

**RHODOTORULA TORULOIDES AS AN INDUSTRIAL PLATFORM FOR FOOD AND
CHEMICAL PRODUCTION USING ALTERNATIVE FEEDSTOCKS**

CROSS-REFERENCE TO RELATED APPLICATION

- 5 This application claims priority of Singapore Patent Application No. 10202400461Q filed on 21 February 2024, the contents of which are incorporated herein in their entirety by reference.

FIELD OF THE INVENTION

- 10 The present invention relates to a *Rhodotorula toruloides* strain that is enhanced through metabolic engineering and laboratory-directed evolution for use in producing food and chemical production using alternative feedstocks. More particularly, the engineered strain which overexpresses Carnitine acetyltransferase (*CAT*) and/or transcription factor *HAA1* genes, and the evolved strain have enhanced acetate utilization and/or fatty acid and/or lipid
15 and/or biomass production compared to wildtype *R. toruloides*.

BACKGROUND OF THE INVENTION

- The global shift to sustainability has driven the exploration of alternative feedstocks beyond sugars for biomanufacturing [Zhang, C., et al., 2022]. Single-cell oils (SCO) and single-cell
20 protein (SCP) produced by microorganisms can be used as food ingredients. SCO is intracellular storage lipids composing mainly triacylglycerols (TAGs) [Chawla, K., et al., 2022; Maina, S., et al., 2017; Mhlongo, S.I., et al., 2021; Ochsenreither, K., et al., 2016]. SCO can be used to replace edible oils that are traditionally found in the plant and animal kingdoms [Demir, M. and A.G. Gündes, 2020]. SCP is semi-purified proteins from the cells of
25 microorganisms. It has high protein and vitamins contents and is increasingly regarded as a future food ingredient [Demir, M. and A.G. Gündes, 2020]. SCP has already been used in food products, such as meat substitute, food spreads, flavouring agents and food supplements [Jach, M.E., et al. 2022; Ritala, A., et al., 2017]. SCP-based products with the brand name QUORN, PEKILO, MARMITE, VEGEMITE and CENOVIS have a long history of
30 safe human consumption [Ritala, A., et al., 2017; García Martínez, J.B., et al., 2022; Nyssölä, A., et al., 2022]. The advantages of SCO and SCP include high productivity, low CO₂ and water footprints, and pesticide-free production [Mhlongo, S.I., et al., 2021; Ochsenreither, K., et al., 2016; Demir, M. and A.G. Gündes, 2020].

SCO and SCP production in the past have been relying on sugars as the feedstock except phototropic cultivation of microalga. The first generation (G1) feedstock is produced from sugar crops, such as sugarcane and sugar beet while the second generation (G2) feedstock is derived from the more abundant cellulose-rich biomass from forestry and agriculture [Bratosin, B.C., S. Darjan, and D.C. Vodnar 2021]. Driven by the Paris Agreement on climate change with the goal to limit global warming to less than 1.5°C, most countries in the world have pledged to achieve CO₂ neutrality by mid-century. With the natural abundance and rapid advance in catalytic, biocatalytic and microbial electrochemical synthesis, C1 (CO₂, CO, methane, methanol and formate) and C2 (acetate and ethanol) substrates are increasing hailed as an alternative feedstock for microbial fermentation [Zhang, C., et al., 2022; Jiang, W., et al., 2021; Fang, X., et al., 2020]. The utilization of these feedstocks will be more effective in reducing carbon emissions and supporting sustainability. Methanol can be utilized efficiently as the sole carbon source by a small number of methylotrophic yeast and bacteria [Espinosa, M.I., et al., 2020]. For example, engineered *Pichia pastoris* allowed the production of free fatty acids and fatty alcohol from methanol, with a free fatty acid titre and yield of 23.4 g/L and 0.078 g/g, respectively [Cai, P., et al., 2022]. Similarly, *Ogataea polymorpha* has been engineered to produce 15.9 g/L free fatty acid from sole methanol [Gao, J., et al., 2022]. With lower toxicity and higher energy density, C₂ feedstock may be a better next generation feedstock than methanol. In addition, acetate is an abundant by-product in biomass hydrolysate. For example, 49.8 mM and 34.3 mM acetate were found in soybean hull and spruce acid hydrolysate [Hu, C., et al., 2009; Chen, L., et al., 2021]. Therefore, biomass conversion rate to food and chemicals should be significantly improved if acetate can be efficiently utilized as the co-substrate.

Different methods have been applied to increase the production of lipids from acetate, including bioprocess engineering, substrate co-feeding and metabolic engineering. For example, metabolic engineering and co-substrate fermentation in *Y. lipolytica* has significantly increased lipid production from acetate by overexpressing acetyl-CoA synthetase (ACS), acetyl-CoA carboxylase (ACC1) and fatty acid synthase [Chen, L., et al., 2021]. The lipid content was increased by 95% through acetate and glycerol co-substrate fermentation. Coupled metabolic engineering with metabolic control in fermentation, *Y. lipolytica* was able to attain 115 g/L of lipid content, 0.16 g/g lipid yield and 0.8 g/L/h of productivity [Xu, J., et al., 2017].

Rhodotorula toruloides is an emerging host for metabolic engineering and synthetic biology due to its high lipid and cell mass productivity [Zhao, Y., et al., 2021; Park, Y.-K., et al., 2018]. Although acetic acid is naturally produced by many microbes, its presence could be a

limiting factor in microbial fermentation. For example, ethanol production strongly inhibits *Pichia stipites* growth at 83 mM [Gong, Z., et al., 2015]. Acetate had little effect on lipid production in *R. toruloides* up to 70 mM (4.2 g/L). In fact, acetate at 120 mM has been shown to improve cell performance and lipid content in *R. toruloides* Y4 [Hu, C., et al., 2009; 5 Chen, L., et al., 2021]. Microbes have diverse capacity to utilize acetate. For example, *Cryptococcus curvatus*, *Trichosporon cutaneum* AS 2.571 and *T. fermentans* CICC 1368 are much more efficient in utilizing acetate for lipid production than *R. toruloides* and *Y. lipolitica* [Gong, Z., et al., 2015].

There is a need to develop improved lipid production, fatty acid and/or cell mass profile in a 10 microorganism.

SUMMARY OF THE INVENTION

R. toruloides C3 is a strain isolated in Singapore by Temasek Life Sciences Laboratory Limited. The strain C3 was used as the host for yeast genetic manipulation. It's genome was 15 sequenced and the nucleotide sequences of the contigs are set forth in SEQ ID NOs: 33-50. Lipid production in *R. toruloides* C3 was enhanced through metabolic engineering and laboratory directed evolution. The *R. toruloides* C3 genome encodes 4 candidate carnitine acetyltransferase (CAT) genes, *CAT1*, *CAT2*, *CAT3* and *CAT4*, the overexpression of which led to enhanced acetate resistance and utilization. *CAT1* encoding a peroxisome targeted 20 enzyme was the most effective. Overexpression of *HAA1*, which encodes a bread yeast homolog of transcriptional activator, also significantly enhanced acetate utilization although *CAT1* and *HAA1* showed no additive effect. Over-expression of *CAT 1-4* genes also significantly increased PUFA content. In 2L-scale fed-batch fermentation using 3% acetate with 0.5% dextrose monohydrate as the co-carbon source, the *CAT1*-overexpressing strain 25 increased acetic acid consumption, biomass titer and lipid titer respectively. Maximal dry biomass titer reached 156.5 g/L, 22% higher than control and this was reached 24 hr earlier. Lipid contents were significantly higher than control over the course of fermentation, reaching a maximum of 19.8 g/L. There was a 99.5% increase of cumulative consumption of acetic acid. The biomass yield, however, was not increased by the overexpression of *CAT1*.

30 At maximal biomass titer, biomass yield was 0.55 g/g acetic acid (at 216 hr), lower than 0.68 g/g observed with control at 240 hr, possibly due to the need for more substrate to produce lipid. The low yield was, in part, attributable to high glycerol generation. At maximal biomass titer, glycerol titer was 65.6 g/L and 48.3 g/L respectively for *CAT1*-OE and control strain. Consistent with the flask culture data (Table 3), The average ALA content of samples from 35 all time-points was 2.58% of total fatty acids for pKCL2 control strain, compared with 9% in *CAT1*-OE strain, an increase of 249%.

Deletion of Cat1 peroxisome targeting signal only marginally weakened the biomass and lipid production-promoting activities of Cat1, suggesting it functions mainly by increasing the cytoplasmic pool of L-carnitine-CoA, which may lead to higher levels of cytoplasmic and mitochondria acetyl-CoA.

5

An adaptive laboratory evolution-derived strain, designated C3_64B, was able to produce 99.63 g/L of biomass, with 13%-16% lipid in 2L fed-batch fermentation. Interestingly the evolved strains had a drastically increased α -linolenic acid content, from ~4% to 15%. In a semi-continuous fermentation process, C3 strain produced > 100 g/L dry cell mass with 44-45% protein, ~ 10% lipid and 6.0-7.2 g/L/h productivity when glucose was used as the sole carbon source. Thus, *R. toruloides* is a promising host for SCP and SCO production, using either sugars or acetic acid as the feedstock.

Accordingly, in a first aspect the disclosure provides a genetically engineered *R. toruloides* cell, wherein the cell has been transformed by at least one polynucleotide molecule; the at least one polynucleotide molecule comprising a carnitine acetyltransferase (*CAT*) selected from the group consisting of *CAT1*, *CAT2*, *CAT3* and *CAT4*; and/or transcription factor *HAA1* gene, operably linked to at least one promoter, wherein overexpression of *CAT* and/or *HAA1* provides increased acetate utilization, fatty acid and/or lipid and/or biomass production compared to wildtype *R. toruloides*. In some embodiments the *CAT* gene is *CAT1*.

In some embodiments, the *CAT* and/or *HAA1* genes are operably linked to a strong promoter to enhance the expression level of the genes and achieve a protein expression level that is at least 5 times higher, preferably 10, 100 or 100 times higher. In some embodiments the strong promoter is a constitutive promoter such as, for example selected from the group comprising translation-elongation factor 1 alpha (*TEF1*) and glycerol 3-phosphate dehydrogenase (*GPD1*).

In some embodiments, the *TEF1* promoter and the *GPD1* promoter are from *R. graminis*. Preferably the promoters are from *R. graminis* WP1 strain. In some embodiments the *TEF1* promoter comprises the nucleotide sequence set forth in SEQ ID NO: 11. In some embodiments the *GPD1* promoter comprises the nucleotide sequence set forth in SEQ ID NO: 13.

35

In some embodiments, the isolated genetically engineered *R. toruloides* cell has a genome sequence that is at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity, to the contig polynucleotide sequences set forth in SEQ ID NOs: 33-50. In some
5 embodiments, the genetically engineered *R. toruloides* cell is genetically engineered *R. toruloides* C3 strain. In some embodiments the genome nucleotide sequence of *R. toruloides* C3 strain is set forth in SEQ ID NOs: 33-50.

In some embodiments, the *CAT* gene is *CAT1* and encodes the amino acid sequence set
10 forth in SEQ ID NO: 1.

In some embodiments, the *HAA1* gene encodes the amino acid sequence set forth in SEQ ID NO: 3.

15 In some embodiments, the *CAT1* gene comprises a polynucleotide sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity, due to the degeneracy of the genetic code, to the polynucleotide
20 sequence shown in SEQ ID NO: 2.

In some embodiments, the *HAA1* gene comprises a polynucleotide sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%,
25 at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity, due to the degeneracy of the genetic code, to the polynucleotide sequence shown in SEQ ID NO: 4.

In a second aspect the disclosure provides an isolated directed-evolution-derived *R. toruloides* C3 strain having enhanced acetate utilization, decreased glycerol by-product accumulation, and stress resistance compared to wildtype *R. toruloides*. Preferably, in some
30 embodiments the isolated directed-evolution-derived *R. toruloides* strain evolution is performed by adapting a culture of *R. toruloides* C3 against a gradual increase of acetate concentration in the culture medium.

35

In some embodiments, the evolved C3 strain can tolerate up to 6% acetic acid.

In some embodiments, the evolved C3 strain can tolerate up to 6% acetic acid and up to 4% glycerol. In one embodiment, the evolved C3 strain is C3_64B strain.

- 5 In some embodiments, the evolved C3 strain can produce higher α -linolenic acid (ALA) content compared to genetically engineered *CAT1* C3 strain because it can utilize glycerol as a carbon source.

- 10 In a third aspect the disclosure provides an isolated *R. toruloides* C3 cell, wherein the cell has a genome sequence that is at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity, to the polynucleotide sequences set forth in SEQ ID NOs: 33-50.

- 15 In a fourth aspect the disclosure provides a method of producing increased fatty acids, lipids and/or biomass by a genetically engineered *R. toruloides* cell, or by a directed-evolution-derived *R. toruloides* cell, comprising culturing a plurality of genetically engineered *R. toruloides* cells or directed-evolution-derived *R. toruloides* cells of any one of the embodiments described above in medium under conditions for fatty acids, lipid production and/or increased biomass, wherein the medium comprises acetate as carbon source, 20 wherein said genetically engineered cell or directed-evolution-derived cell is capable of increased production of fatty acids, lipids and/or biomass compared to a strain of identical genome background or non-evolved *R. toruloides* cell.

- 25 In some embodiments, the *R. toruloides* cell is an *R. toruloides* C3 cell, preferably a *CAT1* C3 cell, more preferably a C3_64B cell.

In some embodiments, the increased fatty acid is alpha-linolenic acid.

- 30 In some embodiments, the said method further comprises having glycerol in the medium.

In a fifth aspect the disclosure provides the use of a *R. toruloides* cell of any aspect of the invention in a method of producing enhanced lipid and/or biomass production compared to wildtype *R. toruloides*.

- 35 In some embodiments, the *R. toruloides* cell is an *R. toruloides* C3 cell, preferably a *CAT1* C3 cell, more preferably a C3_64B cell.

BRIEF DESCRIPTION OF THE DRAWINGS

Certain embodiments of the present disclosure are described more fully hereinafter with reference to the accompanying drawings.

- Figure 1** shows microbial diversity on acetate utilization. RT1, RT2 and C3: *R. toruloides* strains; WP1: *R. graminis* WP1; GS115: *Pichia pastoris* (Komagataella phaffii); SSY: *Saccharomyces cerevisiae*. A single colony of each strain was inoculated into 5 mL YPD in 50 mL Falcon tubes and incubated at 28 °C overnight. After dilution to OD₆₀₀ 0.2, 1 mL diluted culture was inoculated into 49 mL acetate assimilation medium with 3% acetate in 250 mL conical flasks and cultured at 30 °C for 3 and 5 days.
- Figure 2** shows optimization of medium for cell mass and lipid production for C3 strain. **(a)** Different medium compositions. The reported media were modified to have glucose substituted by 3% acetic acid as the sole carbon source, **(b)** Effect of different pH on biomass. **(c)** Medium with different nitrogen sources. Y1 : yeast extract 2.15 g/L, urea 2g/L; Y2: yeast extract 2.15 g/L, NH₄Cl 3.68 g/L ; Y3 : yeast extract 2.15 g/L, KNO₃ 6.63 g/L; Y4: yeast extract 4.3 g/L, urea 2 g/L; Y5: yeast extract 4.3 g/L, NH₄Cl 3.68 g/L; Y6: yeast extract 4.3 g/L, KNO₃ 6.63 g/L; Y7: yeast extract 6.45 g/L, urea 2g/L; Y8: yeast extract 6.45 g/L, NH₄Cl 3.68 g/L; Y9: yeast extract 6.45 g/L, KNO₃ 6.63 g/L. **(d)** Effect of acetate concentrations on biomass and lipid production. Error bars are the standard deviation of triplicates.
- Figure 3** shows the performance of Wt C3 and the final adapted population, C3_64B was inoculated into media with various acetic acid to glycerol ratio. GAAM24 contains 2% acetate/4% glycerol and so on. **(a)** OD 600 nm of strain C3 and C3_64B and **(b)** Acetate consumption of strain C3 and C3_64B.
- Figure 4** shows the effect of overexpression of various genes were over-expressed under the Rg *TEF1* promoter in C3CK strain, a derivative of C3 containing the KU70 gene knockout and an inducible cre recombinase gene as described in Koh, C.M.J., Liu, Y., Moehninsi et al. (2014). Three different single colonies of each construct were inoculated into 5 ml YPD in 50 ml Falcon tube and incubated at 28-30 °C overnight. Cells were diluted to 0.05 OD₆₀₀ in 50 ml medium in a 250 ml conical flask. **A.** Biomass of different strains cultured in AMP3-Y6 medium with 3% acetic acid at 28 °C for 5 days. **B.** Biomass of different strains cultured at 30 °C in AMP3-Y4.3 with 4% acetic acid (ie 6.63 g/L KNO₃, 5.7 g/L KH₂PO₄, 1 g/L MgSO₄·7H₂O , 0.015 g/L FeCl₃·6H₂O, 40 g/L acetic acid and 4.3 g/L yeast extract. The p values indicate the probability that an observed difference between control strain and over-expression strains. ** and * indicate p<0.01; 0.05, respectively.

Figure 5 shows the effect of MgSO_4 level on dry cell mass. A single colony was inoculated into 5 mL YPD in 50 mL Falcon tube and incubated at 28 °C overnight. All of the 5 mL overnight culture was inoculated into 45 mL YPD in 250 mL conical flasks and incubated for 24 hours at 30 °C. 5 mL of this subculture was inoculated into 45 mL of each different medium with different $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ content in 250 mL conical flasks and incubated for 72 hours at 30 °C. To remove calcium precipitates, cell pellets were treated with 1 mL 1 M citric acid for 30 mins, and washed twice with distilled water before being dried at 65°C for 48 hours. The basic medium contained yeast extract 4.3 g/L, urea 2 g/L, KH_2PO_4 5.7 g/L, $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ 0.015 g/L, acetic acid 30 g/L. All media were adjusted to pH 7.0 with NaOH.

Figure 6 shows the effect of medium pH on cell biomass production. A single colony of C3 was inoculated into 5 mL YPD in 50 mL Falcon tube and incubated at 28 °C overnight. All the 5 mL overnight culture was inoculated into 45 mL YPD in 250 mL conical flasks and incubated for 24 hours at 30 °C. 5 mL of this subculture was inoculated into 45 mL of APM2 medium of different pH value in 250 mL conical flasks and incubated for 72 hours at 30 °C. All the cell pellets were treated with 1 mL 1M citric acid for 30 mins, washed twice with distilled water before being at 65°C for 48 hours.

Figure 7 shows the effect of acetate and glycerol concentrations. Acetate assimilation medium (AAM, pH 6.5) described by Gong et al. (2015) was used to culture strain C3 at various acetate and glycerol concentrations. Cell were cultured for 3 days at 30°C with agitation (210 rpm) in 250 mL flasks. **(a)** C3 was cultured in AAM medium with 0-3% acetic acid. **(b)** C3 was cultured in AAM medium with 1% acetic acid and various levels of glycerol.

Figure 8 shows the effects of co-substrates. Strain C3 seed culture (5 mL) prepared in YPD medium was inoculated into 50 mL APM2 basal medium 250 mL conical flasks. APM2 basal medium contained 10 g/L acetic acid and 25 g/L of respective co-substrate as the carbon source. Culture was re-inoculated into APM2 with no glucose but 10 g/L acetic acid only after Day 3, except for glucose setup. Upper line in each graph is the co-substrate level. Lower line in each graph represents acetate level.

Figure 9 shows a comparison of cell biomass and lipid production of different evolved isolates. C3 Wt and 22 individual colonies of 25th generation of the adapted population (C3_64) were cultured in 50 mL GJM3 in 250 mL conical flasks for 5 days at 30°C and 210 rpm.

Figure 10 shows the effects of overexpression of CAT1, HAA1, DGA1 on biomass **(a)** and lipid **(b)** production.

Figure 11 shows the effect of overexpression of ACS1. C3_pKC2.3GPD1_ACS1: The ACS1 CDS was expressed using the RgGPD1 promoter and terminator. C3_pKC2.3TEF1_ACS1/MTS: the RfACS1 ORF was fused with the mitochondria using the malate dehydrogenase 2 mitochondria targeting sequence, MAAATRQSSRLARSFSTSARAN (SEQ ID NO: 15) and overexpressed under the control of RgTEF1 promoter and terminator. C3_pKC2 is the control vector. All constructs were integrated at the CAR2 locus by homologous recombination. Error bars indicate standard deviation of 3 biological replicates. Strains were cultured in AAM medium with 3% acetate for 5 days.

Figure 12 shows the effect of peroxisome targeting on CAT1 functions. The C-terminal PTS signal of CAT1, AKL tripeptide, was deleted by PCR and expressed under the same RgTEF1 promoter. 3 independent knock-in mutants at the CAR2 locus of C3CK strain were cultured in 50 ml AMP3-Y6 medium with 4% acetic acid at 28 °C for 5 days. *** indicate p value <0.001 between control vector pKCL2 and overexpression constructs. **A.** Cell density over 5 days; **B.** Dry Cell Mass; **C.** lipid content.

Figure 13 shows semi-continuous fed-batch fermentation. Wt C3 frozen stock at -80°C was inoculated to 5 mL APM2 medium and cultured for about 16 hr at 30°C. 100 mL APM2 medium of about 10 OD₆₀₀ units is prepared by culturing the overnight culture (1 mL) in APM2 medium in conical flasks. Seed culture was inoculated to 1 litre of APM2 medium at a dilution rate of 10% in a 2 L bioreactor and cultured at 30°C. pO₂ was maintained at 30% by cascade control of agitation speed and air-flow. pH was maintained at 5.5 using HCl and NaOH. Glucose level was maintained at 40-60 g/L by adjusting the flow rate of the feeding solution, between 4-8 mL/hr. In the semi-continuous fed-batch phase, 700-1000 mL fermentation broth was removed when OD₆₀₀ reached about 320 and replaced with 500-800 mL Replacement Medium (CRM). This bleeding/feeding process was repeated about every 6 hours so as to maintain the cell density at 300-400 OD₆₀₀ units. Biomass productivity and glucose consumption (glucose consumed per gram biomass) refer the transient value of over the two sampling times. (SI Te to confirm the volume for RM). The figures on top of the biomass titers indicate the protein contents of the respective time point.

(a) Effect of nitrogen source and content in CM. The content of yeast extract (YE) and urea in the RM in each phase are in the respective window. **(b)** Semi-continuous fed-batch fermentation of Wt C3 strain. RM contained 17.2 g/L yeast extract and bleeding/feeding was repeated every 6 hours. **(c)** Effect of nitrogen source on lipid content.

Figure 14 shows fed-batch fermentation of CAT4 and C3-64B strains. 100 mL of Seed culture in YPD was inoculated into 1 L AMP3-Y6 medium with 3% acetic acid and

supplemented with 5 g/L glucose, 0.1 g/L L-Histidine, 0.01 g/L; L-Aspartic acid; 10 mL vitamins of AAM in a 2 L bioreactor. The pH was adjusted to pH 7 using 5 M NaOH and pO₂ was maintained at 30%. pH during fermentation was maintained at 7.0 using glacial acetic acid. Feeding is started about 40 h at the flow rate so that glucose and acetic acid can be maintained at about 5 g/L and 30 g/L respectively. (A) Biomass titer; (B) Growth rate; (C) Biomass yield; (D) Lipid titer; (E) Lipid content; (F) Lipid yield.

DETAILED DESCRIPTION OF THE INVENTION

Further details of the invention will now be described with reference to the following non-limiting examples. Unless otherwise defined herein, technical and scientific terms used in the present description have the meanings that are commonly understood by those of ordinary skill in the art.

A. Definitions

As used herein, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise.

As used herein, the term "comprising" may include the embodiments "consisting of" and "consisting essentially of". The terms "comprise(s)," "include(s)," "having," "has," "can," "contain(s)," and variants thereof, as used herein, are intended to be open-ended transitional phrases that require the presence of the named features/steps and permit the presence of other features/steps. However, such description should be construed as also describing compositions, mixtures, or processes as "consisting of" and "consisting essentially of" the enumerated features/steps, which allows the presence of only the named features/steps, along with any impurities that might result therefrom, and excludes other features/steps.

The terms "amino acid" or "amino acid sequence," as used herein, refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

As used herein, the terms "polypeptide", "peptide" or "protein" refer to one or more chains of amino acids, wherein each chain comprises amino acids covalently linked by peptide bonds, and wherein said polypeptide or peptide can comprise a plurality of chains noncovalently and/or covalently linked together by peptide bonds, having the sequence of native proteins, that is, proteins produced by naturally-occurring and specifically non-recombinant cells, or

genetically-engineered or recombinant cells, and comprise molecules having the amino acid sequence of the native protein, or molecules having deletions from, additions to, and/or substitutions of one or more amino acids of the native sequence. A "polypeptide", "peptide" or "protein" can comprise one (termed "a monomer") or a plurality (termed "a multimer") of amino acid chains.

The terms "nucleotide", "nucleic acid" or "nucleic acid sequence", as used herein, refer to an oligonucleotide, polynucleotide, or any fragment thereof, to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

The term "operably linked" as used herein refers to a first molecule that can be joined to a second molecule, wherein the molecules are so arranged that the first molecule affects the function of the second molecule. The term "operably linked" includes the juxtaposition of two or more components (e.g., a promoter and another sequence) such that both components function normally and allow for the possibility that at least one of the components can mediate a function that is exerted upon at least one of the other components. The two molecules may or may not be part of a single contiguous molecule and may or may not be adjacent. For example, a promoter is operably linked to a polynucleotide molecule if the promoter modulates transcription of the polynucleotide molecule of interest in a cell. In additional embodiments, two portions of a transcription regulatory element are operably linked to one another if they are joined such that the transcription-activating functionality of one portion is not adversely affected by the presence of the other portion. Two transcription regulatory elements may be operably linked to one another by way of a linker nucleic acid (e.g., an intervening non-coding nucleic acid) or may be operably linked to one another with no intervening nucleotides present.

A vector can include one or more carnitine acetyltransferase (CAT) or *HAA1* nucleic acid(s) in a form suitable for overexpression of the nucleic acid(s) in a host cell. Preferably the recombinant expression vector includes one or more regulatory sequences operatively linked to the nucleic acid sequence(s) to be expressed. The term "regulatory sequence" includes promoters, enhancers, ribosome binding sites and/or IRES elements, and other expression control elements (e.g., polyadenylation signals). The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or polypeptides encoded by nucleic acids as described herein (e.g., CAT or *HAA1* proteins).

The recombinant expression vectors of the invention can be designed for over-expression of CAT or HAA1 proteins in prokaryotic or eukaryotic cells. For example, polypeptides of the invention can be expressed in yeast (e.g., *R. toruloides*). Having now generally described the invention, the same will be more readily understood through reference to the following
5 examples which are provided by way of illustration, and are not intended to be limiting of the present invention.

A person skilled in the art will appreciate that the present invention may be practiced without undue experimentation according to the methods given herein. The methods, techniques and chemicals are as described in the references given or from protocols in standard
10 biotechnology and molecular biology textbooks.

EXAMPLES

The following examples are intended to exemplify the present disclosures and are not limitations of the claimed invention. All molecules, compositions, methods, assays, and results disclosed in the examples and other sections of the specification, figures, and claims
15 form part of the disclosure of the invention.

Example 1

Materials and methods

1.1 Strains, medium compositions

The strains used were *R. toruloides* ATCC 10657 and *R. toruloides* ATCC 10788 which were
20 obtained from American Type Culture Collection (USA), *R. graminis* strain WP1 which obtained from Fungal Genetics Stock Center (University of Missouri, USA) and *R. toruloides* C3 is a strain isolated in Singapore by Temasek Life Sciences Laboratory Limited. Yeast strains were maintained in YPD agar and medium at 28-30°C. *Escherichia coli* XL1-BLUE competent cells were used for plasmid amplification and *R. toruloides*. The strain C3 was
25 used as the host for yeast genetic manipulation.

Various media were used in this work. Acetate Assimilation Medium (AAM) has been described by Gong, Shen [Gong, Z., et al., 2015]. AAM is usually supplemented with 10 mL/L AAM-vitamins composed of 50 mg/L thiamine hydrochloride, 50 mg/L riboflavin, 50
30 mg/L nicotinic acid, 50 mg/L pantothenic acid, 10 mg/L pyridoxine hydrochloride, 20 mg/L biotin, 20 mg/L folic acid, 50 mg/L 4-aminobenzoic acid, 50 mg/L cyanocobalamin.

Glycerol acetate assimilation medium (GAAM12) contains 8.6 g/L yeast extract, 5.7 g/L KH_2PO_4 , 5.7 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2 g/L NaCl, 80 mg/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 10 g/L acetate, and 20 g/L

glycerol. Different levels of acetate and glycerol were used in GAAM during the course of laboratory-directed evolution. For example, GAAM with 2% acetate and 4% glycerol was denoted as GAAM24.

- 5 The original APM medium contains 8.6 g/L yeast extract, 5.7 g/L KH_2PO_4 , 5.7 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2 g/L NaCl, 80 mg/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 10 g/L, 80 g/L glucose monohydrate. APM2 medium is same as AMP except 8.6 g/L yeast extract was reduced to 4.3 g/L and 2 g urea was supplemented as the inorganic nitrogen source. The Replacement Medium (RM) used semi-continuous fed-batch fermentation contains 17.2 g/L yeast extract, 5.7 g/L
- 10 KH_2PO_4 , 5.7 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2 g/L monosodium glutamate (MSG), 80 mg/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 80 g/L glucose monohydrate, 15 g/L NH_4Cl . The Continuous Feeding Medium (CFM) for the semi-continuous fed-batch fermentation contains 800 g/L glucose monohydrate and 35 g/L NH_4Cl . Acetate Feeding Medium (AFM), composed of 30 g/L acetic acid glacial, 5 g/L glucose, 100 g/L yeast extract, 20 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 66 g/L KNO_3 , 150 mg/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$,
- 15 1 g/L L-Histidine, 100 mg/L L-aspartic acid and 10 mL AAM vitamin solution.

1.2 Shake-flask cultures

- A single colony established on an agar plate was inoculated into 5 mL YPD in a 50 mL
- 20 Falcon tube; incubated at 28°C overnight with agitation (210 rpm) before inoculated into 50 mL appropriate medium in a 250 mL conical flask. The culture usually started at 0.2 OD_{600} and cultured at 28-30°C with agitation at 200 rpm for 3-5 days. Biomass, lipid content and acetate consumption were determined as described previously [Liu, Y., et al., 2021].

25 1.3 Plasmid construction and yeast transformation

The primers and plasmids used are listed in Table 1 and Table 2, respectively.

Table 1: Primers used in the Examples

Primer	Sequence (5'-3')	SEQ ID NO:
3GPD1ACS1F	caccacatttctccagctcgaaccATGGCCGACGTCGACTACAA	19
3GPD1ACS1R	ggtacggaggcgctggatattCACTCGGACGAGAGCTTG	20
3TEF1CAT1F	cgtcctcgccgccctgcagcacgccATGCACGCAGCCCGCAGAA	21
3TEF1CAT1R	aagaagcagctcgtgaggatattTACAGCTTGGCCTTGTCGGC	22
3TEF1CAT2F	cgtcctcgccgccctgcagcacgccATGCTTGTTGCACGGCTG	23
3TEF1CAT2R	aagaagcagctcgtgaggatattTCACGCCTTCGCCTTCC	24
3TEF1CAT3F	gtcctcgccgccctgcagcacgccATGCCCTCCCGCGCCAGAC	25

3TEF1CAT3R	aagaagcagctcgtgaggatataTCACAGCTCCACCGTCG	26
3TEF1CAT4F	cgtcctcgccgccctgcagcacgccATGCTCCTGTGCGGCACCG	27
3TEF1CAT4R	aagaagcagctcgtgaggatataCTACAGCTTCGCCTTGGCC	28
3TEF1HAA1F	tcgccgccctgcagcacgccATGGTCCTCATCGACGGG	29
3TEF1HAA1R	aagaagcagctcgtgaggatataTCAGCCGAAAGCGCCGCT	30
DGA1F1	cctgcaatcatggccttaataTGACGGCTTGTCTCTCC	31
DGA1R1	tcgaccttctcaagcactagAACCAACTTTATCCAGCAGC	32

Table 2: Plasmids used in the Examples

Plasmid	Genotype and purpose
pKC2.RgGPD1-ACS1	Sp ^R , <i>CAR2L-P_{3GPD1}-ACS1-Hyg^R-CAR2R</i> , for overexpression of ACS1
pKC2.RgTEF1-CAT1	Sp ^R , <i>CAR2L-P_{3TEF1}-CAT1-Hyg^R-CAR2R</i> , for overexpression of CAT1
pKC2.RgTEF1-CAT2	Sp ^R , <i>CAR2L-P_{3TEF1}-CAT2-Hyg^R-CAR2R</i> , for overexpression of CAT2
pKC2.RgTEF1-CAT3	Sp ^R , <i>CAR2L-P_{3TEF1}-CAT3-Hyg^R-CAR2R</i> , for overexpression of CAT3
pKC2.RgTEF1-CAT4	Sp ^R , <i>CAR2L-P_{3TEF1}-CAT4-Hyg^R-CAR2R</i> , for overexpression of CAT4
pKC2.RgTEF1-HAA1	Sp ^R , <i>CAR2L-P_{3TEF1}-HAA1-Hyg^R-CAR2R</i> , for overexpression of HAA1
pKC2.RgGPD1-DGA1	Sp ^R , <i>CAR2L-P_{3GPD1}-DGA1-Hyg^R-CAR2R</i> , for overexpression of DGA1

- Acetyl-CoA synthetase CDS (*ACS1*) was amplified by RT-PCR using cDNA pool as the template and oligonucleotides 3GPD1ACS1F and 3GPD1ACS1R as the primers. Carnitine acetyltransferase 1-carnitine acetyltransferase genes, *CAT1*, *CAT2*, *CAT3* and *CAT4*, and transcriptional activator CDS (*HAA1*) were amplified using genomic DNA *R. toruloides* as the template. The *CAT1-CAT4* CDSs were amplified using primers 3TEF1CAT1F/ 3TEF1CAT1R, 3TEF1CAT2F/ 3TEF1CAT2R, 3TEF1CAT3F/3TEF1CAT3R and 3TEF1CAT4F/ 3TEF1CAT4R, respectively. The *HAA1* sequence was amplified using primers 3TEF1HAA1F/3TEF1HAA1R. Amplified *ACS1* PCR products was fused with pKCL2.RgGPD1 digested with NcoI and EcoRV by Gibson cloning to create plasmid pKC2.RgGPD1ACS1. Amplified *CAT1-4* and *HAA1* PCR products was fused with pKCL2.RgTEF1 digested with NcoI and EcoRV by Gibson cloning to create plasmid pKC2.RgTEF1CAT1-CAT4 and pKC2.RgTEF1HAA1 respectively. All plasmids were verified by sequencing. *DGA1* was amplified from pE1GLA2 [Liu, Y., et al., 2021] using primers DGA1F1 and DGA1R1 (Table 1) and fused with pKC2.RgTEF1CAT1 using Gibson assembly method.
- To explore the effects of double gene overexpression on lipid production, the DGA1 and HAA1 cassettes are released from the backbones by cutting with restriction enzyme PacI and PmeI. pKC2.3TEF1CAT1 was digested with SpeI, blunt-ended using T4 DNA

polymerase and then digested with Pac1. The above two fragments were ligated using Quick Ligation and transformed into *E. coli* XL1-Blue. The T-DNA vectors were transformed into *R. toruloides* strain C3 by *Agrobacterium tumefaciens*-mediated transformation (ATMT) as described previously [Liu, Y., et al., 2013; Liu, Y., et al., 2018]. Gene expression cassettes were site-specifically-integrated to the *CAR2* locus (encoding phytoene synthase/ lycopene cyclase) to eliminate the positional effect. Overexpression mutants were selected by albino phenotype and at least 3 biological replicates were applied in the assays [Liu, Y., et al., 2013; Liu, Y., et al., 2018].

1.4 Laboratory directed evolution to enhance acetate utilization

Directed evolution of C3 strain for enhancing acetate utilization and stress resistance was performed by adapting the culture against a gradual increase of acetate concentration in the culture medium. A single colony of C3 was inoculated into 5 mL of YPD in a 50 mL Falcon tube and culture for 24 hours at 30°C with 210 rpm agitation. The whole 5 mL culture was inoculated into 45 mL of GAAM12 and cells were cultured at 30°C, 210 rpm for 5 days. The culture was sub cultured into fresh media of GAAM12 for 4 more times for a total of 5 cycles. This adaptation process was repeated with 4 subcultures in GAAM24 (2% acetate and 4% glycerol); 5 subcultures in GAAM34 (3% acetate and 4% glycerol), 5 subcultures in GAAM44 (4% acetate and 4% glycerol); GAAM54 (5% acetate and 4% glycerol); 5 subcultures in GAAM64 (6% acetate and 4% glycerol). The final culture was streaked on a YPD plate and 22 colonies were assessed for biomass growth, lipid content and fatty acid profile in shake-flask cultures. Strains were named as C3-64x, which x represents various pure isolates tested for fatty acid content and profiles.

25

1.5 Fed-batch fermentation

2L-scale fermentation was done using the Biostat B Plus bioreactor system (Sartorius, Germany). Seed cultures were prepared by inoculating 1 mL pre-seed culture into 100 mL medium of interest in a 250 mL conical flask and cultured 28-30 °C for 40-48 hours with vigorous agitation. The pre-seed culture was made by inoculating a single colony into 5 mL YPD medium in a 50 mL Falcon tube and cultured at 28-30 °C with vigorous agitation for 18-24 hours.

35

Using acetate and glucose as the carbon source

For fermentation using acetate and glucose as the carbon sources, 100 mL of seed culture was inoculated aseptically into 1L AMP3-Y6 with additional supplements (30 g/L acetic acid, 5 g/L glucose, 5.38 g/L yeast extract, 5.7 g/L KH_2PO_4 , 2.0 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 6.6 g/L KNO_3 , 15 mg/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.1 g/L L-Histidine, 10 mg/L L-Aspartic acid, neutralized to pH 7 with NaOH) in a 2L vessel. The pH of the culture was maintained at 7, using only 100% acetic acid. Base pump for the controller has been disabled. pO_2 was maintained at 30%, cascade-controlled with the stirrer (maximum 1200 rpm) and air-flow (maximum 3 L/min). At 40-48 hours after inoculation when growth rate of culture slowed down significantly, 100 mL of Acetate Feeding Medium (AFM, containing 30 g/L acetic acid, 5 g/L glucose monohydrate, 100 g/L yeast extract, 20 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 66 g/L KNO_3 , 150 mg/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g/L histidine, 0.1 g/L aspartic acid, 100 ml/L AAM vitamin stock solution) was pumped into the vessel at flow rate of 0.9 mL/min. This process was repeated every 40 to 48 hrs when acetic consumption levelling off. The levels of glucose and acetic acid were monitored by HPLC every 24 hr and maintained at 30 and 5 g/L respectively by topping up with stock solution of 80% glucose monohydrate solution and 50% acetic acid, respectively.

Using glucose as the carbon source

In glucose only fermentation, 100 mL of the seed culture was aseptically inoculated into 1 L of APM2 in a 2 L vessel. The pH of the culture was maintained at 5.5, using 3 M HCl and 25% NH_4OH . The pO_2 was maintained at 30%, cascaded firstly by the stirrer (maximum 1200 rpm) and secondly by airflow (maximum 3 L/min). Continuous feeding was set at 4.17 mL/h for the first 24 hours to reach an OD_{600} of 160 to 200 units. Glucose levels were monitored by HPLC and maintained at 40 to 60 g/L. The flow-rate of the continuous feeding medium was adjusted to 8.34 mL/h for the next 18-24 hours. At 40-48 hours after inoculation, OD_{600} of the culture should reach at least 320 units. To maintain a high growth rate, approximately 20% of the culture volume was removed aseptically and topped up with the Replacement Medium. This "bleeding and feeding" process was repeated every 6 hours.

1.6 Analytical methods

Biomass and lipid content were determined as previously described [Liu, Y., et al., 2018]. The concentration of sugars, acetate and glycerol were determined by HPLC as previously described [Liu, Y., et al., 2018]. HPLC equipped with Animex HPX-87H column (Bio-Rad Laboratories, Inc., USA) and Refractive Index Detector were applied. The samples were

analysed using 5 mM of sulfuric acid as a mobile phase and a flow rate of 0.7 mL/min. The column temperature was maintained at 50°C.

Fatty acid profiles were analysed using GCMS. A HP-88 fused silica capillary column (Agilent Technologies, Inc., USA) fitted in GCMS QP2010 Ultra (Shimadzu, USA) was used.

5 Samples were run with helium as the carrier gas which was set with a pressure of 8.9 psi and flow rate of 16.0 mL/min. The GC column flow was set at 1.00 mL/min. Oven temperature started at 70°C, held for 1 minute, and increased 25°C per min to 200°C. After holding for 1 minute, the temperature was increased 5°C per minute to 230°C. After holding for 1 minute, the temperature was increased from 50°C to 250°C and held for 3 minutes. The
10 total run time for each sample was 17.6 minutes.

Protein content was analysed by Dumas method. Elemental N of freeze-dried biomass was quantified using a Perkin Elmer Avio 500 Inductively Coupled Plasma-Optical Emission Spectrometer (ICP-OES). Protein content was calculated from the nitrogen content using a
15 conversion factor of 6.25 [Hayes, M., 2020].

Example 2

Identification of an elite acetate-utilizing yeast strain

Although *R. toruloides* strains showed relatively poor performance in acetate utilization [Gong, Z., et al., 2015], we were intrigued if there were significant microbial strain diversity in
20 acetate utilization. We cultured 4 *Rhodotorula* strains, one *Sacharomyces cerevisiae* and one *Pichia pastoris* strain in acetate assimilation medium with 3% acetic acid as the sole carbon source. *R. toruloides* strain C3 showed a biomass titer that was close to the best acetate utilizing yeast species reported [Gong, Z., et al., 2015]. Its biomass titer was more
25 than twice that of ATCC 10788 and 7 times that of strain ATCC 10657. Notably, all *Rhodotorula* strains showed much better growth in the medium than *S. cerevisiae* and *P. pastoris*. Therefore, strain C3 was selected for further studies.

Medium optimization for acetate utilization

Several medium compositions reported for yeast lipid production were modified, with acetic
30 acid replacing glucose as the carbon source. As shown in Figure 2a, dry cell biomass was the highest when cultured in the modified APM medium followed by B2001 medium [Bhosale, P. and R.V. Gadre, 2001]. The former yielded 2.6 times as much cell biomass as the similarly substituted Y4 medium [Li, Y., Z. Zhao, and F. Bai, 2007]. The modified Yeast Nitrogen Base (YNB) medium yielded the lowest cell biomass. Therefore, the APM

composition was selected for further optimization due to its potential cost advantage over others. As this medium contains 5.7 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, which formed a precipitate with acetate, we were intrigued to see if this could be reduced. Cell mass production was marginally enhanced when $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ as reduced to 0.5-2% (Fig. 5). Subsequently, the effect of medium pH was investigated by culturing C3 cells in this modified APM medium, hereafter referred to as APM3 which is composed of 6.5 g/L yeast extract, 5.7 g/L KH_2PO_4 , 1 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.015 g/L $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 30 g/L acetic acid and pH adjusted to 7.0 using NaOH. A preliminary scanning of pH effect showed that C3 preferred neutral pH conditions (Fig. 6). This was confirmed in another set of experiments with medium pH set between 6.5 and 8.5. The results showed that pH 7.0 was the best for cell biomass production (Figure 2b). Therefore, the neutralized APM3 medium (pH 7.0, 3% acetic acid) was selected for further studies. This result is different from *Y. lipolitica*, which was found to grow better at alkaline pH conditions [Gao, R., et al., 2020].

The C:N ratio and nitrogen source are important factors affecting cell growth and lipid accumulation [Lopes, H.J.S., et al., 2020; Weng, L.-C., et al., 2014]. To determine the optimal nitrogen source for strain C3 growth, the optimized APM3 composition described above was partially substituted with different levels of urea, NH_4Cl and KNO_3 . As shown in Figure 2c, the highest biomass (10.0 g/L) and lipid (4.3 g/L) production were observed in medium with 6.63 g/L KNO_3 in combination with 6.545 g/L yeast extract. With the same concentration of yeast extract in the medium, both the biomass and lipid production in the medium with 6.63 g/L KNO_3 was more than twice as much as that in the medium with 2 g/L urea or 3.77 g/L NH_4Cl . In particular, the biomass of strain C3 grown in medium with 6.6 g/L KNO_3 in combination with 6.55 g/L yeast extract was about five times as much as that grown in medium with either 2 g/L urea or 3.77 g/L NH_4Cl in combination with 6.55 g/L yeast extract. Similarly, lipid production in medium with 6.63 g/L KNO_3 in combination with 6.45 g/L yeast extract was more than four times as much as that grown in medium with either 2 g/L urea or 3.68 g/L NH_4Cl in combination with 6.55 g/L yeast extract. Biomass production was positively correlated with yeast extract input although lipid titer is only about 10% higher if yeast extract was increased by 50%, from 4.3 g/L (AMP3-Y6) to 6.5 g/L (AMP3-Y9) (Fig. 2c).

Lipid production is an energy intensive bioprocess, demanding a high level of carbon source. To see if a higher level of acetate can be tolerated in C3 cell culture, the strain was cultured in media with five different concentrations of acetic acid, from 3% to 7% (w/v). Although slightly more biomass and lipid were produced with 4% acetic acid than with 3% acetic acid in the 5-day culture, biomass was significantly higher with 3% acetate in the 3-day culture,

suggesting faster cell growth in 3% acetate or below (Fig. 2d). Cell growth was severely inhibited when acetate was above 5%.

In preferred embodiments the *R. toruloides* C3 strain is grown in AMP3-Y6 medium pH 7 (4.3 g/L yeast extract, 6.63 g/L KNO₃, 5.7 g/L KH₂PO₄, 1 g/L MgSO₄·7H₂O, 0.015 g/L FeCl₃·6H₂O, 30 g/L acetic acid, pH 7.0).

Supplementation of co-substrate is an effective strategy to enhance acetate utilization [Chen, L., et al., 2021]. Acetate assimilation medium (AAM) described previously was tested in some of our early experiments. In this medium, C3 strain grew poorly when acetate was added at 2% or 3% (w/v). Glycerol at 2% (w/v) significantly enhanced biomass production although it became inhibitory at higher concentrations (Fig. 7). The positive effect of glycerol on acetate utilization was confirmed in the APM2 medium composition that was supplemented with glycerol as acetate was fully depleted on day 3. It took a long time to fully consume acetate when other co-substrates were used (Fig. 8). Furthermore, we tested the effects of amino acids and trace elements in the AMP3-Y6 formulation. Overall, the supplementation had negligible effect except for L-histidine (1 g/L) and L-aspartic acid (0.1 g/L) (data not shown).

EXAMPLE 3

Directed evolution to enhance acetate and glycerol utilization

To see whether C3 strain could be enhanced for acetate utilization by laboratory-directed evolution, the strain was cultured in increasing concentration of acetate. Starting in GAAM with 1% acetate and 2% glycerol as co-substrate, referred to as GAAM 12 medium, acetate concentration was increased stepwise after 5 cycles adaptation in any given medium. The carbon source was increased from 1% acetic acid with 2% glycerol to 6% acetic acid with 4% glycerol in a total of 25 cycles of cultures in 5 sets of media. We compared the growth of the 25th generation to the Wt C3 strain in different acetic acid concentrations. Significant improvement in cell optical density and acetate consumption were observed in all four media tested (Fig. 3). When glycerol was kept at 4%, acetate consumption was increased by 32.9%, 85.1%, 25.2% and 174.3% in medium with 2%, 3%, 4%, 6% acetic acid, respectively. The cell optical densities were also increased significantly. Notably, the fastest cell growth was obtained with 2-3% acetic acid loading.

We analyzed the cell growth and lipid production of 22 isolates of the 25th generation in a glucose based medium that we normally used for lipid production [Liu, Y., et al., 2021]. Most strains grew slower in this glucose-based medium, with lower cell biomass and lipid content in the 5-day culture (Fig. 9). Surprisingly, most strains have 4-6 fold higher alpha-linolenic acid (ALA) content, an omega-3 fatty acid that is nutritionally valuable. ALA reached more than 15% of total fatty acids in the cells in fed-batch fermentations.

EXAMPLE 4

Metabolic engineering to enhance lipid production from acetate

Acetyl-CoA is a metabolic intermediate that is central for many biochemical reactions, including pathways for protein, carbohydrate, lipid and sterol biosynthesis. Over-expression of acetyl-CoA synthetase (ACS) enhances iso-butanol production from acetate in *E. coli* as it enhances acetyl-CoA production by directly coupling acetate to CoA at the cost of one ATP [Song, H.S., et al., 2018]. In *S. cerevisiae*, overexpression of transcription factor Haa1 enhances acetate resistance and utilization [Swinnen, S., et al., 2017]. Acetate kinase-phosphotransacetylase (ACK-PTA) makes the 2nd route for acetate utilization [Kiefer, D., et al., 2021]. Carnitine acetyltransferase (CAT) shuttles acetyl units between cytoplasm, mitochondria and peroxisomes by reversibly converting acetyl-CoA to acetyl-carnitine [Hynes Michael, J., et al., 2011; Rong, L., et al., 2022]. Most eukaryotic cells have multiple isoforms of this enzyme, suggesting the vital role this enzyme plays. The *S. cerevisiae* genome encodes three CAT genes but only *CAT2* contributes to stress resistance to small organic acids and oxidative agents [Franken, J., et al., 2008]. By BLAST search of the *R. toruloides* transcriptomes using the *S. cerevisiae* CAT as the queries, four different CAT homologs and one obvious homolog for *ACS1* and *HAA1* were identified (*CAT1*, SEQ ID NO: 2; *CAT2*, SEQ ID NO: 6; *CAT3*, SEQ ID NO: 8; *CAT4*, SEQ ID NO: 10; *ACS1*, SEQ ID NO: 16; *HAA1*, SEQ ID NO: 4).

To enhance acetate utilization in *R. toruloides*, the homologs for *CAT*, *ACS1* and *HAA1* were individually over-expressed by fusing the CDS between the Rg*GPD1* promoter or the intron-containing Rg*TEF1* promoter of *R. graminis* WP1 strain and the respective transcriptional terminator [Liu, Y., et al., 2021; Liu, Y., et al., 2016]. The constructs were site-specifically inserted at the *CAR2* locus. Little improvement in cell growth was observed with most constructs if the resulted strains were cultured in the optimized acetate media, such as the APM3-Y6 with 3% acetic acid (Fig. 4a). Clear enhancement of cell growth was observed when acetic acid was raised to 4-5% and when the nutrient poor AAM medium was used (Fig. 4b). Overexpression of *CAT1* showed the strongest enhancing effect, followed by

HAA1 while *CAT3* showed the weakest effect in flask cultures. Cell biomass and lipid production were also improved by the overexpression of *CAT1* and *HAA1* although no additive effect was observed for the two genes (Fig. 10). Overexpression of *CAT1* and *HAA1* led to 7.2% and 9.1% improvement in biomass titer; and 14.5% and 5.1% improvement in lipid content, respectively. Thus, *CAT1* and *HAA1* were effective targets for enhancing acetate utilization in *R. toruloides*. The differences in the activity of the four CAT enzymes is not clear at present.

Similar to a previous report for *Y. lipolytica*, *DGA1* overexpression significantly enhanced cell growth. *ACS1* overexpression had weak effect on cell growth on acetate, be it targeted to cytoplasm or mitochondria. In AAM medium with 3% acetate, both Wt and *ACS1*-overexpressing cells showed a lag phase of growth of about 24 hours and only marginal improvement in OD₆₀₀ nm was observed in the *ACS1*-overexpressing cells at the end of 5 days culture. On the other hand, over-expressing the *ACS1* variant with a mitochondria transit peptide of the malate dehydrogenase 2 (MDH2) (MAAATRQSSRLARSFSTSARAN; SEQ ID NO: 15) effectively relieved the lag phase. Surprisingly, the OD₆₀₀ of this strain levelled off much earlier and the final cell density ended up significantly lower than Wt (Fig. 11). This data suggest that *ACS1* overexpression is not an effective strategy to enhance acetate utilization in *R. toruloides*.

Surprisingly, overexpression of the CAT genes resulted in significantly higher content of C18:2 and C18:3 (ALA) (Table 3).

Table 3: Effect of CAT overexpression on lipid composition

Fatty acid profile (% FA)	pKCL2	CAT1	CAT2	CAT3	CAT4	
C16 (%)	35.58±3.01	24.22±0.4	24.27±1.08	26.15±0.74	25.33±1.59	
C18 (%)	14.67±1.11	13.94±1.57	14.94±1.84	14.88±1.62	14.7±0.96	
C18:1(%)	39.61±3.57	45.19±2.43,	44.68±1.16	44.54±1.77	44.73±1.36	
C18:2(%)	8.1±0.83	13.32±1.03***	12.57±0.17***	11.88±0.67***	12.63±0.95***	
C18:3(%)	2.04±0.12	3.33±0.51***	3.53±0.18***	2.54±0.28**	2.6±0.47	

Note: Three different colonies of each construct were inoculated into 5 ml YPD in 50 ml Falcon tube and incubated at 28°C overnight. Cells were diluted to 0.05 OD₆₀₀ in 50 ml AMP3-Y6 medium with 3% acetic acid in 250 ml conical flask and cultured for 5 days at 30°C. CAT1, CAT2, CAT3 and CAT4 refer to C3CK strain inserted with the respective CAT

gene fused with the RgTEF promoter and knocked in at the *CAR2* locus. pKCL2 is the strain with control construct. **: $p < 0.05$; ***: $p < 0.01$

The average ALA content of samples from all time-points was 2.58% of total fatty acids for pKCL2 control strain, compared with 9% in *CAT1*-OE strain, an increase of 249%.

An *in silico* analysis of subcellular targeting using the TargetP - 2.0 software program suggests that CAT1, CAT2 and CAT3 is peroxisomal, mitochondrial and cytoplasmic, respectively, while CAT4 would be dual-targeted to peroxisomes and mitochondria. Thus, CAT4 is the homolog for the *Candida albicans* CTN2 or the *Aspergillus nidulans* ACUJ, which is dual targeted to peroxisomes and mitochondria due to the presence of a N-terminal MTS and a C-terminal PTS1 [Hynes Michael, J., et al., 2011]. Our work suggests, for the first time, that overexpression of CAT, particularly the peroxisomal isozyme, is an effective strategy to balance acetyl-CoA levels in different cellular compartments, leading to enhanced acetate utilization and lipid production in *R. toruloides*. In addition, acetate resistance and metabolism are regulated by the conserved HAA1 transcriptional factor in *R. toruloides*.

To see if peroxisome targeting is important for the observed CAT activity (Fig. 4), the PTS of CAT1 was deleted. The CAT1- Δ AKL marginally lost its activity in cell growth promotion in acetate medium. Both CAT and CAT1- Δ AKL overexpression significantly increased cell mass and lipid content (Fig. 12). This suggests CAT1 exerts its function both in cytoplasm and peroxisome. The weak effect of PTS deletion may suggest the peroxisomal and mitochondrial CAT activities encoded by other 3 homologs are strong enough to mask the CAT1 mutation.

Our work suggests, for the first time, that overexpression of CAT, particularly the peroxisome or mitochondria targeted isozymes like CAT1 and CAT4, is an effective strategy to balance acetyl-CoA levels in different cellular compartments, leading to enhanced acetate metabolism, cell growth and lipid production in *R. toruloides*. In addition, acetate resistance and metabolism are regulated by the conserved HAA1 transcriptional factor in *R. toruloides*. In preferred embodiments, *R. toruloides* is genetically engineered to overexpress CAT1 and/or HAA1 genes for enhanced acetate utilization and lipid productions compared to wildtype *R. toruloides*. In some embodiments, the genes are overexpressed under the influence of a promoter such as RgTEF1 or RgGPD1. More preferably, *R. toruloides* C3 strain is genetically engineered.

Fed-batch fermentation

Studies on *R. toruloides* have been focusing on lipid production due to its high lipid productivity and lipid content. Like most oleaginous microorganisms, nitrogen starvation is a condition for inducing lipid biosynthesis [Ren, H.-Y., et al., 2013; Donzella, S., et al., 2019].

5 With the increasing interest to use SCP for food and animal feed applications [Jach, M.E., et al., 2022], we investigated the possibility of SCP production from *R. toruloides* C3 strain. After a series of process optimization, dry cell mass reached 98.3 g/L in 44 hours from an initial cell density of about 1 OD₆₀₀ unit. Cell mass productivity was 3.1 g/L/h in the first 20 h and slowed to 1.7 g/L/h in the next 24 h (Fig. 13a). To improve SCP productivity, semi-

10 continuous fermentation mode [Yuan, Q., et al., 2022; Carlozzi, P. and E. Touloupakis, 2021] was investigated. We found nitrogen type and content were crucial for cell mass productivity. Cell biomass titre kept decreasing when the Replacement Medium contained 10.6 g/L yeast extract while increasing the yeast extract to 13.0 g/L reversed the downward trend. Inclusion of urea (at 4.3-6.9 g/L) further increased the biomass titre. However, urea was less effective

15 in maintaining the high biomass productivity and dry cell mass titre. Dry cell mass productivity could be maintained above 6 g/L/h when 17.2 g/L yeast extract was used in the RM (Fig. 13b). Under such conditions, dry cell mass titres were above 100 g/L and glucose consumption per gram biomass produced could be significantly reduced, from about 2.5 g/g to less than 2 g/g. The average lipid and protein content of cells ranged from 10% - 15% and

20 40% - 53%, respectively (Figs. 13 and 14). Fermentation results of other strains are summarized in Table 4.

Table 4: Comparison of fermentation parameters of various yeasts and carbon source

Strain	Mode	Max Biomass Titer (g/L)	Max Lipid Content (%)	Average ALA Content (%)	Average Lipid Yield ¹ (g/g)	Average Biomass Yield ¹ (g/g)	Max Lipid Yield ² (g/g)	Max Biomass Yield ² (g/g)	Lipid productivity (g/Lh)	Biomass productivity (g/Lh)	Reference
Rhodotorula toruloides C3	Glucose Semi-continuous	115.7	22.01	4.20	0.04	0.48	0.12	0.63		2.03-7.18	This study
Rhodotorula toruloides C3	Acetate:glucose (6:1) Fed-batch	151.5	19.3	7			0.07			0.61	This study
Rhodotorula toruloides C3-64B	Acetate:glucose (6:1) Fed-batch	118.3	16	16.6			0.02			0.8	This study
Rhodotorula toruloides C3CK CAR2::RgTEF:Luc2	Acetate:glucose (6:1) Fed-batch	125	16.2	2.58	0.09	0.68		0.88		0.52	This study
Rhodotorula toruloides C3CK CAR2::RgTEF:CAT1	Acetate:glucose (6:1) Fed-batch	156.5	19.8	9%	0.06	0.55		0.55		0.72	This study

Cryptococcus curvatus ATCC 20509	Acetate Fed-batch	48.7	60		0.18						[50]
Cryptococcus curvatus MUCL 29819	Acetate Fed-batch		50.1		0.15						[51]
Rhodotorula glutinis CGMCC 2.703	Acetate-glucose Fed-batch	68.2	52.5		0.1				0.186		[49]
Yarrowia lipolytica MUCL28849	Acetate-glucose Fed-batch	30.8	40.7		0.13				0.16		[52]
Yarrowia lipolytica MTYLD65 ³	3% acetate semicontinuous	194	115		0.16				0.8	1.35	[19]

Note:

1. The average lipid or biomass yield refer to the average product produced per gram of substrate (glucose or acetic acid) consumed. In this study, the input for glucose, which is about 14.7% of total substrate used, has not been included in the calculation.
2. The maximal lipid or biomass yield refers to maximal value between any two sampling points (about 24 hours).
3. Po1g strain overexpressing *ACC1* and *DGAI*.
4. NA: not applicable

10 Summary

SCO and SCP produced from microorganisms are increasingly being regarded as alternative food staples for the future. Cell biomass and lipid yields (g/L) and productivities (g/L/h⁻¹) are strongly dependent on microbial genetics, culture medium composition and environmental conditions [Reihani, S.F.S. and K. Khosravi-Darani, 2019]. Filamentous *F. venenatum* was reported to have a biomass yield of about 0.35 g/g using date biomass-derived sugar as the feedstock [Hosseini, S., et al., 2009]. The biomass yield of Crabtree positive *S. cerevisiae* is much lower (0.12 g/g) [Reihani, S.F.S. and K. Khosravi-Darani, 2019]. Crabtree negative yeasts, such as *Kluyveromyces fragilis* and *Candida utilis*, have a strong advantage for biomass and SCP production. Biomass yield of 0.74 g/g and 0.65 g/g, respectively, was reported using cheese whey and defatted rice polishing as the feedstock [Reihani, S.F.S. and K. Khosravi-Darani, 2019].

R. toruloides is a Crabtree negative oleaginous yeast that can utilize hexose, pentose, glycerol and a variety of fatty acids as the carbon source [Sitepu, I.R., et al., 2014; van der

Hoek, S.A., et al., 2022]. Our biomass and lipid productivity data suggest that the land productivity of a microbial cell factory based on *R. toruloides* would be thousands more productive than land crop when glucose is used as the feedstock. Therefore, harnessing the natural high productivity of *R. toruloides* could contribute significantly to enhance food security.

In contrast to a previous report regarding *R. toruloides* acetate utilization [Gong, Z., et al., 2015], we showed that *R. toruloides* strain C3 was able to utilize acetate as efficiently as other microbes, such as *Trichosporon cutaneum*, *Y. lipolytica* and *Cryptococcus curvatus* [Gong, Z., et al., 2015; Gao, R., et al., 2020]. Acetate is one of the promising future feedstocks for SCO and SCP production owing to its natural abundance, low toxicity and technical feasibility of production using CO₂ and H₂. Furthermore, acetate utilization in *R. toruloides* was enhanced through metabolic engineering and directed evolution. Amongst the six genes tested, *CAT1* was the most effective in enhancing acetate utilization, leading to significantly better biomass and lipid production. In addition, the transcriptional factor Haa1 was equally effective. Our approach is different from previous reports in *Y. lipolytica*, which targeted acetyl-CoA synthetase (ACS), acetyl-CoA carboxylase (ACC) and fatty acid synthase (FAS). Overexpression of the 3 genes led to 25.7% increase in lipid content when using acetate and glycerol co-substrate in fed-batch fermentation, achieving a maximum OD₆₀₀ of 29.2 and a lipid content of 41.7% [Chen, L., et al., 2021].

R. toruloides appears to require significantly different conditions for optimal growth. The effect of pH on lipid production by *R. glutinis* CGMCC 2.703 (re-named as *R. toruloides*) has been studied. The optimal pH for cell growth is 6.6 while pH 7.0 was found optimal for lipid production. Biomass and lipid titre of 67.2 g/L and 29.3 g/L was achieved, respectively, in pH-stat fed-batch fermentation of acetate. The biomass yield, lipid titer, and lipid content was increased to 68.2 g/L, 35.8 g/L, and 52.5%, respectively, using a two-stage pH regulation culture strategy [Zhang, W., et al., 2019]. Our results with *R. toruloides* C3 are quite similar to *R. glutinis* CGMCC 2.703, preferring neutral pH while *Y. lipolytica* prefers slightly alkaline conditions (pH 8-9). Notably, *R. toruloides* C3 had superior lipid yield from acetate. Maximal lipid yield (ie, substrate conversion rate) was about 0.3 g/g although the current average lipid yield was only 0.13 (Table 4; Fig. 14).

Dry cell mass for the *CAT1*-over-expressing strain was 156.5 g/L, much better than that of *R. glutinis* CGMCC 2.703 and *Y. lipolytica* also (37 g/L in small scale fed-batch fermentation) (Gao et al., 2020). This lipid yield is lower than the 0.16 g/g reported in the *Y. lipolytica* strain

overexpressing *ACC1* and diacylglycerol acyltransferase 1 (*DGA1*) [Xu, J., et al., 2017]. We noticed that glycerol was produced as a significant by-product, in both the CAT4 and C3-64 strains (Fig. 14G). This suggests that certain steps in the coupling of mono/di-acylglycerol with acy-CoA were blocked. Overexpression of *DGA1* was not effective in enhancing growth or lipid production in acetate medium (data not shown). The limiting step for TAG synthesis remains unknown at present.

We have demonstrated that laboratory-directed evolution using acetate and glycerol as co-substrate is able to increase acetate resistance; enhance utilization and decrease the glycerol by-product accumulation. Strain C3-64B produces high ALA content compared to Wt C3 strain and CAT1-overexpression strain because strain C3-64B could utilize glycerol as a carbon source. Unexpectedly, this adaptation strategy was able to increase ALA content. In 2L scale fermentation, ALA level was raised from about 4% up to 17.33% of total fatty acid when acetate was used as the main carbon source. Furthermore, the fatty acid profile can be improved by culturing *R. toruloides* cells under increasing acetate-glycerol ratio. Notably, α -linolenic acid content could be significantly increased using this approach.

References

- Béligon, V., et al., Improvement and modeling of culture parameters to enhance biomass and lipid production by the oleaginous yeast *Cryptococcus curvatus* grown on acetate. *Bioresource Technology*, 2015. **192**: p. 582-591.
- Bhosale, P. and R.V. Gadre, Optimization of carotenoid production from hyper-producing *Rhodotorula glutinis* mutant 32 by a factorial approach. *Letters in Applied Microbiology*, 2001. **33**(1): p. 12-16.
- Bratosin, B.C., S. Darjan, and D.C. Vodnar Single Cell Protein: A Potential Substitute in Human and Animal Nutrition. *Sustainability*, 2021. **13**, DOI: 10.3390/su13169284.
- Cai, P., et al., Methanol biotransformation toward high-level production of fatty acid derivatives by engineering the industrial yeast *Pichia pastoris*. *Proceedings of the National Academy of Sciences*, 2022. **119**(29): p. e2201711119.
- Carlozzi, P. and E. Touloupakis, Bioplastic production by feeding the marine *Rhodovulum sulfidophilum* DSM-1374 with four different carbon sources under batch, fed-batch and semi-continuous growth regimes. *New Biotechnology*, 2021. **62**: p. 10-17.
- Chawla, K., et al., Metabolic engineering of oleaginous yeasts to enhance single cell oil production. *Journal of Food Process Engineering*, 2022. **45**(7): p. e13634.

- Chen, L., et al., Increased Lipid Production in *Yarrowia lipolytica* from Acetate through Metabolic Engineering and Cosubstrate Fermentation. *ACS Synthetic Biology*, 2021. **10**(11): p. 3129-3138.
- Christophe, G., et al., Production of Oils from Acetic Acid by the Oleaginous Yeast *Cryptococcus curvatus*. *Applied Biochemistry and Biotechnology*, 2012. **167**(5): p. 1270-1279.
- Demir, M. and A.G. Gündes, Single-cell oil production by *Mortierella isabellina* DSM 1414 using different sugars as carbon source. *Biotechnology Progress*, 2020. **36**(6): p. e3050.
- Donzella, S., et al., Engineering cytoplasmic acetyl-CoA synthesis decouples lipid production from nitrogen starvation in the oleaginous yeast *Rhodospiridium azoricum*. *Microbial Cell Factories*, 2019. **18**(1): p. 199.
- Espinosa, M.I., et al., Adaptive laboratory evolution of native methanol assimilation in *Saccharomyces cerevisiae*. *Nature Communications*, 2020. **11**(1): p. 5564.
- Fang, X., et al., A three-dimensional hybrid electrode with electroactive microbes for efficient electrogenesis and chemical synthesis. *Proceedings of the National Academy of Sciences*, 2020. **117**(9): p. 5074-5080.
- Fontanille, P., et al., Bioconversion of volatile fatty acids into lipids by the oleaginous yeast *Yarrowia lipolytica*. *Bioresource Technology*, 2012. **114**: p. 443-449.
- Franken, J., et al., Carnitine and carnitine acetyltransferases in the yeast *Saccharomyces cerevisiae*: a role for carnitine in stress protection. *Current Genetics*, 2008. **53**(6): p. 347.
- Gao, R., et al., Enhanced lipid production by *Yarrowia lipolytica* cultured with synthetic and waste-derived high-content volatile fatty acids under alkaline conditions. *Biotechnology for Biofuels*, 2020. **13**(1): p. 3.
- Gao, J., et al., Rescuing yeast from cell death enables overproduction of fatty acids from sole methanol. *Nature Metabolism*, 2022. **4**(7): p. 932-943.
- García Martínez, J.B., et al., Methane Single Cell Protein: Potential to Secure a Global Protein Supply Against Catastrophic Food Shocks. *Frontiers in Bioengineering and Biotechnology*, 2022. **10**.
- Gong, Z., et al., Efficient conversion of acetate into lipids by the oleaginous yeast *Cryptococcus curvatus*. *Biotechnology for Biofuels*, 2015. **8**(1): p. 189.
- Hayes, M., Measuring Protein Content in Food: An Overview of Methods. *Foods*, 2020. **9**(10): p. 1340.
- Hosseini, S., et al., Production of mycoprotein by *Fusarium venenatum* growth on modified vogel medium. *Asian Journal of Chemistry*, 2009. **21**(5): p. 4017.

- Hu, C., et al., Effects of biomass hydrolysis by-products on oleaginous yeast *Rhodospiridium toruloides*. *Bioresource Technology*, 2009. **100**(20): p. 4843-4847.
- Hynes Michael, J., et al., Role of Carnitine Acetyltransferases in Acetyl Coenzyme A Metabolism in *Aspergillus nidulans*. *Eukaryotic Cell*, 2011. **10**(4): p. 547-555.
- 5 Jach, M.E., et al., Yeast protein as an easily accessible food source. *Metabolites*, 2022. **12**(1): p. 63.
- Jiang, W., et al., Metabolic engineering strategies to enable microbial utilization of C1 feedstocks. *Nature Chemical Biology*, 2021. **17**(8): p. 845-855.
- Kiefer, D., et al., From Acetate to Bio-Based Products: Underexploited Potential for
10 Industrial Biotechnology. *Trends in Biotechnology*, 2021. **39**(4): p. 397-411.
- Koh, C.M.J., Liu, Y., Moehninsi et al. Molecular characterization of KU70 and KU80 homologues and exploitation of a KU70-deficient mutant for improving gene deletion frequency in *Rhodospiridium toruloides*. *BMC Microbiol* 14, 50 (2014).
- Li, Y., Z. Zhao, and F. Bai, High-density cultivation of oleaginous yeast
15 *Rhodospiridium toruloides* Y4 in fed-batch culture. *Enzyme and Microbial Technology*, 2007. **41**(3): p. 312-317.
- Liu, Y., et al., Developing a set of strong intronic promoters for robust metabolic engineering in oleaginous *Rhodotorula* (*Rhodospiridium*) yeast species. *Microbial Cell Factories*, 2016. **15**(1): p. 200.
- 20 Liu, Y., et al., Understanding and exploiting the fatty acid desaturation system in *Rhodotorula toruloides*. *Biotechnology for Biofuels*, 2021. **14**(1): p. 73.
- Liu, Y., et al., Characterization of glyceraldehyde-3-phosphate dehydrogenase gene RtgPD1 and development of genetic transformation method by dominant selection in oleaginous yeast *Rhodospiridium toruloides*. *Applied Microbiology and Biotechnology*, 2013.
25 **97**(2): p. 719-729.
- Liu, Y., et al., Identification of novel genes in the carotenogenic and oleaginous yeast *Rhodotorula toruloides* through genome-wide insertional mutagenesis. *BMC Microbiology*, 2018. **18**(1): p. 14.
- Lopes, H.J.S., et al., C/N ratio and carbon source-dependent lipid production profiling
30 in *Rhodotorula toruloides*. *Applied Microbiology and Biotechnology*, 2020. **104**(6): p. 2639-2649.
- Maina, S., et al., Microbial oil production from various carbon sources by newly isolated oleaginous yeasts. *Eng Life Sci*, 2017. **17**(3): p. 333-344.
- Mhlongo, S.I., et al., The Potential of Single-Cell Oils Derived From Filamentous
35 Fungi as Alternative Feedstock Sources for Biodiesel Production. *Frontiers in Microbiology*, 2021. **12**.

- Nyyssölä, A., et al., The role of single cell protein in cellular agriculture. Current Opinion in Biotechnology, 2022. **75**: p. 102686.
- Ochsenreither, K., et al., Production Strategies and Applications of Microbial Single Cell Oils. Frontiers in Microbiology, 2016. **7**.
- 5 Park, Y.-K., J.-M. Nicaud, and R. Ledesma-Amaro, The Engineering Potential of *Rhodospiridium toruloides* as a Workhorse for Biotechnological Applications. Trends in Biotechnology, 2018. **36**(3): p. 304-317.
- Reihani, S.F.S. and K. Khosravi-Darani, Influencing factors on single-cell protein production by submerged fermentation: A review. Electronic Journal of Biotechnology, 2019. **37**: p. 34-40.
- 10 Ren, H.-Y., et al., A new lipid-rich microalga *Scenedesmus* strain R-16 isolated using Nile red staining: effects of carbon and nitrogen sources and initial pH on the biomass and lipid production. Biotechnology for Biofuels, 2013. **6**(1): p. 143.
- Ritala, A., et al., Single Cell Protein—State-of-the-Art, Industrial Landscape and Patents 2001–2016. Frontiers in Microbiology, 2017. **8**.
- 15 Rong, L., et al., Engineering *Yarrowia lipolytica* to Produce Itaconic Acid From Waste Cooking Oil. Frontiers in bioengineering and biotechnology, 2022. **10**: p. 888869-888869.
- Sitepu, I.R., et al., Oleaginous yeasts for biodiesel: current and future trends in biology and production. Biotechnol Adv, 2014. **32**(7): p. 1336-1360.
- 20 Song, H.S., et al., Enhanced isobutanol production from acetate by combinatorial overexpression of acetyl-CoA synthetase and anaplerotic enzymes in engineered *Escherichia coli*. Biotechnology and Bioengineering, 2018. **115**(8): p. 1971-1978.
- Swinnen, S., et al., Improvement of yeast tolerance to acetic acid through Haa1 transcription factor engineering: towards the underlying mechanisms. Microbial Cell Factories, 2017. **16**(1): p. 1-15.
- 25 van der Hoek, S.A., et al., Engineering ergothioneine production in *Yarrowia lipolytica*. FEBS Letters, 2022. **596**(10): p. 1356-1364.
- Weng, L.-C., et al., Nitrogen Deprivation Induces Lipid Droplet Accumulation and Alters Fatty Acid Metabolism in Symbiotic Dinoflagellates Isolated from *Aiptasia pulchella*. Scientific Reports, 2014. **4**(1): p. 5777.
- 30 Xu, J., et al., Application of metabolic controls for the maximization of lipid production in semicontinuous fermentation. Proc Natl Acad Sci U S A, 2017. **114**(27): p. E5308-e5316.
- Yuan, Q., et al., Long-term semi-continuous acidogenic fermentation for food wastes treatment: Effect of high organic loading rates at low hydraulic retention times and uncontrolled pH conditions. Bioresource Technology, 2022. **357**: p. 127356.
- 35

Zhang, C., et al., Microbial Utilization of Next-Generation Feedstocks for the Biomanufacturing of Value-Added Chemicals and Food Ingredients. *Frontiers in Bioengineering and Biotechnology*, 2022. **10**. p. 48.

5 Zhang, W., et al., Enhanced lipid production by *Rhodotorula glutinis* CGMCC 2.703 using a two-stage pH regulation strategy with acetate as the substrate. *Energy Science & Engineering*, 2019. **7**(5): p. 2077-2085.

Zhao, Y., et al., *Rhodotorula toruloides*: an ideal microbial cell factory to produce oleochemicals, carotenoids, and other products. *World J Microbiol Biotechnol*, 2021. **38**(1): p. 13.

CLAIMS

1. A genetically engineered *R. toruloides* cell, wherein the cell has been transformed by at least one polynucleotide molecule;
the at least one polynucleotide molecule comprising a carnitine acetyltransferase (*CAT*)
5 selected from the group consisting of *CAT1*, *CAT2*, *CAT3* and *CAT4*; and/or transcription factor *HAA1* genes, operably linked to at least one promoter, wherein overexpression of *CAT* and/or *HAA1* provides increased acetate utilization and/or fatty acid and/or lipid and/or biomass production compared to wildtype *R. toruloides*.
- 10 2. The genetically engineered *R. toruloides* cell of claim 1, wherein the *CAT* and/or *HAA1* genes are operably linked to a constitutive promoter, such as a promoter selected from the group comprising *TEF1* and *GPD1*, to achieve a protein expression level that is at least 5 times higher, preferably 10, 100 or 100 times higher.
3. The genetically engineered *R. toruloides* cell of claim 1 or 2, wherein the cell is *R.*
15 *toruloides* C3 strain.
4. The genetically engineered *R. toruloides* cell of any one of claims 1 to 3, wherein the *CAT* gene is *CAT1* and encodes the amino acid sequence set forth in SEQ ID NO: 1.
5. The isolated genetically engineered *R. toruloides* cell of any one of claims 1 to 4, wherein the *HAA1* gene encodes the amino acid sequence set forth in SEQ ID NO: 3.
- 20 6. The genetically engineered *R. toruloides* cell of any one of claims 1 to 5, wherein the *CAT1* gene comprises a polynucleotide sequence having at least 80% sequence identity, due to the degeneracy of the genetic code, to the polynucleotide sequence set forth in SEQ ID NO: 2.
7. The genetically engineered *R. toruloides* cell of any one of claims 1 to 5, wherein the
25 *HAA1* gene comprises a polynucleotide sequence having at least 80% sequence identity, due to the degeneracy of the genetic code, to the polynucleotide sequence set forth in SEQ ID NO: 4.
8. The genetically engineered *R. toruloides* cell of any one of claims 2 to 7, wherein the *TEF1* promoter and the *GPD1* promoter are from *R. graminis*.
- 30 9. The genetically engineered *R. toruloides* cell of claim 8, wherein the promoters are from *R. graminis* WP1 strain.
10. An isolated directed-evolution-derived *R. toruloides* C3 strain having enhanced acetate utilization, decreased glycerol by-product accumulation, and stress resistance compared to wildtype *R. toruloides*, wherein said evolution was performed by adapting a culture of *R.*
35 *toruloides* against a gradual increase of acetate concentrations in culture medium.

11. The isolated directed-evolution-derived *R. toruloides* C3 strain of claim 10, wherein the evolved C3 strain can tolerate up to 6% acetic acid.
12. The isolated directed-evolution-derived *R. toruloides* C3 strain of claim 10, wherein the evolved C3 strain can tolerate up to 6% acetic acid and 4% glycerol.
- 5 13. The isolated directed-evolution-derived *R. toruloides* C3 strain of any one of claims 10 to 12, wherein the evolved C3 strain can produce higher α -linolenic acid (ALA) content compared to genetically engineered *CAT1* over-expressing C3 strain.
14. The isolated directed-evolution-derived *R. toruloides* C3 strain of any one of claims 10 to 13, wherein the evolved C3 strain is C3_64B.
- 10 15. An isolated *R. toruloides* C3 cell, wherein the cell has a genome contig sequence that is at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity, to the contig polynucleotide sequences set forth in SEQ ID NOs: 33-50.
16. A method of producing increased fatty acids, lipids and/or biomass by a *R. toruloides* cell,
15 comprising culturing a plurality of *R. toruloides* cells of any one of claims 1 to 15 in medium under conditions for fatty acid, lipid production and/or increased biomass, wherein the medium comprises acetate as carbon source.
17. The method according to claim 16, further comprising glycerol in the medium.
18. The method according to claim 16 or 17, wherein the *R. toruloides* cell is an *R. toruloides*
20 *CAT1* over-expressing C3 cell, more preferably a C3_64B cell.
19. The method according to any one of claims 15 to 18, wherein the increased fatty acids, lipids and/or biomass is relative to a strain of identical genome background or non-evolved *R. toruloides* cell.
20. The method according to any one of claims 15 to 19, wherein the increased fatty acid is
25 α -linolenic acid.
21. The use of an *R. toruloides* cell of any one of claims 1 to 14 in a method of producing enhanced fatty acid, lipid and/or biomass production compared to wildtype *R. toruloides*.
22. The use according to claim 21, wherein the *R. toruloides* cell is an *R. toruloides* C3 cell, preferably a *CAT1* over-expressing C3 cell, more preferably a C3_64B cell.