA3 strain. The resulting ΔURA3ΔSDH5 strain was able to grow in YPD medium, but exhibited slower growth and lower biomass accumulation than the parental strain. When 5% w/v glucose was used as the initial substrate to produce SA in a shake flask culture, all glucose was completely consumed by both ΔURA3ΔSDH5 and ΔURA3 within 48 hours. When glucose was exhausted (48 h), the OD600 value of ΔURA3ΔSDH5 was about 5.3, approximately 83% of the corresponding value achieved by ΔURA3. The SA titre of ΔURA3ΔSDH5 reached its highest value of ~3.2 g l−1 at 36 h. At this time, the residual glucose concentration fell below 2 g l−1, and ~12.0 g l−1 ethanol was produced as the main by-product. The amount of other minor by-product such as glycerol and acetic acid are very low, which can be ignored. By contrast, the parental ΔURA3 strain practically did not produce any SA after 36 h of cultivation. To explore whether the inactivation of PkJEN transporters increased SA accumulation, the SDH5 gene was further deleted in the ΔΔPKJEN2 strain. In YPD culture without pH control, the ΔΔPKJEN2ΔSDH5 mutant showed a higher glucose consumption rate at both 12 and 24 h (1.20 and 1.58 g l−1 h−1), corresponding to ~33% and 18% increases over ΔURA3ΔSDH5 respectively (Fig. 7A). In addition, ΔΔPKJEN2ΔSDH5 produced SA titres of 3.2 and 3.6 g l−1 at 24 and 36 h, which were 10% and 12% higher than that of ΔURA3ΔSDH5 respectively. When glucose was depleted (48 h), the SA titre of ΔURA3ΔSDH5 decreased rapidly, while no obvious decrease of the SA titre was observed in ΔΔPKJEN2ΔSDH5. Notably, no significant differences in biomass accumulation were observed between the two strains during the entire fermentation period (Fig. 7A). To test whether the PkJEN proteins are capable of bidirectional SA transport, we further disrupted the SDH5 gene in the PkJEN overexpression mutants to obtain the strains OV_pkJen2−1ΔSDH5 and OV_pkJen2−2ΔSDH5 respectively. Both strains showed similar cell growth and glucose consumption compared to the control strain ΔΔPKJEN2ADE2ΔSDH5. However, the SA production titre of the OV_pkJen2−1ΔSDH5 and OV_pkJen2−2ΔSDH5 strain significantly decreased from the early stage of fermentation. After 24 h of cultivation, the SA titres of OV_pkJen2−1ΔSDH5 and OV_pkJen2−2ΔSDH5 reached 2.2 and 2.3 g l−1, which was, respectively, ~27 and 24% lower than that of the control. A similar
decrease in SA titre was also observed at 36 and 48 h (Fig. 7B), indicating that PkJEN transporters mediate inward transport of carboxylate rather than outward secretion.

Discussion

The acid-tolerant yeast *P. kudriavzevii* can not only grow with common carbon sources such as glucose, ethanol, glycerol or fructose, but can also utilize the di- and tricarboxylate intermediates of the Krebs cycle as the sole carbon sources for cell growth (Suthers et al., 2020). Although *P. kudriavzevii* has been developed for production of organic acids, especially SA, the mechanism of dicarboxylate transport of this yeast is still unclear. The main goal of this study was to identify key dicarboxylate transporters involved in SA transport in *P. kudriavzevii* strain CY902. Comparative transcriptomic analysis of CY902 cells grown on SA and glucose revealed that two JEN family transporters (PkJEN2-1 and PkJEN2-2) are responsible for the inward transport of SA. The JEN family transporters of yeast, which are known to be involved in the specific uptake of monocarboxylate (JEN1 subfamily) and dicarboxylate (JEN2 subfamily), have been intensively studied in several species include *S. cerevisiae* (Casal et al., 1999; Soares-Silva et al., 2003), *Kluyveromyces lactis* (Lodi et al., 2004; Queiros et al., 2007) and *Candida albicans* (Soares-Silva et al., 2004; Vieira et al., 2010). In contrast to all the above-mentioned species which contain JEN1 homologues, both JEN transporters in *P. kudriavzevii* belong to the JEN2 subfamily, since simultaneously inactivation of PkJEN2-1 and PkJEN2-2 did not impair the growth of CY902 on L-lactate or pyruvate. However, the double-knockout cells were no longer able to grow on L-malate, fumarate, succinate or α-ketoglutarate as the sole carbon source. In addition, CY902 exhibited faster carbon source consumption and more biomass accumulation on
succinate than on L-malate, fumarate or α-ketoglutarate, suggesting the succinate is more suitable as a dicarboxylate substrate for PkJEN transporters. Although the expression of both PkJEN2 genes was subject to glucose repression, low-level basal expression was observed. This was different from what was reported for ScJEN1 and CaJEN1, which are fully repressed by glucose (Casal et al., 1999; Soares-Silva et al., 2004). The expression of PkJEN2-1 and PkJEN2-2 can be significantly induced by SA, with 23-fold and fivefold increases compared to growth on glucose. Although ΔPkJEN2-1 had a more serious growth defect than ΔPkJEN2-2 on dicarboxylates, no such differential growth phenotype was observed for their overexpression mutants using the same constitutive promoter, suggesting that PkJEN2-1 is the major dicarboxylate transporter in CY902 due to its higher transcription. Interestingly, we found that ΔΔPkJEN2 almost lost growth on L-malate, fumarate or α-ketoglutarate, and showed a significant decrease in growth on succinate at pH 2.5, indicating that dicarboxylic acid molecules are not able to easily enter CY902 cells through passive diffusion at low pH. Similar behaviour was observed earlier with respect to passive undissociated succinic acid diffusion across the plasma membrane of yeast Hansenula anomala, and the permeability of cells to undissociated acids was reported to be 100-fold lower at pH 3.0 than at pH 6.0 (Corte-Real and Leao, 1990). These results indicate that the inactivation of PkJEN proteins can effectively prevent dicarboxylate from re-entering the cells, which can be used as an effective strategy to increase the yield of low-pH organic acids fermentation by yeast. Furthermore, we found that the dicarboxylate consumption of CY902 and its derivative strains markedly increased at pH 2.5 or 5.0 relative to pH 7.0, indicating that there is a clear pH dependence of dicarboxylate uptake. Previous studies have shown
that the intracellular pH (pHi) of *P. kudriavzevii* is approximately 8.0 (Halm *et al.*, 2004), so we speculated that the higher ΔpH at low pH may contribute to the active import of dicarboxylate via an electroneutral proton-symport mechanism, similar to that reported for ScJEN1 (Soares-Silva *et al.*, 2003). Notably, the L-lactate utilization was not significantly affected at pH 7.0 compared to pH 2.5 and 5.0, which indicates that other pH-independent permeases are responsible for lactate transport in *P. kudriavzevii*. Previous studies have shown that ScJEN1 and ScADY2 alleles (ScADY2Leu219Val and ScADY2Ala252Gly) are responsible for the transport of lactate in *S. cerevisiae* (de Kok *et al.*, 2012). We found that CY902 contains 6 proteins with high sequence similarity to ScADY2, although they are currently annotated as ATO2 (Table S2). Whether these proteins have L-lactate transport activity needs further experiments to verify. In addition, we found that CY902 cells cannot grow on pyruvate at pH 7.0, suggesting that pyruvate transport was different from that of L-lactate and depended on proton gradient. Monocarboxylate transporters (MCTs) are reported to be able to catalyse the bidirectional proton-linked transport of short-chain monocarboxylic acids such as L-lactate, ketone bodies and pyruvate across the plasma membrane of mammalian cells (Perez-Escuredo *et al.*, 2016). We found that eight MCH transporters in CY902 (Table S2) are homologous proteins of mammalian MCTs. We speculate that these proteins may be involved in pyruvate transport, albeit further research is needed to prove this hypothesis.

Additionally, we discovered that PkJEN2-2 has a specific tricarboxylate transport ability not found in PkJEN2-1 and other reported JEN2 homologous proteins. To date, few tricarboxylate transporters have been characterized in yeast. The low- and high-affinity transport systems for citrate uptake have been previously described in *Candida utilis* (Cassio and Leao, 1991, 1993), but the corresponding genes have not yet been cloned. In addition, it has been reported that the six JEN3 subfamily transporters of *Y. lipolytica* have citrate transport activity (Guo *et al.*, 2015), but this finding is still controversial because another independent study showed that the disruption of all six YUEN genes did not affect the growth of *Y. lipolytica* on citrate (Dulermo *et al.*, 2015). In this work, the molecular mechanisms underlying the substrate transport differences between PkJEN2-1 and PkJEN2-2 were explored based on structural analysis and rational protein engineering. We found that inserting the 392WCVQGGLGVVPS403 fragment located on the TMS-X of PkJEN2-2 at the corresponding site of PJEN2-1 protein, endowed the latter with significant citrate transport activity. Structurally, the 392WCVQGGLGVVPS403 fragment of TMS-X is sandwiched between the potential dicarboxylic acid specificity site (Gln\textsuperscript{207}Gln\textsuperscript{208}) on TMS-V and the conserved JEN domain (‘NXXS/THXS/TQD XXXT’) on TMS-VII. Notably, all these characteristic structures seem to be located parallel to the side of the JEN channel (Fig. 5E), suggesting that they may act synergistically to determine the transport activity and substrate specificity of the channel.

In addition, we found that the Asn\textsuperscript{341} residue on TMS-X of the modelling template PiPT, which has been reported to play a role in coordinating the phosphate substrate in the protein channel (Pedersen *et al.*, 2013), overlaps with the TMS-X mutation region of PkJEN2-2 (Fig. 5D). This suggesting that the mutation region on TMS-X may...
Unlike ScJEN1, we did not observe that overexpression of PkJEN2 in the cytoplasm leads to glucose concentration below 3 g l$^{-1}$ that ScJEN1 promoter activity is induced at glucose concentrations of 0.1 g l$^{-1}$ or less (Chambers et al., 2004). Our finding also suggests that the repression phenotype was deactivated in the cytoplasm due to the presence of dicarboxylic acid determination site to better recognize and accommodate the dicarboxylic acid molecule. However, further understanding of the mechanisms has been limited due to the lack of a real crystal structure of PkJEN2-2 and this needs to be pursued in the future. Although the focus of this study was to explore the dicarboxylate transport activity of PkJEN transporters, the citrate transport activity found in PkJEN2-2 can be used for other important industrial applications, such as citric acid production or deacidification of fruit wine.

Another aim of this study was to investigate potential applications of PkJEN transporters in SA production. Previous work has demonstrated that the ScJEN1 transporter is capable of bidirectional transport of lactate, and constitutive overexpression of ScJEN1 has been used as a strategy to enhance lactic acid production (Branduardi et al., 2006; Pacheco et al., 2012). However, ScJEN1-mediated lactic acid efflux seems to occur only when the external concentration of lactic acid is low. When the lactic acid titre reaches 8–10 g l$^{-1}$, the efflux effect is no longer obvious (Branduardi et al., 2006). Unlike ScJEN1, we did not observe that overexpression of PkJEN2-1 or PkJEN2-2 can promote SA efflux. On the contrary, the overexpression decreased the SA titre by about 25% before glucose was exhausted, implying that either PkJEN2-1 or PkJEN2-2 only has an inward unidirectional carboxylate transport activity. Thus, the expression of these transporters is unfavourable for the accumulation of extracellular SA. In this work, we found that inactivation of both PkJEN transporters lead to a higher glucose consumption rate and SA titre before glucose was depleted, which may be due to the removal of the basal expression of PkJEN proteins, thereby reducing the cell burden caused by the excreted SA re-entering the cell. Although the inhibitory effect of glucose on transporters largely alleviated this adverse effect, we found that the repression phenotype was deactivated in the late fermentation stage. A previous study reported that ScJEN1 promoter activity is induced at glucose concentrations below 3 g l$^{-1}$ and reaches its maximum at a glucose concentration of approximately 0.1 g l$^{-1}$ (Chambers et al., 2004), indicating that JEN proteins can be expressed before glucose is exhausted. This information indirectly explains the phenotype we observed, where by the SA titre of the control strain decreased rapidly at the late stage of fermentation. Therefore, knocking out the PkJEN gene may be an effective strategy to establish the optimal balance between carbon source consumption and product synthesis in the later stage of SA fermentation. In addition, it was reported that glycerol dose not inhibit the expression of JEN genes (Chambers et al., 2004). Given that glycerol has been commonly used for bio-based SA production as a cheap and abundant by-product of biodiesel production (Gao et al., 2016; Yu et al., 2019), the strategy of engineering PkJEN transporters may be more important in the production of SA from non-repressing carbon sources such as glycerol in the future. In addition, SA prevails in its anionic form at physiological pH, so it needs to be exported from the cell in an ATP-dependent process. Since the two PkJEN proteins do not mediate the outward transport of SA, the exact transporter that mediates SA efflux in P. kudriavzevii is still unclear. In S. cerevisiae, the ATP-binding cassette (ABC) transporter PDR12 has been reported to be capable of mediating the active efflux of organic acid anions from the cell, and SA can strongly induce its expression (Hatzixanthis et al., 2003). We found two PDR12 orthologues (PK3493 and PK4475) in the CY902 genome, but these two genes showed practically no expression in cells grown on glucose or SA (Table S2), implying that another presently unknown SA export mechanism exists in P. kudriavzevii.

**Experimental procedures**

**Strains and culture conditions**

The P. kudriavzevii isolate CY902 was deposited with the China General Microbiological Culture Collection Center (CGMCC, [http://www.cgmcc.net/](http://www.cgmcc.net/)) under the identity number 20885. CY902 and its derivative strains were routinely cultured in standard Yeast Peptone Dextrose medium (YPD; 1% w/v yeast extract, 2% w/v peptone, and 2% w/v glucose) at 30°C. The Synthetic Defined medium lacking Uracil (SD-URA; 0.67% w/v yeast nitrogen base with-out amino acids, 2% w/v glucose, and 0.077% w/v uracil drop-out supplement) was used to select URA$^+$ strains. *Escherichia coli* DH5α (Takara Bio, Shiga, Japan) was used as the cloning host and was grown in Luria–Bertani (LB; 0.5% NaCl, 1% tryptone, 0.5% yeast extract) medium supplemented with ampicillin (100 μg ml$^{-1}$) at 37°C.

**Isolation of acid-tolerant yeast strain**

The microorganisms living on the peel of wild fruits collected from the Yunnan Province of China were screened on YPD-agar plates (pH 2.5, adjusted using phosphoric acid). Surviving strains with morphological characteristics typical of yeasts were selected for further screening. Yeast cells were inoculated into inorganic acid-SD medium (2% w/v glucose as carbon source, medium adjusted to pH 2.0 with hydrochloric acid) and incubated at 30°C until the cells reached the late stage of fermentation. Therefore, knocking out the PkJEN gene may be an effective strategy to establish the optimal balance between carbon source consumption and product synthesis in the later stage of SA fermentation. In addition, it was reported that glycerol dose not inhibit the expression of JEN genes (Chambers et al., 2004). Given that glycerol has been commonly used for bio-based SA production as a cheap and abundant by-product of biodiesel production (Gao et al., 2016; Yu et al., 2019), the strategy of engineering PkJEN transporters may be more important in the production of SA from non-repressing carbon sources such as glycerol in the future. In addition, SA prevails in its anionic form at physiological pH, so it needs to be exported from the cell in an ATP-dependent process. Since the two PkJEN proteins do not mediate the outward transport of SA, the exact transporter that mediates SA efflux in P. kudriavzevii is still unclear. In S. cerevisiae, the ATP-binding cassette (ABC) transporter PDR12 has been reported to be capable of mediating the active efflux of organic acid anions from the cell, and SA can strongly induce its expression (Hatzixanthis et al., 2003). We found two PDR12 orthologues (PK3493 and PK4475) in the CY902 genome, but these two genes showed practically no expression in cells grown on glucose or SA (Table S2), implying that another presently unknown SA export mechanism exists in P. kudriavzevii.

**Experimental procedures**

**Strains and culture conditions**

The P. kudriavzevii isolate CY902 was deposited with the China General Microbiological Culture Collection Center (CGMCC, [http://www.cgmcc.net/](http://www.cgmcc.net/)) under the identity number 20885. CY902 and its derivative strains were routinely cultured in standard Yeast Peptone Dextrose medium (YPD; 1% w/v yeast extract, 2% w/v peptone, and 2% w/v glucose) at 30°C. The Synthetic Defined medium lacking Uracil (SD-URA; 0.67% w/v yeast nitrogen base without amino acids, 2% w/v glucose, and 0.077% w/v uracil drop-out supplement) was used to select URA$^+$ strains. *Escherichia coli* DH5α (Takara Bio, Shiga, Japan) was used as the cloning host and was grown in Luria–Bertani (LB; 0.5% NaCl, 1% tryptone, 0.5% yeast extract) medium supplemented with ampicillin (100 μg ml$^{-1}$) at 37°C.

**Isolation of acid-tolerant yeast strain**

The microorganisms living on the peel of wild fruits collected from the Yunnan Province of China were screened on YPD-agar plates (pH 2.5, adjusted using phosphoric acid). Surviving strains with morphological characteristics typical of yeasts were selected for further screening. Yeast cells were inoculated into inorganic acid-SD medium (2% w/v glucose as carbon source, medium adjusted to pH 2.0 with hydrochloric acid) and incubated at 30°C until the cells reached the late stage of fermentation. Therefore, knocking out the PkJEN gene may be an effective strategy to establish the optimal balance between carbon source consumption and product synthesis in the later stage of SA fermentation. In addition, it was reported that glycerol does not inhibit the expression of JEN genes (Chambers et al., 2004). Given that glycerol has been commonly used for bio-based SA production as a cheap and abundant by-product of biodiesel production (Gao et al., 2016; Yu et al., 2019), the strategy of engineering PkJEN transporters may be more important in the production of SA from non-repressing carbon sources such as glycerol in the future. In addition, SA prevails in its anionic form at physiological pH, so it needs to be exported from the cell in an ATP-dependent process. Since the two PkJEN proteins do not mediate the outward transport of SA, the exact transporter that mediates SA efflux in P. kudriavzevii is still unclear. In S. cerevisiae, the ATP-binding cassette (ABC) transporter PDR12 has been reported to be capable of mediating the active efflux of organic acid anions from the cell, and SA can strongly induce its expression (Hatzixanthis et al., 2003). We found two PDR12 orthologues (PK3493 and PK4475) in the CY902 genome, but these two genes showed practically no expression in cells grown on glucose or SA (Table S2), implying that another presently unknown SA export mechanism exists in P. kudriavzevii.
acid, sulfuric acid or phosphoric acid, respectively) or organic acid-SD medium (7% v/v lactic acid, 7% w/v D-malic acid or 7% w/v succinic acid as the sole carbon source respectively. The pH value of medium is about 2.0) at an initial OD<sub>600</sub> of 0.1. The S. cerevisiae BY4743 strain was used as a strain control, and SD medium without pH adjustment was used as a solvent control. The growth phenotype of acid-tolerant yeast was observed after 36 h of cultivation at 30°C.

**Genome assembly and annotation**

High-quality genomic DNA of the CY902 strain was extracted using the CTAB (cetyltrimethylammonium bromide) method and used for subsequent commercial Oxford Nanopore (Biomarker Technologies Co. Ltd, Beijing, China) and Illumina sequencing (Novogene Biotechnology Co. Ltd, Tianjin, China). The Oxford Nanopore long-read library was prepared according to the Ligation Sequencing Kit 1D (SQK-LSK109) protocol and sequenced using a MiniION DNA sequencer (Oxford Nanopore, Oxford, UK). The small Illumina paired-end library (2 × 125 bp) with an insert size of 380bp was sequenced using the HiSeq X Ten platform. For genome assembly, Canu (version 1.5) (Koren et al., 2017) was used to create initial reads, and the assembly was performed using Wtdbg2 (version 2.1) (Ruan and Li, 2020). The consensus assembly was generated by 3 rounds of Pilon (version 1.21) polishing using the Illumina reads. The completeness and accuracy of the genome assembly were evaluated by aligning the original Illumina reads to the final draft using the Burrows–Wheeler Alignment tool (BWA, v0.7.17) (Li and Durbin, 2009). The prediction of protein-coding genes was carried out using the BRAKER2 (Hoff et al., 2019) and PASA (Haas et al., 2003) pipelines, based on a strand-specific Illumina RNA-Seq library (paired-end, 2 × 150 bp) of CY902 wild-type cells grown in YPD medium to provide experimental prediction evidence. The predicted peptide sequences were then used as queries for blast searches against Swiss-Prot (release 2017-04, https://www.uniprot.org/), Pfam (release16.0, http://pfam.xfam.org/) and Saccharomyces Genome Database (SGD, https://www.yeastgenome.org/) for functional annotations with an E-value cut-off of 10<sup>-5</sup>. The identification and classification of the putative transporters in CY902 was based on information on homologous transporters from the Saccharomyces cerevisiae strain S288C listed in TransportDB (v2.0, http://www.membranetransport.org/transportDB2/index.html). The sequence data reported in this manuscript have been submitted to the NCBI Nucleotide Database with the following accession numbers: P. kudriavzevii CY902 nuclear and mitochondrial genome sequences (CP039612.1-CP039617.1); P. kudriavzevii CY902 MiniION nanopore genomic sequencing reads (SRA accession SRX5767040); P. kudriavzevii CY902 Illumina genomic sequencing reads (SRX5762150); and P. kudriavzevii CY902 strand-specific RNA-Seq Illumina reads (SRX5760412).

**CRISPR/Cas9 plasmid construction**

The pWS-Pk-URA-GFP plasmid was a kind gift from Dr. Zengyi Shao (Iowa State University, USA). It contains a 91 bp autonomously replicating sequence (ARS) and a uracil auxotrophy complementation marker (URA3) from P. kudriavzevii (Sun et al., 2020). Using the pWS-Pk-URA-GFP plasmid as the backbone, we constructed a CRISPR/Cas9 plasmid named pWSPK-Cas9 (GenBank MW296878) for gene editing in CY902 (Fig. S3A). The plasmid construction was based on a previous publication (Tran et al., 2019), with minor modifications as follows. Briefly, two DNA fragments required for the expression of functional Cas9 and single guide RNA (sgRNA) were commercial synthesized (GenScript, Nanjing, China) and incorporated into the plasmid backbone, including (I) the Cas9 expression cassette, which comprised the Cas9 coding sequence optimized for P. kudriavzevii codon usage flanked by the natural RNR2 promoter (Ryan et al., 2014) and CYC1 terminator; and (II) sgRNA expression cassette, a DNA fragment fused with the RNA<sup>sg</sup> RNA polymerase (RNAP) III promoter, RNA<sup>sg</sup> sequence, sgRNA and polyT sequence. The two DNA fragments were integrated into the pWS-Pk-URA-GFP backbone using a Seamless Cloning Kit (Beyotime Biotechnology, Shanghai, China).

**Transformation of P. kudriavzevii by electroporation**

To genetically modify the CY902 strain, the URA3 orthologous gene (PK2075) in CY902 was deleted using a previously reported homologous recombination procedure to generate the uracil auxotrophic mutant ΔURA3 (Xiao et al., 2014; Park et al., 2018). Unless otherwise specified, all subsequent genetic modifications in CY902 started with the ΔURA3 strain. The transformation of P. kudriavzevii was performed by an electroporation method modified from a previous publication (Wu and Letchworth, 2004). Briefly, 500μL of an overnight culture was transferred to 3 ml of YPD medium and incubated at 30°C and 250 rpm until the OD<sub>600</sub> reached 0.8-1.5. The cells were collected by centrifugation at 12 000 g for 1 min and washed twice with distilled water. The yeast cells were then resuspended in 1 ml of pretreatment reagent (100 mM Li-acetate, 10 mM DTT, 0.6 M sorbitol, and 10 mM Tris-HCl, pH 7.5) and incubated at room temperature for 30 min. The cells were then centrifuged and washed once with 1 M ice-cold sorbitol.
Subsequently, the cells were resuspended in 60 µl of ice-cold 1 M sorbitol and mixed with 500–1000 ng total DNA dissolved in <6 µl of ultrapure water. The mixture was then transferred to a 2-mm gap electroporation cartridge and incubated on ice for 5 min. A Gene-pulsar II (Bio-Rad, Hercules, CA, USA) was used to apply electroporation pulses at 2.7 kV. The electroporated cells were immediately diluted with 1 ml of ice-cold 1 M sorbitol and recovered at 30°C and 250 rpm for 60 min. The culture was then centrifuged at 12,000 g for 1 min to concentrate the cells, and 100 µl aliquots were spread on SD-URA plates, which were grown at 30°C for 2–3 days to select positive URA" transformants. To verify the genotype, a single colony was suspended in 10 µl of 20 mM NaOH solution and boiled at 98°C for 40 min to release genomic DNA, and 1 µl of the resulting DNA solution was used directly as the template for PCR.

RNA-Sequencing and data analysis

The ΔURA3 cells were grown overnight, and then transferred into 50 ml liquid SD complete medium to reach an initial cell density of OD_{600} ≈ 0.1, and cultured for 24 h to the logarithmic growth phase (OD_{600} ≈ 2-3). The cells were collected by centrifugation at 6000 g for 5 min, washed twice with phosphate buffered saline (PBS) and used to inoculate fresh 50 ml of SD medium containing either 2% w/v glucose (SoleG for short), 2% w/v glucose plus 2% w/v SA (GS) or 2% w/v SA as the sole carbon source (SoleS) respectively. To reduce the interference of the passive diffusion of undissociated carboxylic acids into the cells at low pH on the identification of SA transporters, the pH of the cultures was adjusted to 5.0 using NaOH as described previously (Queiros et al., 2007). After the cells were collected by centrifugation for total RNA isolation and Illumina RNA-Seq library construction. Paired-end libraries (2 × 125 bp) were generated by a commercial provider (Novogene Biotechnology Co. Ltd, Tianjin, China) using an Illumina HiSeq X Ten platform. Two biological replicates were sequenced for each condition. Clean reads were aligned to the CY902 genome using Bowtie 2 (version 2.2.5) (Langmead and Salzberg, 2012). The aligned reads were stored in SAM format, and the raw counts for reads mapped to unique genes were then tallied using the HTSeq-count script (0.6.0) in intersection-non-empty resolution mode. The abundance of each transcript was calculated using the Fragments Per Kilobase per Million (FPKM) method. Differentially expressed genes (DEGs) were identified using a final cut-off of log_{2} Ratio≥1. Sequencing data were deposited in the Gene Expression Omnibus database (GEO, http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE159457.

Strain engineering

To construct single-gene null mutants (ΔPkJEN2-1 and ΔPkJEN2-2), the Seamless Cloning Kit (Beyotime Biotechnology, Shanghai, China) was used to construct a pWSKP-Cas9 plasmid containing a specific sgRNA for each gene. The sgRNAcas9 software (v3.0.5; Xie et al., 2014) was used to select a sgRNA with a length of 20 nt and evaluate potential off-target cleavage sites in the CY902 genome. In addition, the Integrative Genomics Viewer (IGV, http://www.igv.org/) was used to avoid the sgRNA sequence containing any heterozygous SNV sites. To construct the template DNA for gene knockout, homology arms of approximately 500 bp selected from the upstream and downstream sequences of each gene were fused by overlap-extension PCR. The dedicated pWSKP-Cas9 and linear template DNA were used to co-transform the ΔURA3 strain by electroporation as mentioned above (Fig. S3B). To construct the double-gene null mutant ΔΔPkJEN2, the transformed pWSKP-Cas9 plasmid in ΔPkJEN2-1 was cured by culturing the colonies on a YPD plate overnight, which reinstated the uracil auxotrophy of the ΔPkJEN2-1 strain. Subsequently, the PkJEN2-2 gene was further deleted in the ΔPkJEN2-1 strain using the same gene deletion method as described above. To construct the overexpression strains for PkJEN2-1 and PkJEN2-2, cassette random insertion (RI) (Fig. S3C) and site-specific insertion (SI) strategies (Fig. S3D) were used separately. For the random insertion strategy, four individual fragments were prepared, including (i) the URA3 expression cassette amplified from the pWS-Pk-URA-GFP plasmid, (ii) the natural constitutive promoter (~800 bp) of the glyceraldehyde-3-phosphate dehydrogenase gene TDH3 (PK1425), (iii) the coding sequence of PkJEN2-1 or PkJEN2-2, and (iv) the natural terminator (~500 bp) of the galactose permease gene GAL2 (PK0762). These four DNA fragments were fused into the complete gene overexpression cassette by overlap-extension PCR (standard PCR procedures were used, except that each DNA template was added at a concentration of 60 ng µl⁻¹ at the same time), and then the cassette was randomly inserted into the chromosome of ΔΔPkJEN2 to generate the strains OV_{PkJEN2-1-RI} and OV_{PkJEN2-2-RI}. For the site-specific insertion strategy, the gene overexpression cassette (containing the PkJEN2-1 or PkJEN2-2 coding sequence flanked by the natural TDH3 promoter and GAL2 terminator) was first fused with the upstream (~500 bp) and downstream (~500 bp) homology arms of the phosphoribosylaminomimidazole carboxylase gene ADE2 to form the complete DNA template. Then, CRISPR/Cas9-assisted homologous recombination was used to specifically insert the gene overexpression cassette into the ADE2 locus, thereby generating
the strains OVpJEN2-1-SI and OVpJEN2-2-SI. The null mutant of ADE2 is adenine auxotrophic, which will cause a red-colour phenotype on YPD medium due to the accumulation of a red intermediate metabolite when cells are deprived of adenine (Bharathi et al., 2016). The red phenotype can be used to quickly screen positive transformants. The ADE2 gene was also deleted in ΔURA3 and ΔΔpJEN2 to generate ΔURA3ΔADE2 and ΔΔpJEN2ΔADE2 control mutants. To construct SA producing strains, the SDH5 gene was separately deleted in ΔURA3, ΔΔpJEN2, ΔΔpJEN2ΔADE2, OVpJEN2-1-SI and OVpJEN2-2-SI using the same CRISPR/Cas9-aided homologous recombination procedure as described above. All primers and spacer (N20) used in the strain construction are listed in the Tables S1 and S2. All the strains used in this study are listed in Table 1.

**Determination of transport substrate specificity**

P. kudriavzevii cells were cultured overnight, centrifuged and resuspended in water to an OD_{600} of 1.0. The cultures were then serially diluted with distilled water and 3 μl of each dilution was spotted onto SD-agar plates with either 2% w/v glucose, 2% v/v l-lactic acid, 2% v/v pyruvic acid, 2% w/v l-malic acid, 2% w/v fumaric acid, 2% w/v succinic acid, 2% w/v α-ketoglutaric acid, 2% w/v citric acid or 2% w/v DL-isocitric acid trisodium salt as the sole carbon source respectively. The pH of the medium was adjusted to 5.0 or 7.0 with NaOH or 2.5 with hydrochloric acid. To measure the consumption of carbon sources, cultures were also carried out in liquid SD medium. After 24 h of cultivation, the biomass of each culture was assessed by measuring the OD_{600} value, and the residual carbon source content in the medium was determined using an Agilent 1260 high-performance liquid chromatography (Agilent Technologies Inc., Santa Clara, CA, USA) equipped with a Bio-Rad Aminex HPX-87H column (Bio-Rad, Hercules, CA, USA). The column was operated at a flow rate of 0.5 ml min⁻¹ at 60°C, using 10 mM sulfuric acid as the mobile phase.

**Homology modelling and rational protein design**

The initial three-dimensional structure models of S. cerevisiae JEN1 protein, as well as P. kudriavzevii PkJEN2-1 and PkJEN2-2 proteins, were created using the I-TASSER server (https://zhanglab.ccmb.med.umich.edu/I-TASSER; Roy et al., 2010) with the default parameters. The top-scoring model created by I-TASSER was submitted to the GalaxyRefine program (http://galaxy.seoklab.org/index.html; Ko et al., 2012) for further model refinement. All the structures were visualized and manipulated using UCSF Chimera (v1.13.1). To verify the substrate specificity determining site of the transporter, the coding sequences for specific short peptides on TMS-V, TMS-VIII and TMS-X of PkJEN2-2 were separately integrated into primers and introduced into the corresponding sites of the PkJEN2-1 gene by overlap-extension PCR. The wild-type and rationally mutated PkJEN2-1 genes were expressed under the control of the TDH3 promoter and GAL2 terminator, and the entire gene cassette was incorporated into the ADE2 locus via CRISPR/Cas9-
ailed homologous recombination. The growth phenotype, carbon source consumption and biomass accumulation were assessed to determine the substrate specificity of the mutants. The experimental method was as described above, except that adenine sulfate was added to the medium at a final concentration of 120 mg l⁻¹.

**Shake flask cultures for SA production**

Yeast cells grown overnight were used to inoculate a 250 ml flask with 50 ml YPD medium (5% w/v glucose) to an OD600 of 0.1 and cultivated at 30°C and 250 rpm for 12, 24, 36 and 48 h. The pH was not controlled and dropped to a value of around 3.8. The SA concentration in the medium was measured by high-performance liquid chromatography according to a previously reported protocol (Gao et al., 2016).

**Phylogenetic analysis**

The phylogenetic tree was constructed based on an alignment of 35 JEN proteins obtained from 10 species of yeast and 5 species of filamentous fungi (Lodi et al., 2007). Homologues of PkJEN protein were confirmed using local BLASTp (version Blast 2.2.28) with an E-value <10⁻⁵ as the cut-off. Phylogenetic analysis of PkJEN proteins was carried out in MEGAX using the maximum likelihood method with 1000 bootstrap replicates.

**Statistical significance tests**

Unless otherwise noted, all experiments were performed in triplicate, and statistical significance was determined using one-way ANOVA in R (version 3.1.1).

**Acknowledgements**

This research was supported by grants from the National Key R&D Program of China (2019YFA0905300) and the National Natural Science Foundation of China (No. 31700066).

**Conflict of interest**

This work has been included in a patent application by the Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences.

**References**

Ahn, J.H., Jang, Y.S., and Lee, S.Y. (2016) Production of succinic acid by metabolically engineered microorganisms. *Curr Opin Biotechnol* 42: 54–66.

Beauprez, J.J., De Mey, M., and Soetaert, W.K. (2010) Microbial succinic acid production: natural versus metabolic engineered producers. *Process Biochem* 45: 1103–1114.

Bharathi, V., Girdhar, A., Prasad, A., Verma, M., Taneja, V., and Patel, B.K. (2016) Use of ade1 and ade2 mutations for development of a versatile red/white colour assay of amyloid-induced oxidative stress in *Saccharomyces cerevisiae*. *Yeast* 33: 607–620.

Branduardi, P., Sauer, M., De Gioia, L., Zampella, G., Valli, M., Mattanovich, D., and Porro, D. (2006) Lactate production yield from engineered yeasts is dependent from the host background, the lactate dehydrogenase source and the lactate export. *Microb Cell Fact* 5: 4.

Cao, M., Fatma, Z., Song, X., Hsieh, P.H., Tran, V.G., Lyon, W.L., et al. (2020) A genetic toolbox for metabolic engineering of *Issatchenkia orientalis*. *Metab Eng* 59: 87–97.

Casal, M., Paiva, S., Andrade, R.P., Gancedo, C., and Leao, C. (1999) The lactate-proton symport of *Saccharomyces cerevisiae* is encoded by JEN1. *J Bacteriol* 181: 2620–2623.

Casal, M., Paiva, S., Queiros, O., and Soares-Silva, I. (2008) Transport of carboxylic acids in yeasts. *FEMS Microbiol Rev* 32: 974–994.

Casal, M., Queiros, O., Talaia, G., Ribas, D., and Paiva, S. (2016) Carboxylic acids plasma membrane transporters in *Saccharomyces cerevisiae*. *Adv Exp Med Biol* 892: 229–251.

Cassio, F., and Leao, C. (1991) Low- and high-affinity transport systems for citric acid in the yeast *Candida utilis*. *Appl Environ Microbiol* 57: 3623–3628.

Cassio, F., and Leao, C. (1993) A comparative study on the transport of L(-)malic acid and other short-chain carboxylic acids in the yeast *Candida utilis*: evidence for a general organic acid permease. *Yeast* 9: 743–752.

Chambers, P., Issaka, A., and Palecek, S.P. (2004) *Saccharomyces cerevisiae* JEN1 promoter activity is inversely related to concentration of repressing sugar. *Appl Environ Microbiol* 70: 8–17.

Chen, G.Q., and Patel, M.K. (2012) Plastics derived from biological sources: present and future: a technical and environmental review. *Chem Rev* 112: 2082–2099.

Corte-Real, M., and Leao, C. (1990) Transport of malic acid and other dicarboxylic acids in the yeast *Hansenula anomala*. *Appl Environ Microbiol* 56: 1109–1113.

Douglass, A.P., Offei, B., Braun-Galleani, S., Coughlan, A.Y., Martos, A.A.R., Ortiz-Merino, R.A., et al. (2018) Population genomics shows no distinction between pathogenic *Candida krusei* and environmental *Pichia kudriavzevii*: one species, four names. *PLoS Pathog* 14: e1007138.

Dulemo, R., Gamboa-Melendez, H., Michely, S., Theveneau, F., Neuveglise, C., and Nicaud, J.M. (2015) The evolution of Jen3 proteins and their role in dicarboxylic acid transport in *Yarrowia*. *Microbiology Open* 4: 100–120.

Ferone, M., Raganati, F., Olivieri, G., and Marzocchella, A. (2019) Bioreactors for succinic acid production processes. *Crit Rev Biotechnol* 39: 571–586.

Gao, C., Yang, X., Wang, H., Rivero, C.P., Li, C., Cui, Z., et al. (2016) Robust succinic acid production from crude glycerol using engineered *Yarrowia lipolytica*. *Biotechnol Biofuels* 9: 179.
acid transporters in the pathogen Candida albicans. Mol Microbiol 75: 1337–1354.

Wu, S., and Letchworth, G.J. (2004) High efficiency transformation by electroporation of Pichia pastoris pretreated with lithium acetate and dithiothreitol. Biotechniques 36: 152–154.

Xiao, H., Shao, Z., Jiang, Y., Dole, S., and Zhao, H. (2014) Exploiting Issatchenkiia orientalis SD108 for succinic acid production. Microb Cell Fact 13: 121.

Xie, S., Shen, B., Zhang, C., Huang, X., and Zhang, Y. (2014) sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. PLoS One 9: e100448.

Yocum, R.R., Dole, S., and Pero, J.G. (2013) Production of organic acids by fermentation at low pH. MYRIANT CORP (US). Patent No. US2013/059828.

Yu, Y., Zhu, X., Xu, H., and Zhang, X. (2019) Construction of an energy-conserving glycerol utilization pathways for improving anaerobic succinate production in Escherichia coli. Metab Eng 56: 181–189.

Yuzbashiev, T.V., Yuzbasheva, E.Y., Laptev, I.A., Sobolevskaia, T.I., Vybornaya, T.V., Larina, A.S., et al. (2011) Is it possible to produce succinic acid at a low pH? Bioeng Bugs 2: 115–119.

Zhang, Y., Zhang, M., Zhu, W., Yu, J., Wang, Q., Zhang, J., et al. (2020) Succinate accumulation induces mitochondrial reactive oxygen species generation and promotes status epilepticus in the kainic acid rat model. Redox Biol 28: 101365.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.